

THE ECONOMIC VALUE OF NATURAL LANGUAGE COMMUNICATION:
THEORY AND EXPERIMENTS

by

Siyu Wang
Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Economics

Committee:

_____ Director

_____ Department Chairperson

_____ Program Director

_____ Dean, College of Humanities
and Social Sciences

Date: _____ Spring Semester 2016
George Mason University
Fairfax, VA

The Economic Value Of Natural Language Communication: Theory And Experiments

A Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

by

Siyu Wang
Master of Arts
George Mason University, 2013

Director: Daniel Houser, Professor
Department of Economics

Spring Semester 2016
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my loving husband Garrett Harmon, my two wonderful parents Hong Wang and Hong Huo.

ACKNOWLEDGEMENTS

I would like to thank the many friends, relatives, and supporters who have made this happen. My loving husband, Garrett Harmon, assisted me in my research. My mother-in-law, Debbie Harmon, helped me with word processing and notes. Dr. Daniel Houser, and the other members of my committee were of invaluable help. Finally, thanks go out to the Fenwick Library for providing a clean, quiet, and well-equipped repository in which to work.

TABLE OF CONTENTS

| | Page |
|--|------|
| List of Tables | v |
| List of Figures | vi |
| List of Equations | vii |
| List of Abbreviations | viii |
| Abstract | ix |
| Chapter One: A Model of Multi-dimensional Intention Signaling | 1 |
| Section I. Introduction | 1 |
| Section II. Literature Review | 4 |
| Section III. Model | 9 |
| Section IV. Conclusion | 23 |
| Chapter Two: An Experimental Analysis On Natural Language Communication | 24 |
| Section I. Introduction | 24 |
| Section II. Experiment | 26 |
| Section III. Results | 30 |
| Section IV. Conclusion | 47 |
| Chapter Three: Advice and Social Learning: An Experimental Investigation | 50 |
| Section I. Introduction | 50 |
| Section II. Literature Review | 52 |
| Section III. The Experiment: Design and Procedure | 55 |
| Section IV. Hypothesis | 59 |
| Section IV. Discussion | 60 |
| Appendix | 62 |
| References | 77 |

LIST OF TABLES

| Table | Page |
|---|------|
| Table 1. Number of Subjects per Treatment Cross Subjects Pools | 29 |
| Table 2. Demographics of Subjects in Two Pools | 30 |
| Table 3. Pairwise Comparison of Coordination Rate | 31 |
| Table 4. Comparison of Coordination Rate with Single-gender and Mixed-gender Group | 33 |
| Table 5. Example of Messages in Each Category..... | 35 |
| Table 6. Effect of Communication on Coordination Rate | 40 |
| Table 7. Treatments of Social Learning Experiment | 55 |
| Table 8. Rotation Scheme For Subjects | 58 |

LIST OF FIGURES

| Figure | Page |
|--|------|
| Figure 1. Pink-Blue Game | 19 |
| Figure 2. Response to Two-dimensional Communication in Pink-Blue Game | 22 |
| Figure 3. Instruction Table..... | 26 |
| Figure 4. An Overview of Coordination Rate..... | 31 |
| Figure 5. Coordination Rate with Single-gender and Mixed-gender Groups | 33 |
| Figure 6. Number of Messages Classified as Each Communication Style (By Culture) . | 36 |
| Figure 7. Fraction of Players Choose the Color They Signaled | 37 |
| Figure 8. Coordination Rate in Three Communication Cases | 42 |
| Figure 9. Communication by Matching Type and University | 43 |

LIST OF EQUATIONS

| Equation | Page |
|------------------|------|
| Equation 1 | 45 |
| Equation 2 | 45 |

LIST OF ABBREVIATIONS

| | |
|--|-----------------------------|
| One-dimensional Restricted Communication | 1D Restricted Communication |
| Two-dimensional Restricted Communication | 2D Restricted Communication |

ABSTRACT

THE ECONOMIC VALUE OF NATURAL LANGUAGE COMMUNICATION: THEORY AND EXPERIMENTS

Siyu Wang, Ph.D.

George Mason University, 2016

Dissertation Director: Dr. Daniel Houser

Language is a powerful and complex human tool facilitating social and economic decisions. That the rich structure of language has evidently survived an evolutionary process has led some economists to argue for the importance of understanding its influence on human decision-making. Recently, an extensive experimental literature reveals, in relation to a well-designed predetermined language space, efficient economic outcomes emerge more readily when players can communicate using natural language. This dissertation investigates why natural language communication promotes efficient economic outcomes. In doing so, I bridge theory and experiments in a way that sheds light on the rich structure of language.

In Chapter I, I build a theoretical framework which predicts that, in order to coordinate, people both communicate with and respond to two key features of language: intentions and attitudes, where attitude indicates the strength of a message sender's desire to have her intention followed. I demonstrate that three types of equilibria coexist in this

environment: agreement equilibrium, negotiation equilibrium and communication failure equilibrium. Overall, I demonstrate that the use of intentions and attitudes in natural language communication significantly improves coordination from one-dimensional intention signaling.

In Chapter II, I use controlled laboratory experimentation in both the United States and China to test the predictions of the model detailed in Chapter I. I find (i) natural language messages do include both signaled intentions and attitudes; (ii) people respond both to intentions and attitudes when making decisions; and (iii) the use of attitude significantly improves coordination. Moreover, while males and females recognize and respond to intentions and attitudes equally well, we find females are more likely to send demanding signals, while males are more likely to send messages focused on intentions.

In Chapter III, I investigate why advice often promotes social learning better than observing history of past actions and corresponding outcomes. For example, Schotter (2003) documents several experiments which show that advice is more effective in shaping the decisions that people make and tends to push those decisions in the direction of the predictions of rational theory. I ask, why does advice impact decisions differently than observations? Inspired by Manski (2004), I hypothesize that advice is different because it provides information about past generations' expectations regarding counterfactual outcomes. To test this I designed an experiment using intergenerational games. In the experiment, players can learn from others by observing the history of decisions and outcomes, receiving Recommendation Only advice or Full advice (include both recommendation and justification statements). I hypothesize that the written

justifications include not only chosen actions and corresponding outcomes, but also counterfactuals related to expected outcomes of actions not chosen. Consequently, the Full Advice treatment enhances social learning and improves social outcome in relation to the Recommendation Only and History treatments.

CHAPTER ONE: A MODEL OF MULTI-DIMENSIONAL INTENTION SIGNALING

Section I. Introduction

Language is a powerful and complex human tool facilitating social and economic decisions. Over the last decades there has been significant theoretical and empirical progress towards understanding how communication improves coordination. While much of this early progress has occurred within the context of intention signaling models, natural language communication has a more richly detailed structure that can convey multiple meanings concisely. That this structure has evidently survived an evolutionary process has led some economists to argue for its importance not only to understanding grammar (e.g. Selten and Warglien 2007) but also for understanding human social and economic decisions (e.g. Rubinstein 2000). Farrell (1993) argues that richness of language should play a more prominent role in game theory. Recent empirical evidence suggests Farrell was correct. In relation to constrained signaling, efficient economic outcomes emerge more readily when players can communicate using rich natural language (see Charness and Dufwenberg 2006, 2010; Cooper and Kühn, 2014 and 2015; Cason and Mui 2015). Our aim here is to present and test a formal framework that predicts that people both use and respond to multi-meaning natural language in a way that improves coordination. In doing so, we bridge theory (Farrell 1993, Rabin 1994) and

experiments (Charness and Dufwenberg 2006, 2010; Cooper and Kühn, 2014 and 2015; Cason and Mui 2015) in a way that sheds light on why and how rich language promotes efficient outcomes in coordination environments.

To illustrate the importance of multi-meaning, rich natural language, consider routine discussions over where to dine or which film to view. If everyone conveys indifference, by stating that “it’s up to you,” or “I do not care,” the conversation is likely to be long and inefficient. At the same time, the same is true if all report a preferred option without a clear attitude. For example, one person says, “I feel like French,” and another responds, “I’m thinking Indian,” a different long conversation ensues. People typically solve these problems by making statements that combine intention signaling with attitudes,¹ that is, they indicate strength of desire to have their suggestions followed. For example, one person’s, “I really want French tonight” meeting another’s “I’m thinking Indian, but it’s up to you” will lead to immediate agreement. Intuitively, it makes good sense to communicate in a way that allows common ground to be rapidly found.

One can incorporate *attitudes* into an intention-signaling model with rich language that includes two-dimensional meanings. One dimension is *intention*, which signals the intended move in the game. The other dimension is *attitude*, which indicates the strength of desire to have the intentions followed. An *attitude* could be demanding, meaning that the message sender demands that the receiver follow the stated intentions.

¹ In the psychology literature, attitude is understood as one component of a broader model of behavior, which includes social norms and behavioral intentions. The way we define attitude in this paper is closely related to the behavioral intentions part of the concept. (See Osgood et al. 1957, Fishbein 1967c, Fishbein 1975.)

Alternatively, the message could be deferring, meaning that the senders offer to follow the receiver's stated intentions.

To formalize the idea, consider a two-stage game in which players send each other simultaneous cheap talk messages in the first stage and make simultaneous actions in the second stage. Suppose the language space players use to communicate is two-dimensional, common, and complete. That is, players understand each other's messages, each of which includes a signaled *intention* (so-called *E-meaning*; Rabin 1991, 1994) as well as an *attitude* (which we denote as *A-meaning*).

First, we apply the notion of *Credibility* from Farrell (1993)² to a communication environment with two-dimensional meanings to define an *agreement-equilibrium*, which requires that players act in accordance with the *intentions* of their messages when the action is a best response to the *intentions* of the partner's message.

Furthermore, we introduce the notion of *Attitude Ranking*, which posits that players indicating weaker desire to have their *intentions* followed (that is, those who are more willing to defer) cede decision authority to the player indicating stronger desire (that is, those who are more demanding). With this assumption, we demonstrate the existence of another group of equilibria, denoted *negotiated-equilibria*, which require that players act in accordance with the relatively more demanding player's *intentions* when paired players' messages differ in both intentions and *attitudes*. We show that *negotiated-equilibria*, in which one player concedes to the other player's requested equilibrium can happen in one-shot simultaneous communication.

² Farrell (1993) introduced *Credibility*, the intuition of which is that communication works well when there is no incentive to lie. Demichelis and Weibull (2008) include lexicographic preference for honesty, second to material preferences, to capture a similar idea.

With a simplified two-dimensional language including two *intentions* and two *attitudes*, communication failure can occur whenever players indicate differing *intentions* along with the same *attitude*. It is easy to see, however, that if the *attitude* is continuous (instead of binary) then communication failure occurs with probability approaching zero. The intuition is that in this case *attitudes* differ with probability approaching one, meaning they can surely be used to arbitrate disagreement over *intentions*. To the extent that time is costly, because communicating with *attitude* allows more rapid agreement, it is also in this sense efficient.

Economic theory has long conjectured that natural language communication facilitates efficient economic outcomes because it includes a rich, multi-meaning structure that people use and to which people respond (e.g. Marschack, 1965). To our knowledge, our paper is the first to apply the idea of multi-dimensional meaning to the context of coordination.

Section II. Literature Review

II.1. Intention Signaling

Both theory and experimental evidences suggest that the effectiveness of intention signaling in promoting coordination not only depends on the payoff structure of the underlying game, but also the structure of communication mechanisms (e.g. Cooper et al. 1989, 1992; Holt and Davis 1990; Van Huyck et al. 1993; Charness 2000; Clark et al. 2001; Charnes and Grosskopf, 2004; Duffy and Feltovich 2002, 2006; Blume and Ortman 2007). For instance, Cooper et al. (1989) found one-way communication (i.e. only one player can send a message) resolves coordination problems in the Battle of Sexes game,

while two-way simultaneous communication (both players can send messages) is of less value in solving 2-person coordination problems. Cooper et al. (1992) confirmed that two-way communication does not always decrease the frequency of coordination failure in different types of coordination games. Clark et al. (2001) tested Aumann's conjecture with two-way signals and showed, consistent with theory, informative communication does not necessarily lead to Pareto-efficient equilibrium outcomes.

Blume and Ortmann (2007) studied minimum and median games³ where multiple Pareto-ranked equilibria exist. In their experiment, all (nine) players send numerical messages simultaneously to indicate which of the (seven) options they intend to choose. Their results suggest that cheap talk facilitates coordination in games with a unique efficient equilibrium.

Choi and Lee (2014) studied the role of network structure of pre-play communication in determination of outcomes in coordination games. In addition to providing evidence that increasing the length of communication both improves the chance of successful coordination and reduces the asymmetry in the distribution of outcomes, they also discovered substantial variations in both efficiency and equity of coordination across networks.

Farrell (1987, 1988) took the important step of studying intention signaling in simple sequential games with complete information, where an underlying game is preceded by one or more rounds of structured pre-play communication in which players

³ Van Huyck, Battalio and Beil (1990) designed the so-called minimum and median games, where coordination among multiple players is required for efficient equilibrium. They showed that adding a pre-play auction enables players to coordinate on Pareto-efficient equilibrium.

make nonbinding announcements about their intended moves. Farrell suggested cheap talk need not be ignored, and moreover it could make equilibrium focal.

Rabin (1991, 1994) extended the analyses of Farrell (1987, 1988). In Rabin's model, players make repeated, simultaneous statements about their intended moves before they play a coordination game. He made it possible to combine the assumptions that players maximize expected payoffs given beliefs with a variety of behaviorally motivated restrictions on beliefs. He generalized Farrell's (1987) result to an unlimited number of rounds of communication.

Several studies have looked at the strategic aspects of intention signaling. Crawford (2003) allows for the possibility of bounded strategic rationality and rational players' responses to the misrepresentation of intentions. Demichelis and Weibull (2008) generalized cheap talk models by introducing a small cost of lying, and find this so-called lexicographic communication game is evolutionarily stable if and only if it results in a unique Pareto-efficient outcome. Ellingsen and Östling (2010) used a level-k model of strategic thinking to describe players' beliefs in games with one-way and two-way communication of intentions. They found that communication facilitates coordination in common interest games with positive spillovers and strategic complementarities.

II.2 Multidimensional Cheap Talk

In the last decade increasing effort has been made by theorists to investigate multidimensional signaling. The multidimensional messages in this group of literature were often sent from the informed player(s) to the uninformed player(s).

Battaglini (2002) proved that, contrary to the uni-dimensional case, with multidimensional cheap talk, full revelation of information in all states of nature is generically possible, even when the conflict of interest is arbitrarily large. Also, Levy and Razin (2007) analyzed the Crawford and Sobel (1982) cheap talk game with multidimensional states and found communication on one dimension can reveal information about the others. Finally, Chakraborty and Harbaugh (2010) showed that with multidimensional talk, experts with state-independent preferences can always make credible statements.

That multidimensional cheap talk can improve information transmission has also been confirmed in the laboratory. Lai, Lim and Wang (2015) designed experimental games that capture the logic of Battaglini (2002). They allow two senders to transmit information to a receiver over a 2-by-2 space. Their findings confirm that more information is extracted in a multidimensional setting. The extent to which information is transmitted depends on whether dimensional interests are aligned between a sender and the receiver, as well as the size of the message space. Vespa and Wilson (2015) implemented a multi-sender cheap talk model and showed that competing senders provide enough information for nearly full revelation. Battaglini and Makarov (2014) studied how behavior changes with the addition of an audience as well as from varying the degree of conflict between a sender's and a receiver's preferences. They found evidence that the addition of an audience alters the communication between sender and the receiver in a way consistent with the theoretical predictions: participants become more truthful as senders and more trusting as receivers. The multidimensional

communication in the literature mainly exists in the sender-receiver environments, which is distinct from the environment reported in this paper.

II.3. Natural Language Communication

Formal game theory modeling of natural language communication is limited. A closely related influential theoretical advance was provided by Farrell (1993), in which he carefully discusses the meaning of a message. He argues that a message may have meaning in one of three ways. First, a meaning may be established by use: Wittgenstein (1957) urged “Don’t ask the meaning; ask the use.” The meaning of messages that are used in equilibrium, for example, is established by Bayes’ rule, which tells us their meaning-in-use. Second, messages may have meanings that can be determined, or at least somewhat restricted, by introspection. This yields restrictions on out-of-equilibrium beliefs in generic signaling games, but not cheap talk games. Finally - and this is the key element absent in previous analyses - a message may have a focal meaning, especially if it is phrased in a preexisting language. Farrell (1993) uses rich language assumption to propose equilibrium refinements. It will become clear below that our model builds closely on this previous research by Farrell and Rabin.

Several experimental papers have suggested that free-form communication works better than restricted-form in facilitating efficient outcomes. Charness and Dufwenberg (2006) studied the impact of free-form communication on trust and trustworthiness. Charness and Dufwenberg (2010) conducted the games with identical designs except players were only given two restricted messages to choose from, One stated “I promise to choose Roll” (“bare promises”), while the other was blank. The impact on trust and

trustworthiness was weaker with restricted as compared to free-form communication. In Cason and Mui's (2015) experiment, a leader can decide whether to extract surplus from a victim and shares it with a beneficiary. They found that the successful joint resistance rate increases from 15% to 58% when moving from more restricted communication to rich communication. Cooper and Kühn (2014) explored how communication fosters cooperation.⁴ They find free-form chat improves cooperation much more than limited communication. Also, Cooper and Kühn (2015) investigates whether making it easier for subjects to reach an agreement by allowing more rounds of restricted communication can increase cooperation. Their results indicate that restricted communication is not a good substitute for the use of chat even after allowing subjects to have more rounds of interaction.

In broad brush-stroke, one message of the empirical literature is that the efficiency of economic outcomes is surprisingly high in the presence of pre-play communication, and more so when communication includes natural language. The question left unanswered is: why? The evidence we are going to show in this paper, which points to the importance of richly-structured multi-meaning free-form messages, is another step towards answering this question.

Section III. Model

Two individuals are playing a two-stage game $G^*(C, \mathcal{G})$ where they can communicate in the first stage C and play a coordination game in the second stage \mathcal{G} . We

⁴ Also called collusion in Cooper and Kühn (2014) context.

call $G^*(C, \mathcal{G})$ an extended game of G if and only if players play G in stage \mathcal{G} , and correspondingly the game G is called the underlying game of $G^*(C, \mathcal{G})$.

We focus on the case where the underlying game G is a symmetric game. Let the finite set A include all pure strategies a_i of game G . Let the set S include all the mixed strategy s_i , where $s_i \in \Delta(A)$ is a randomization over the set A of pure strategies. $U_i(a_i, a_j)$ is player i 's payoff as a function of pure strategies and $\pi_i(s_i, s_j)$ is the expected payoff derived from $U_i(\cdot)$, $i, j \in \{1, 2\}, j \neq i$. The set \mathbf{E} contains the Nash equilibria of game G . Nash equilibria are denoted as E_1, \dots, E_K and the strategies that constitute equilibrium $E_k \in \mathbf{E}$ are denoted as $(s_i^{E_k}, s_j^{E_k})$. Let \mathbf{E} be a subset of \mathbf{E} containing Pareto-efficient Nash equilibria, denoted by $\mathcal{E}_1, \dots, \mathcal{E}_K$. The strategies that constitute a Pareto-efficient equilibrium $\mathcal{E}_k \in \mathbf{E}$ are denoted by $(s_i^{\mathcal{E}_k}, s_j^{\mathcal{E}_k})$.

In the first stage C , each player simultaneously sends a cheap talk message $m_i \in M$, where M is a publicly known nonempty message set. By the definition of cheap talk, messages in M are non-binding, non-verifiable and costless. A function $\varphi: M \times M \rightarrow S$ specifies the second stage strategy $s_i = \varphi_i(m_i, m_j)$ given player i sends message m_i and receives m_j , $i \in \{1, 2\}, j \neq i$.

Combining both stages, the finite set A^* contains the pure strategies of $G^*(C, \mathcal{G})$, $a_i^* = (m_i, \varphi_i(m_i, m_j))$. A mixed strategy $\sigma_i \in \Delta(A^*)$ defines the probability assigned to pure strategy A^* . $V_i(a_i^*, a_j^*)$ is player i 's payoff as a function of pure strategies and $\Pi_i(\sigma_i, \sigma_j)$ is the expected payoff derived from $V_i(\cdot)$. We have

$\Pi_i(\sigma_i, \sigma_j) = \pi_i(\varphi_i(m_i, m_j), \varphi_j(m_j, m_i))$ given talk is costless. Furthermore, because communication is non-binding, any strategy profile (σ_i, σ_j) which constitutes a Nash equilibrium of G in the second stage \mathcal{G} is a Nash equilibrium of $G^*(C, \mathcal{G})$, which we denote as $(\sigma_i^{E_k}, \sigma_j^{E_k})$. Similarly, any strategy profile (σ_i, σ_j) which constitutes a Pareto-efficient Nash equilibrium of G in the second stage \mathcal{G} is a Pareto-efficient Nash equilibrium of $G^*(C, \mathcal{G})$, which we denote as $(\sigma_i^{\mathcal{E}_k}, \sigma_j^{\mathcal{E}_k})$.

III.1. Language Space

We assume players share a common language space M , which consists of (1) a meaningful vocabulary; and (2) a shared understanding among players that it is appropriate to interpret statements according to their literal meaning.⁵ We focus on the focal meaning of messages, a concept first suggested by Farrell (1993).⁶ We extend Farrell and Rabin's analysis⁷ to formally discuss the meaning of messages in two dimensions and add another restriction, *Attitude Ranking*. We start by defining the *E-meaning*⁸ (*equilibrium*) and *A-meaning* (*attitude*) of messages as follows:

DEFINITION 1: *The E-meaning refers to the equilibrium where the player signals her intention to play. We denote the set of messages that jointly signal Nash equilibrium E_k as $Q(E_k)$, $\forall k \in \{1, \dots, K\}$.*

⁵ It is important to assume players communicate with common language not only because it eliminates the inessential multiplicity of equilibria in cheap talk games, but also because it restricts the plausible interpretations of out-of-equilibrium messages.

⁶ In Farrell (1993), the meaning of messages are expanded to a rich pre-existing common language set, where the message's literal meaning is clear and only *Credibility* restricts communication.

⁷ Farrell and Rabin's analysis refers to Farrell (1993) and Rabin (1994).

⁸ Rabin (1994) defines a complete and common space for the *E-meaning* of language. We extend this idea to another dimension of language - attitude.

DEFINITION 2: *The A-meaning of a message indicates the strength of a player's desire to have her intention followed by others. We denote the set of messages conveying attitude \mathcal{A}_w as $D(\mathcal{A}_w)$, $\forall w \in \{1, \dots, W\}$.*

We denote \mathcal{A} as the set containing all possible *A-meanings* players can signal in the communication stage C. Suppose \mathcal{A} includes W elements, which are denoted as $\mathcal{A}_1, \dots, \mathcal{A}_W$ respectively. In the following analysis we concentrate on the cases where $W \geq 2$.

We continue by defining a complete language space both in *E-meaning* and *A-meaning*. First, denote by $M(G^*)$ the language space players use to communicate during the communication stage of game $G^*(C, \mathcal{G})$.

DEFINITION 3: *The common language space $M(G^*)$ is complete with respect to its E-meaning if and only if:*

- (1) *For all $m_i \in M$, there exists an equilibrium $E_k \in \mathbf{E}$ of G such that $m_i \in Q(E_k)$;*
- (2) *For any two equilibria $E_k \in \mathbf{E}, E_l \in \mathbf{E}$ and $k \neq l$, $Q(E_k) \cap Q(E_l) = \emptyset$.*

DEFINITION 4: *The common language space $M(G^*)$ is complete with respect to its A-meaning if and only if:*

- (1) *For all $m_i \in M$, there exists attitude $\mathcal{A}_w \in \mathcal{A}$ such that $m_i \in D(\mathcal{A}_w)$;*
- (2) *For all attitudes $\mathcal{A}_w \in \mathcal{A}, \mathcal{A}_y \in \mathcal{A}$ and $w \neq y$, $D(\mathcal{A}_w) \cap D(\mathcal{A}_y) = \emptyset$.*

Part (1) of Definition 3 says that all messages contain an *E-meaning*, and Part (2) illustrates that no message can simultaneously signal two different equilibria. Analogously, Definition 4 says that all messages convey an *A-meaning* and that no single message can convey more than one attitude.

In the development below, following Rabin (1994),⁹ we assume players' messages consist solely of proposals to play Pareto-efficient Nash equilibria.

III.2. Equilibria

III.2.1. Agreement-equilibrium

We now move to how players form an "agreement" in a two-dimensional language space. Rabin (1994) says players have reached an *agreement* when they mention the same equilibrium in the pre-play communication stage. We extend the notion of *agreement* into a language space with both *E-meaning* and *A-meaning*, and proceed to define the concept of an *agreement-equilibrium*.

ASSUMPTION 1 (Credibility): *If it is optimal for a player to honor the E-meaning of her own message given the other player honors the E-meaning of her own message, then the E-meanings of messages from both players will be honored.*

Roughly put, *Credibility* assumes players won't deviate from their words if there is no incentive for them to do so.¹⁰ For example, if paired players suggested the same equilibrium in their messages and that equilibrium is optimal for the both of them, then we assume that they will choose that equilibrium in the second stage.

⁹ This is referred to as "Pareto talk" by Rabin (1994). Rabin considers this one of the simplest examples of negotiating language, where it consists solely of proposals to play Pareto-efficient Nash equilibria.

¹⁰ Farrell (1988, 1993) summarizes the intuition of *Credibility* as "communication cannot work well when there are incentives to lie." Demichelis and Weibull (2008) add a lexicographic preference for honesty (small costs of lying) to their model to capture the similar idea.

PROPOSITION 1: *Paired players jointly play \mathcal{E}_k in the second stage, if in the first stage they made proposal $\mathbf{m} = (m_1, m_2)$, where $m_i \in Q(\mathcal{E}_k), i \in \{1, 2\}, \forall \mathcal{E}_k \in \mathcal{E}$ of underlying game G .*

If paired players signal a shared intention to play the same Pareto-efficient Nash equilibrium of underlying game G then, by the assumption of *Credibility*, they have an implicit agreement to play the signaled equilibrium. This type of equilibrium of G^* is called an agreement-equilibrium and can be characterized as follows:

DEFINITION 5: *A strategy profile (σ_1, σ_2) , where $\sigma_i \equiv (m_i, s_i), \forall i \in \{1, 2\}$, constitutes agreement-equilibrium of game $G^*(C, \mathcal{G})$ iff*

- (1) *Players make proposal set $\mathbf{m} = (m_1, m_2)$ in stage C , where $m_i \in Q(\mathcal{E}_k) \forall i \in \{1, 2\}$;*
- (2) *Players play $\mathcal{E}_k = (s_1^{\mathcal{E}_k}, s_2^{\mathcal{E}_k})$ in stage \mathcal{G} .*

III.2.2. Negotiated-equilibrium

We continue by considering how a message's *A-meaning* can be used to achieve equilibrium when paired players signal the intention to play different equilibria.

ASSUMPTION 2 (Attitude Ranking): *Suppose players share a complete and transitive ordering of the elements in \mathcal{A} . Given $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$ and $m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_y), i, j \in \{1, 2\}, j \neq i, k \neq l$, if \mathcal{A}_w indicates stronger desire of having the message followed than \mathcal{A}_y , then players will play \mathcal{E}_k in stage \mathcal{G} .*

We assume that players share a complete and transitive ordering for *A-meanings*. This requires that the players have the same interpretation of attitude with the preexisting common language. For example, we assume that everyone in the game agrees that the message “Let’s do X no matter what” expresses a stronger desire to have the message followed than a message like “It’s up to you, if you have no strong opinion, let’s do X.”

Assumption 2 says that if there are two different *E-meanings* and two different *A-meanings* in the two messages from paired players, then both players will honor the *E-meaning* of the message which includes the *A-meaning* indicating stronger desire to have one’s message followed. We are now in position to explain how simultaneous “negotiated agreement” can occur within a single round of simultaneous multi-way communication. Given *Attitude Ranking*, players are able to use *A-meanings* to resolve the conflicts between *E-meanings*.

DEFINITION 6: A strategy profile (σ_1, σ_2) , where $\sigma_i \equiv (m_i, s_i), \forall i \in \{1, 2\}$, constitutes negotiated-equilibrium of game $G^*(C, \mathcal{G})$ iff

- (1) Player i makes a proposal which satisfies $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$ in stage C , and player j makes a proposal m_j such that $m_j \notin Q(\mathcal{E}_k)$ but $m_j \in D(\mathcal{A}_y)$, where \mathcal{A}_w indicates stronger desire to the message followed than \mathcal{A}_y ;
- (2) Players play $\mathcal{E}_k = (s_1^{\mathcal{E}_k}, s_2^{\mathcal{E}_k})$ in the stage G .

III.2.3. Communication failure

We close the analysis of equilibria by describing the case where different *E-meanings* but the same *A-meaning* are contained in the messages. That is, neither the

conditions of *agreement-equilibrium* nor *negotiated-equilibrium* are achieved with pre-play communication.

ASSUMPTION 3 When $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$ and $m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_w)$, the message set $\mathbf{m} = (m_i, m_j)$ will be ignored, $i, j \in \{1, 2\}, i \neq j, \forall \mathcal{E}_k \in \mathcal{E}, \mathcal{E}_l \in \mathcal{E}$ and $\mathcal{E}_k \neq \mathcal{E}_l \forall \mathcal{A}_w \in \mathcal{A}$.

Assumption 3 says that when paired players' messages convey different *E-meanings* but the same *A-meaning*, since it is not possible to coordinate based on the message.

DEFINITION 7: A strategy profile (σ_1, σ_2) , where $\sigma_i \equiv (m_i, s_i)$, constitutes communication failure for game $G^*(C, \mathcal{G})$ if

- (1) For $\mathbf{m} = (m_1, m_2)$, $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w), m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_w), \forall \mathcal{E}_k \in \mathcal{E}, \mathcal{E}_l \in \mathcal{E}$ and $\mathcal{E}_k \neq \mathcal{E}_l \forall \mathcal{A}_w \in \mathcal{A}$.
- (2) Players play the game as if there is no communication stage C .

Communication failure does not necessarily lead to coordination failure. Definitions 5, 6 and 7 illustrate all possible message combinations that paired players are able to send in cheap talk environments where messages have two dimensions of meaning.

III.3. Communication Strategies

We now move to the analysis of communication strategies in the first stage. The assumptions we made allow us to restrict attention to only three different communication

situations, i.e. signaling same *E-meaning*, signaling different *E-meanings* with different *A-meanings*, and signaling different *E-meanings* with the same *A-meaning*. Suppose it is common knowledge that everyone responds to messages based on these assumptions. Given players' beliefs about what messages they might receive, there exist ex-ante best messaging strategies. However, the strategies highly depend on the payoff structure of the underlying coordination game. For ease of exposition, we concentrate on pure coordination games. The advantage is that the identical payoffs for two players isolate the function of cheap talk as a pure coordination device as compared to its use for any other information revealing purposes (e.g., indicating different marginal utility of money, or different social preferences or risk attitudes). Note it is straightforward to generalize this framework to other environments.

In a pure coordination game where players are indifferent among all Pareto-efficient Nash equilibria, it restricts attention to communication strategies meant to avoid communication failure. We denote the two pure-strategy Pareto-efficient equilibria as \mathcal{E}_k and \mathcal{E}_l , and one Pareto-dominated mixed-strategy equilibrium as \mathcal{E}_m . To simplify the characterization of equilibria, we assume there are exactly two attitudes available, \mathcal{A}_w and \mathcal{A}_y , where \mathcal{A}_w indicates stronger desire to the message followed than \mathcal{A}_y .

PROPOSITION 2: *With a pure coordination game as underlying game G , a player maximizes her expected payoff by sending message $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$, if she believes that she is least likely to receive message $m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_w)$, $i \in \{1,2\}$, $j \neq i$, $k \neq l$.*

Proposition 2 says that in order to maximize expected utility in a cheap talk game G^* with a pure coordination game with the equal payoff for two pure-strategy Nash equilibrium, one needs to minimize the chance of encountering \mathcal{E}_m by sending a message with a different *E-meaning* but the same *A-meaning* of the message one is least likely to receive.

The only time players cannot achieve a Pareto-efficient outcome is when they face *communication failure*. For each combination of *A-meaning* and *E-meaning* player j can send, there is a corresponding combination of *A-meaning* and *E-meaning* that leads to *communication failure*. Player i desires to minimize the chance of *communication failure* given her belief over the distribution of her counterparts' messages.¹¹

III.4. An Example: Pink-Blue Game

Our interest is in determining whether people use and respond to natural language messages in a way that is consistent with the model above. To do this, following Holm (2000),¹² we designed a "Pink-Blue" Game, where options are labeled as Pink and Blue, and counterparts' genders are revealed to players at the beginning. We use an announcement to ensure it is common knowledge that, for the purpose of this game, "Pink is a color often preferred by females, and Blue is a color often preferred by males." To the Holm (2000) environment, we added a pre-play free-form communication stage C, where players are able to send each other simultaneous messages. After the message

¹¹ Further, notice that the more categories of *attitude* there are, the less possible paired players encounter failure. If *attitude* is continuous, players could reach coordination with probability approaching one.

¹² Holm (2000) discusses the idea that introducing Pink and Blue labels into the gender-revealed coordination game could significantly improve coordination rate.

exchange, players make decisions in the Pink-Blue game, the payoff matrix for which is described by Figure 1.

| Player 1 \ Player 2 | | | |
|---------------------|------|------|------|
| | | Blue | Pink |
| | Blue | x, x | 0, 0 |
| | Pink | 0, 0 | x, x |

Figure 1. Pink-Blue Game

As indicated by Figure 1, players earn $x > 0$ if they choose the same color, and zero otherwise. We adopt the game to create two different situations, both when there exists a focal point (when players of the same gender are matched) and when there does not exist a focal point (when players of different genders are matched) before communication. We want to investigate how communication plays a role in each of those two cases. We intentionally focused on a pure coordination game so that social preferences do not complicate our analysis. In this environment the use and effect of communication is more readily discerned.

The Pink-Blue game has three Nash equilibria: two pure-strategy equilibria (Pink, Pink), (Blue, Blue) as well as a mixed-strategy equilibrium where each player plays each action with probability 0.5. Neither of the pure-strategy equilibrium Pareto-dominates the other, but both dominate the mixed strategy equilibrium. Given that Pink-Blue is a complete information game, players can use pre-play communication exclusively to make

claims about intended moves in order to break the symmetry (as compared to, for example, using this stage to reveal private information).¹³

In order to map the Pink-Blue game to the framework described above, we specify the language space with two possible Pareto-efficient Nash equilibria $\mathcal{E} = \{(\text{Blue}, \text{Blue}) (\text{Pink}, \text{Pink})\}$ as *E-meanings*, and two possible *A-meanings* $\mathcal{A} = \{\text{demanding}, \text{deferring}\}$, where *demanding* indicates stronger desire of “have the message followed” than *deferring*.

HYPOTHESIS 1: *Players communicate with one of four types of messages in game $G^*(C, \mathcal{G})$: Demanding Blue (DmB), Demanding Pink (DmP), Deferring Blue (DfB), Deferring Pink (DfP).*

DmB refers to the case where the sender intended to play Blue and requests the other player also play Blue; *DfB* refers to the case where the sender intended to play Blue, but is ultimately deferring the choice to the player, and analogously for *DmP* and *DfP*. Hypothesis 1 is built on Definitions 1-4. It implies that players in Pink-Blue Game communicate in this well-defined common and complete space with both *E-meaning* and *A-meaning*.

Based on Assumptions 1-3 and Definitions 5-7, we obtain the following three hypotheses and summarize all connections between messages and actions in Figure 2.

¹³ Farrell (1987, 1988) and Rabin (1991, 1994) provide formal analysis of communication as a means to convey intentions and thereby improve coordination among rational players. Farrell (1987, 1988) studies signaling intentions in simple sequential games of complete information where an underlying game is preceded by one or more rounds of structured pre-play communication. Rabin (1991, 1994) extends analysis of Farrell (1987, 1988) and studies coordination in a more general class of games with multiple rounds of communication.

When two players signal the same color, they can achieve an *agreement-equilibrium*. When two players signal different colors with different attitudes, they achieve a *negotiated-equilibrium*. Communication failure occurs when the two players signal different colors with the same Attitude. In this case, they ignore the messages and play a mixed-strategy Nash equilibrium.

HYPOTHESIS 2: *If paired players send messages signaling intentions to play the same color, they will choose that color.*

HYPOTHESIS 3: *If paired players send messages signaling intentions to play different colors with one Demanding and one Deferring attitude, both the demanding player and the deferring player will choose the color the demanding player signaled.*

HYPOTHESIS 4: *If paired players send messages signaling intentions to play different colors but with the same attitude, they will play mixed strategies with half chance choosing Blue, and half choosing Pink.*

Figure 2 illustrates the links between all 16 possible paired communication situations and response actions. In Pink-Blue game, a pair of players is guaranteed to coordinate on a Nash equilibrium in the second stage as long as both players respond to each communication situation in the first stage according to our model's three assumptions. However, one presumably wants to avoid Pareto-dominated outcomes that

result from communication failure. This can be done by optimally using *E-meaning* and *A-meaning* of messages based on beliefs over the distribution of message types.

| Player <i>i</i> \ Player <i>j</i> | | | | |
|-----------------------------------|--|--|--|--|
| | <i>DmB</i> | <i>DfB</i> | <i>DmP</i> | <i>DfP</i> |
| <i>DmB</i> | (Blue, Blue) | (Blue, Blue) | $((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$ | (Blue, Blue) |
| <i>DfB</i> | (Blue, Blue) | (Blue, Blue) | (Pink, Pink) | $((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$ |
| <i>DmP</i> | $((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$ | (Pink, Pink) | (Pink, Pink) | (Pink, Pink) |
| <i>DfP</i> | (Blue, Blue) | $((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$ | (Pink, Pink) | (Pink, Pink) |

Figure 2. Response to Two-dimensional Communication in Pink-Blue Game

HYPOTHESIS 5: *With a focal point of (Blue, Blue) in the second stage game, players will send messages with Blue E-meaning. With a focal point of (Pink, Pink) in the second stage game, players will send messages with Pink E-meaning.*

HYPOTHESIS 6: *Absent a focal point in the second stage game (male-female group), players will send messages including different E-meaning and same A-meaning of messages they believe least likely to receive.*

Hypothesis 6 is derived directly from Proposition 3. Hypothesis 5 is based on the fact that focal point is a solution that people will tend to use in the absence of communication, because it seems natural, special, or relevant to them.¹⁴ When adding a

¹⁴ See more at Schelling (1960), p. 57.

communication stage to games with natural focal points, and thus high rates of coordination in the absence of communication, the improvements communication can make may be limited. Players are nevertheless able to use communication as a way to confirm expectations. If both players are females, (Pink, Pink) is focal, while if both are males (Blue, Blue) is focal. With same-gender groups, the focal point is clear given we made an announcement about the color-gender relevance, the role of communication in this case is only to confirm. While in a mixed-gender group, communication can create a focal meaning. Once paired players agree on which equilibrium to play, there is no incentive to deviate.

Section IV. Conclusion

This paper provides an explanation for why natural language communication promotes coordination better than what's predicted by intention signaling model. Independent from intention signaling, we hypothesize that players also signal the strength of desire to have their intentions followed, which we call *attitudes*. We demonstrate that three types of equilibria coexist in this environment: agreement equilibrium, negotiation equilibrium and communication failure equilibrium. Overall, we demonstrate that the use of *intentions* and *attitudes* in natural language communication significantly improves coordination from one-dimensional intention signaling.

CHAPTER TWO: AN EXPERIMENTAL ANALYSIS ON NATURAL LANGUAGE COMMUNICATION

Section I. Introduction

In the laboratory, free-form communication has been shown to improve efficiency significantly in games including prisoners' dilemmas (Dawes, MacTavish, and Shaklee 1977), public goods (Isaac and Walker 1988), threshold public goods (Palfrey et al. 2015), as well as games of signaling (Cooper and Kagel 2005) and coordination (Brandts and Cooper 2007). Many studies implement pre-play cheap talk as free-form written messages, including both paper-pencil experiment and chat-box (e.g. Cooper and Kagel 2005; Xiao and Houser 2005, 2009, 2011; Charness and Dufwenberg 2006; Brandts and Cooper 2007; Schotter and Sopher 2007; Kimbrough et al. 2008; Ellingsen and Johannesson 2007; Sutter and Strassmair 2009; Heinnig Schmidt et al. 2008; Lundquist et al. 2009, Cason et al. 2012). Some studies also use face-to-face communication (Isaac, Ramey and Williams 1984; Daughety and Forsythe 1987a, 1987b; Binger et al. 1990; Valley et al. 1998).

More recently, a few experimental studies have documented that free-form communication works better than restricted-form in facilitating efficient outcomes (i.e. Charness and Dufwenberg 2006, 2010; Cooper and Kühn 2014, 2015; and Cason and Mui 2015). Charness and Dufwenberg (2006, 2010) studied the impact on trust and

trustworthiness was weaker with restricted language as compared to free-form communication. Cason and Mui's (2015) investigated the environment where a leader can decide whether to extract surplus from a victim and shares it with a beneficiary. They found that the successful joint resistance rate increases from 15% to 58% when moving from more restricted communication to rich communication. Cooper and Kühn (2014) discovered that free-form chat improves cooperation much more than limited communication.

This chapter investigates the reason why rich language promotes cooperation more than restricted intention signaling model with well-controlled laboratory experiments. We conducted pure coordination games (see Figure 1) with pre-played free-form communication, as well as without communication, with one-dimensional restricted communication and two-dimensional restricted communication as comparison. We tested the hypotheses documented in Chapter one. We are interested in studying how people take advantage of rich language space and how they respond to it.

Our results show that people naturally and without prompting communicate with both *intentions* and *attitudes*. Furthermore, we find people respond well to their counterparts' *intentions* and *attitudes* when making their decisions. Overall, the use of attitudes significantly improves coordination. We find that free-form communication promotes coordination as well as two-dimensional restricted communication, and is much better than no communication or one-dimensional restricted communication. In both free-form and two-dimensional communication treatments, people rapidly coordinate on *agreement-equilibria* and *negotiated-equilibria*. When communication failure occurs, the

coordination rate after free-form communication is higher than after two-dimensional restricted language. This result suggests that even though natural language communication can be well modeled by restricted language with two-dimensional meanings, attitudes might be more finely partitioned than the two categories in our model.

This experiment is the first empirical test of whether people naturally use multi-dimensional language when communicating, and whether people respond to multi-dimensional natural language in a way that promotes coordination.

Section II. Experiment

We conduct the Pink-Blue game discussed in Chapter One under four conditions: no communication, one-dimensional restricted communication, two-dimensional restricted communication and free-form communication.

| | | |
|--------|------|------------------|
| | You | Your Counterpart |
| Choice | Pink | Blue |
| Payoff | \$0 | \$0 |

| | | |
|--------|------|------------------|
| | You | Your Counterpart |
| Choice | Pink | Pink |
| Payoff | \$10 | \$10 |

| | | |
|--------|------|------------------|
| | You | Your Counterpart |
| Choice | Blue | Blue |
| Payoff | \$10 | \$10 |

| | | |
|--------|------|------------------|
| | You | Your Counterpart |
| Choice | Blue | Pink |
| Payoff | \$0 | \$0 |

Figure 3. Instruction Table

In baseline, the game proceeds as follows. We begin by randomly assigning people to pairs. We then reveal the gender of each player to the counterpart. Then the paired players simultaneously choose a color. If they choose the same color, then both earn the same amount of money in addition to their show-up payment. Otherwise, both earn only the show-up amount.

To avoid menu effects (for example, the possibility that participants will coordinate on the first listed option), we use a novel presentation without default “first option.” In particular, we use a square payoff figure (Figure 3) that ensures no option is more salient than any other.

To ensure color-labels were treated the same way across sessions and cultures, subjects were informed in all instructions that pink is a color often preferred by females, and blue is a color often preferred by males.

The three treatments with communication proceed just as in the baseline, with the exception that every subject has one chance to send a message to her counterpart simultaneously. Subjects are not able to read their counterparts’ notes until they have finished writing their own messages. Messages are written and delivered after the gender of the counterpart has been revealed but prior to subjects’ Pink/Blue choices. The details of the messages subjects can send in each communication treatment are summarized below.

II.1. Treatments

In the one-dimensional restricted communication treatment, subjects are informed they can choose one message from the following two to communicate with their counterparts: “I am choosing Pink” or “I am choosing Blue.”

In the two-dimensional restricted communication treatment, subjects are informed that each player can choose one message from the following four to communicate with their counterparts: “I am choosing Pink no matter what,” “I am choosing Blue no matter what,” “I am choosing the color your message tells me to. If your message says the same, I am choosing Pink.” or “I am choosing the color your message tells me to. If your message says the same, I am choosing Blue.”

In the free-form communication treatment, each subject is given a piece of blank paper. They are told they can write anything except potentially identifying information about themselves to their counterpart to communicate.

II.2. Procedures

We use a between-subject design, where each subject experiences exactly one treatment. In all treatments, each subject plays three rounds of the same game and faces different counterparts each round. At the beginning of each round, subjects are randomly and anonymously paired. The randomization is not constrained by gender, though most of our subjects experienced games with both genders. At the end of the experiment, one of the three rounds was randomly selected to determine subjects' payoffs.

During the experiment players are seated at separated computer terminals and are given a copy of the experiment's instructions. These instructions are also read aloud, as

we want to ensure subjects understand that the information contained in them is common knowledge.

II.3. Subject Pool

All four treatments have been conducted at George Mason University, Washington, D.C. The baseline and free-form communication treatment were also conducted in Shanghai Jiao Tong University, Shanghai. Subjects only interact with people from the same university. The main reason that we conducted the baseline and free-form communication treatment at two different locations is to study whether the two-dimensional structure only exists in English, or also applies to other languages. We conducted two extra treatments with one-dimensional restricted and two-dimensional restricted communication at George Mason University to test our framework of free-form communication. Table 1 lists the number of subjects in each treatment cross subjects pool.

Table 1. Number of Subjects per Treatment Cross Subjects Pools

| | Baseline | Free-form Communication | One-dimensional Restricted Communication | Two-dimensional Restricted Communication |
|------|----------|----------------------------|--|--|
| GMU | 42 | 52 | 58 | 56 |
| SJTU | 58 | 76 | N.A. | N.A. |

At Mason, subjects were paid \$5 for showing up and \$10 for coordination success. At SJTU, subjects were paid 10 RMB for showing up and 40 RMB for coordination success (1 US dollar \approx 6.14 RMB and 1 US dollar \approx 6.06 RMB when we

conducted the experiment in Shanghai, May 2013 and Dec. 2013). The experiment lasted about 40 minutes, and the average payoff was designed to be slightly above the local hourly wage for subjects from the two subject pools.

Each of the 342 subjects made 3 decisions for a total of 1,026 observations. Table 2 summarizes our subjects' demographics.

Table 2. Demographics of Subjects in Two Pools

| | George Mason Univ. Washington D.C., U.S. | Shanghai Jiao Tong Univ. Shanghai, China |
|---------------------|---|---|
| Number of Subjects | 208 | 134 |
| Percentage of Males | 49.04% | 64.18% |
| Average Age | 20.63 | 21.83 |
| Average GPA | 3.24 | 3.35 |

The gender ratio of our SJTU pool is unbalanced, which reflects the university's unbalanced gender ratio. This is likely due to the fact that SJTU is famous for its engineering department. About seventy percent of GMU subjects were born in the United States, and all of the SJTU subjects were born in China. The age and GPA of GMU and SJTU subjects are comparable to each other.

Section III. Results

III.1. Overview

We start the result section with a comparison of the coordination rates across all treatments. The overview results in this subsection only include GMU data, since SJTU subjects didn't participate in the two restricted communication treatments.

RESULT 1. *All three communication treatments facilitate coordination. Two-dimensional restricted communication facilitates coordination more than one-dimensional restricted communication. The free-form communication facilitates coordination as much as the two-dimensional restricted communication.*

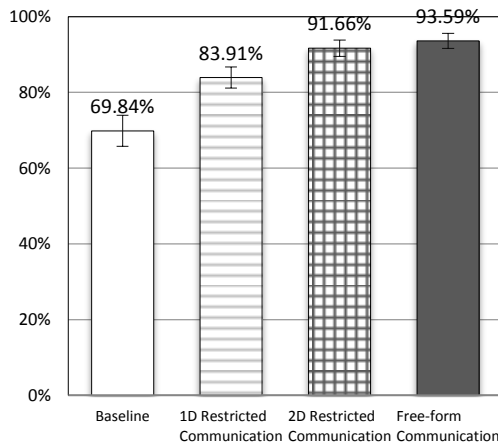


Table 3. Pairwise Comparison of Coordination Rate

| | Baseline | 1D Restricted Communication | 2D Restricted Communication |
|-----------------------------|-------------------|-----------------------------|-----------------------------|
| 1D Restricted Communication | 0.46*** (0.17) | | |
| 2D Restricted Communication | 0.45*** (0.09) | 0.48** (0.19) | |
| Free-form Communication | 0.32*** (0.07) | 0.24** (0.10) | -0.05 (0.28) |

Note: Probit Regression, with individual random effect. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses.

Figure 4. An Overview of Coordination Rate

Figure 4 depicts an overview of the coordination rate.¹⁵ Coordination is improved significantly in all three communication treatments compared to the baseline. Moreover, the coordination rate in the two-dimensional restricted communication treatment is significantly higher than the one-dimensional restricted communication treatment. The coordination rate in free-form communication is as high as two-dimensional restricted communication.

¹⁵ One-dimensional restricted communication is abbreviated as 1D Restricted Communication and two-dimensional restricted communication is abbreviated as 2D Restricted Communication in all the following figures.

The statistical support of Result 1 is summarized in Table 3, where we list pairwise comparisons of coordination rates between treatments. Each cell includes the coefficient and standard error (in parenthesis) from one regression. To accommodate the binary dependent variable and the issue of non-independence with respect to individuals, we used Probit models that control for random effects at the individual level. The dependent variable in the models is coordination success, which takes a value of 1 if two matched players choose the same option and takes a value of zero otherwise. The main independent variable is treatment. Additionally, we controlled for age, years of residence at the current location, GPA, and relationship status.

RESULT 2. *Coordination improvements from the one-dimensional to two-dimensional restricted communication treatment mainly occurred with mixed-gender groups rather than single-gender groups. The coordination rate in two-dimensional restricted communication treatment is not different from free-form communication both with mixed-gender and single-gender groups.*

Figure 5 breaks down the coordination rate into single-gender and mixed-gender pairs. With mixed-gender groups, two-dimensional restricted communication has significantly higher coordination success than one-dimensional restricted communication; no significant difference has been found in single-gender groups. These results indicate that the impact of a second language dimension is much stronger when there is no focal point. Furthermore, the coordination rate in the free-form communication treatment is not significantly different from the two-dimensional restricted communication both with

single-gender and mixed-gender groups. We'll provide further comparison between two-dimensional restricted and free-form communication in later sections. The statistical support of Result 2 is summarized in Table 4. Again, we used Probit models that control for random effects at the individual level with the same demographic variables as in the previous models (in Table 3).

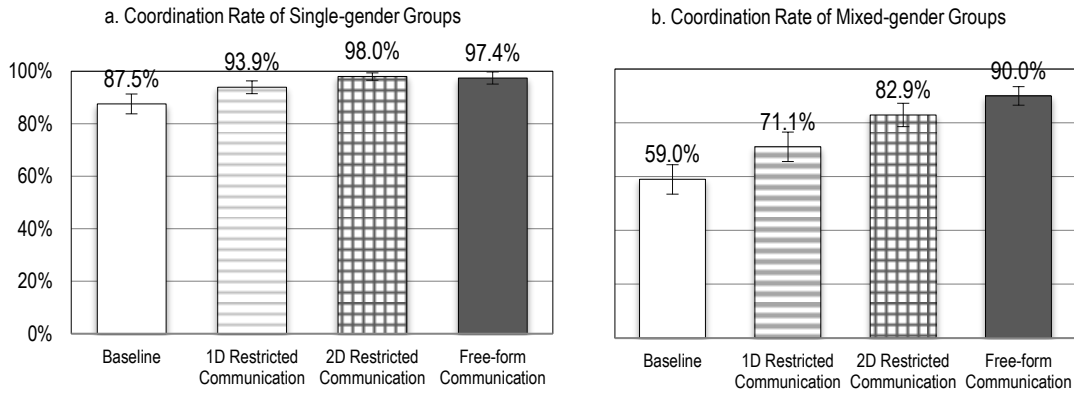


Figure 5. Coordination Rate with Single-gender and Mixed-gender Groups

Table 4. Comparison of Coordination Rate with Single-gender and Mixed-gender Group

| | Single-gender Groups | | Mixed-gender Groups | |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| | 1D Restricted Communication | 2D Restricted Communication | 1D Restricted Communication | 2D Restricted Communication |
| 2D Restricted Communication | 0.62 (0.39) | | 0.50** (0.25) | |
| Free-form Communication | 0.18 (0.19) | -0.39 (0.48) | 0.38** (0.16) | 0.16 (0.30) |

Note: Probit Regression, with individual random effect. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses.

III.2. Nature of Free-form Communication

In this section, we focus on the nature of free-form communication messages we collected. In order to analyze the content of

our messages, we used the Houser and Xiao (2011)¹⁶ procedure to classify messages. Thirty-two evaluators from SJTU and 37 from GMU were recruited outside of our experiment subject pool to categorize the free-form messages we collected. Each message was provided to evaluators both in its initial form and its Chinese (or English) translation and classified by these third party evaluators from both universities. Evaluators read messages and placed them into one of the following five categories: “Demanding Pink,” “Deferring Pink,” “Demanding Blue,” “Deferring Blue” and “None of Above”. The instructions laid out how evaluators should categorize messages: “[Y]ou should choose ‘Demanding Pink’ if you believe the message writer intends to play Pink, and requests that her or his counterpart also chooses Pink. You should choose ‘Deferring Pink’ if you believe the message writer intends to play Pink, but is ultimately deferring the choice to her or his counterpart.” Similar wording was used for “Demanding Blue” and “Deferring Blue.” Evaluators are paid based on whether their categorization of three randomly chosen messages is consistent with the most popular categorization for that session.

To provide a sense of the nature of messages in each category, we have provided one sample message from each category below (see all 384 messages and their classifications in the online Appendix).

¹⁶ While some previous literature on communication uses research assistants as evaluators, Houser and Xiao (2011) discuss the advantage of using a coordination game to classify messages. This classification method is increasingly used to classify messages from free-form communication studies (e.g. Xiao and Houser, 2005; Gächter et al. 2013; Deck et al. 2013; Ong et al. 2012; Chen et al. 2014).

Table 5. Example of Messages in Each Category

| Message | American or Chinese Evaluators | Classification (%) | | | | |
|---|--------------------------------|--------------------|------|------|------|------|
| | | DmB | DfB | DmP | DfP | NoA |
| "Pick Blue no matter what." | A | 94.6 | 2.7 | 0 | 0 | 2.7 |
| | C | 96.9 | 3.1 | 0 | 0 | 0 |
| "Hi, I will follow what you wrote on this paper. If you suggested PINK, I will choose pink, if you wrote Blue, I will choose Blue. If you didn't suggest any color, please choose Blue. " | A | 2.7 | 94.6 | 0 | 0 | 2.7 |
| | C | 0 | 96.9 | 3.1 | 0 | 0 |
| "Pink! Girl's Rule! No if ands or but! I will NOT change! 100% Pink!! I will not choose otherwise even if you 100% say different or tell me to change! I will NOT change! Go Pink!" | A | 0 | 2.7 | 94.6 | 0 | 2.7 |
| | C | 0 | 0 | 93.5 | 6.5 | 0 |
| "Going with whatever color you say. (If you don't mention a color I will go with Pink.)" | A | 2.7 | 2.7 | 0 | 91.9 | 2.7 |
| | C | 0 | 0 | 3.1 | 96.9 | 0 |
| "The point of coming to these experiments is to make money. I don't care about the colors and what-not. So let's try and choose the same option." | A | 0 | 2.7 | 0 | 0 | 97.3 |
| | C | 0 | 0 | 0 | 3.1 | 96.9 |

We used the classification of a message as one of the four categories in our model only if more than 50% of the evaluators from both pools place it in that same category. We classify a message as "None of Above" if more than half of evaluators from both pools believe it belongs to "None of above," or fewer than 50% of evaluators from either pool agree on any of the four named categories.

RESULT 3. *95.8% of messages are classified consistently by evaluators from the U.S. and China as communicating with one of the four types: Demanding Blue, Deferring Blue, Demanding Pink or Deferring Pink.*

Figure 6 displays the distribution of classifications of all the messages we collected from the two subject pools. The number of Blue messages is much higher than Pink messages in SJTU mainly due to the unbalanced gender ratio. Among the 384 free-form messages collected, only ten messages were categorized differently between Shanghai and D.C. subjects, and only six messages were not classified within one of the four named classifications. Thus, 95.83% of messages were classified consistently by evaluators from both pools as one of the four named communication types. Result 3 supports Hypothesis 1 (see Section 3.4). This finding confirms that when communicating in free-form, people naturally write messages in a two-dimensional space that include both signaled intentions and attitudes.

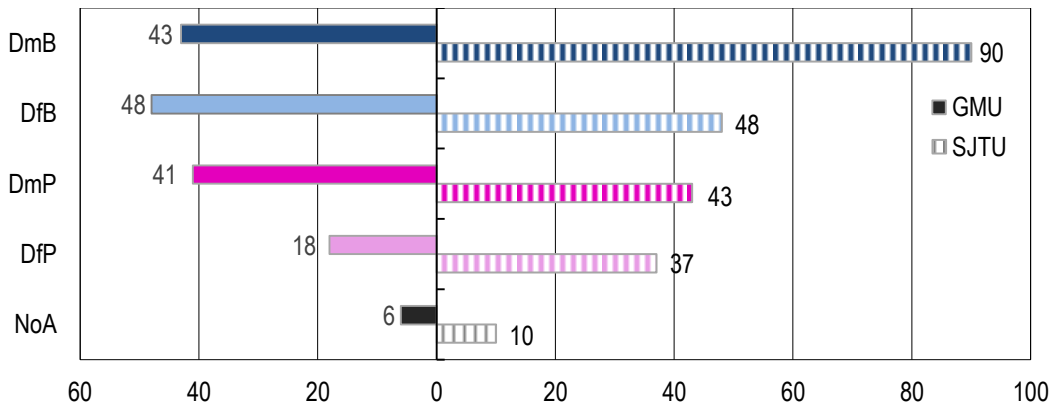


Figure 6. Number of Messages Classified as Each Communication Style (By Culture)

It is important to emphasize that messages were classified based on message content only. Evaluators were provided no further information. In particular, evaluators did not know which messages were paired and knew neither the gender nor the choices of the message writers. Because evaluations are independent of participants' choices, we are able to use the classified *E-meanings* and *A-meanings* to test whether people wrote and responded to messages as our theory predicts.

III.3. Responses to Free-form Messages

We move on to the analysis of how people respond to free-form messages. The following analysis only includes cases where both messages from paired players are classified as one of the four named types (i.e. 95.83% of the messages we collected).

RESULT 4. *When paired players send messages that signal the same color, 99.6% of them play the color they signaled.*

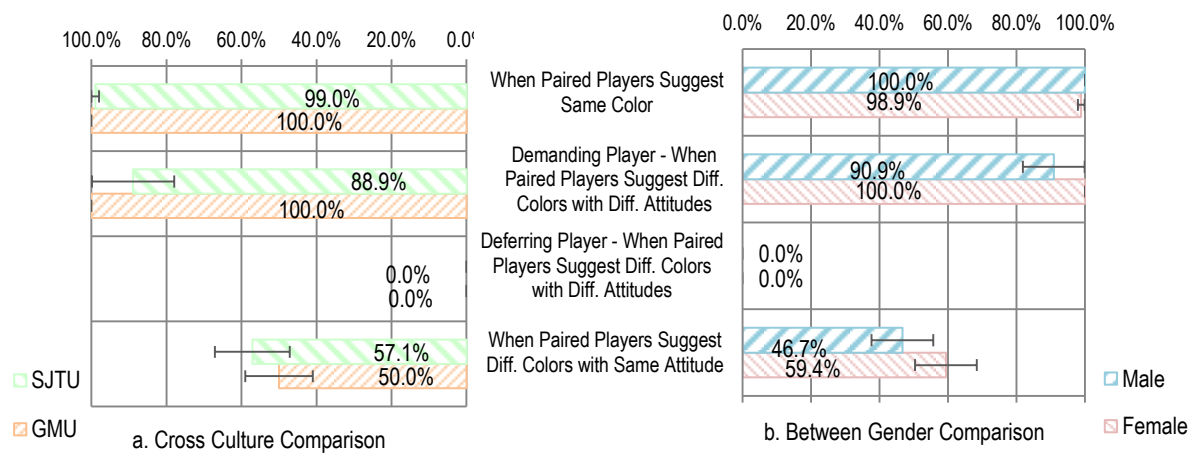


Figure 7. Fraction of Players Choose the Color They Signaled

Figure 7 presents the fraction of players who chose the color signaled in their messages both cross culture (7a) and between gender (7b). As shown in the first row, when paired players send messages that signal the same color, 99.6% of players play the signaled color. Consequently, players achieve agreement-equilibrium 99.2% of the time. Result 4 provides evidence to support Hypothesis 2. Furthermore, this percentage is not significantly different either cross cultures or between genders.¹⁷

RESULT 5. *When paired players send messages that signal different colors, with one player sending a demanding message and the other sending a deferring message, 95.7% of demanding players choose the color they signaled, while no deferring player chooses the color they signaled.*

As displayed in the second and third rows of Figure 7, when paired players sent messages that signal different colors with different attitudes, 95.7% of demanding players choose the color they signaled, and no deferring player chooses the color they signaled. As a result, 95.7% of paired players coordinate on the *negotiated-equilibrium*. Result 5 presents evidence that people take advantage of and respond to the attitudes of language, which supports Hypothesis 3. Neither cultural (shown in Figure 7a, Mann-Whitney U-test: $P=0.24$) nor gender differences have been found to be significant (shown in Figure 7b, Mann-Whitney U-test: $P=0.34$).¹⁸

¹⁷ Regarding the independence of the data included in our following non-parametric tests, only the first observation from each individual in each comparison category is included. For both cross culture and between gender comparisons, the rate of honoring one's stated color is approaching to 100% for each pool respectively. Thus, no P-value is provided here.

¹⁸ For both cross-culture and between gender comparisons, the P-values provided are regarding the demanding players' responses; for deferring players, all of them choose the colors signaled by demanding counterparts. Thus, no P-value can be provided.

RESULT 6. *When paired players send messages that signal different colors with the same attitude, 53.2% of players choose the color they signaled.*

As shown in the last row of Figure 7, when paired players sent messages that signal different colors with the same attitude, about 53.2% of players choose the color that they signaled. This percentage is consistent both across cultures (Mann-Whitney U-test: $P=0.66$) and between gender (Mann-Whitney U-test: $P=0.48$). Result 6 provides evidence to support Hypothesis 4.

RESULT 7. *When the conditions of an agreement equilibrium or a negotiated equilibrium are satisfied in the first stage, the coordination rate in the second stage is significantly higher than otherwise.*

Similar to Result 1 and Result 2, we provide statistical support for this result by conducting Probit regressions that control for random effects at the individual level. The dependent variable in the model is coordination success. The following independent variables are included in the regressions: a *negotiated equilibrium* condition dummy, which takes a value of 1 if the subjects have achieved the conditions for *negotiated equilibrium* in the first stage, and 0 otherwise; a *communication failure* condition dummy, which takes value of 1 if subjects face the condition of communication failure in the first stage, and 0 otherwise; a culture dummy, which takes a value of one for GMU subjects, and 0 for SJTU subjects; a gender dummy, which takes a value of 1 for male subjects, and 0 for female subjects; a gender counterpart dummy, which is equal to 1 for the

subjects who interact with male counterparts in the round, and 0 for the subjects who interact with female counterparts. Again, demographic variables are controlled in all models.

Model 1 of Table 6 shows that the coordination rate in the baseline of an *agreement equilibrium* condition is significantly higher than in communication failure cases, while not significantly different than the *negotiated equilibrium* cases. Model 2 and 3 of Table 6 are pairwise comparisons of when conditions of different equilibria are achieved. Model 2 only includes data of cases where the conditions of either a *negotiated equilibrium* or *communication failure equilibrium* have been achieved. Model 3 only includes data from cases when the conditions of either an *agreement equilibrium* or *communication failure equilibrium* have been achieved.

| Table 6. Effect of Communication on Coordination Rate | | | |
|--|--------------------|------------------|--------------------|
| Dependent Variable: Coordination Success | | | |
| Model | 1 | 2 | 3 |
| Negotiated Equilibrium Conditions | -0.67 (0.52) | | |
| Communication failure Conditions | -1.35*** (0.43) | -0.74* (0.43) | -1.35*** (0.47) |
| Culture | -0.48 (0.46) | -0.25 (0.49) | -0.01 (0.51) |
| Gender | 0.63 (0.44) | 1.12** (0.55) | 0.80 (0.50) |
| Gender of Counterpart | 0.55 (0.40) | 0.96* (0.50) | 0.72 (0.46) |
| Constant | 3.94 (2.97) | 5.05 (3.37) | 2.64 (3.21) |
| Number of Observations | 343 | 104 | 299 |

Note: Probit Regression, with individual random effect. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses.

The regressions show that (i) when the *agreement equilibrium* conditions are satisfied, the coordination rate is not higher than when the *negotiated equilibrium* conditions are satisfied; (ii) when either the *agreement equilibrium* conditions or the *negotiated conditions* are satisfied, the coordination rate is significantly higher than when the *communication failure* conditions are satisfied.

III.4. Two-dimensional Restricted and Free-form Communication

RESULT 8. When *the conditions of an agreement equilibrium or a negotiated equilibrium are satisfied in the first stage, the coordination rate in the second stage is not different between free-form communication and two-dimensional restricted communication. When the conditions of communication failure are satisfied, free-form communication facilitates coordination significantly more than two-dimensional restricted communication.*

Figure 8 breaks down the coordination rate of two-dimensional restricted and free-form communication into the following three different communication cases: paired players signal the same color, paired players signal different colors with different attitudes, and paired players signal different colors with the same attitudes. In the first case, the coordination rate does not differ between the two treatments (Mann-Whitney U-tests: $P=0.133$). When paired players sent messages that signal different colors with different *attitudes*, two-dimensional restricted communication is slightly better than free-form communication. (Mann-Whitney U-tests: $P=0.06$). When paired players sent

messages that signal different colors with the same *attitudes*, free-form communication works much better than two-dimensional restricted communication (Mann-Whitney U-tests: $P < 0.01$).

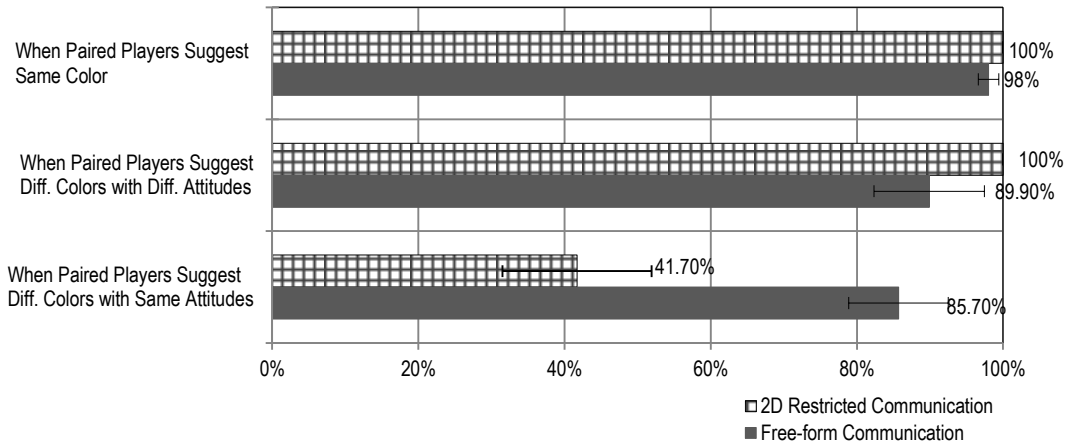


Figure 8. Coordination Rate in Three Communication Cases

One interesting observation is that when communication failure happens in the free-form communication treatment, our subjects successfully managed to coordinate 87.1% of the time with only about half of the players choosing the color they signaled (recall Result 6). However, when communicating with two-dimensional restricted messages, only 41.7% of the pairs managed to coordinate (see Figure 8). How did subjects manage to reach such a high coordination rate by communicating with free-form messages? This might be due to a likely feature of natural language, which is that *attitudes* are more finely grained than just two categories (i.e. demanding or deferring).

III.5. Communication Strategies

The communication styles players use may vary with different gender pairings and across cultures. The next three results summarize some basic findings of communication strategies and their impact on coordination rates.

RESULT 9. *Male-Male matched players send messages with DmB and DfB communication styles more frequently than DmP and DfP. Female-Female matched players send message with DmP and DfP communication styles more frequently than DmB and DfB.*

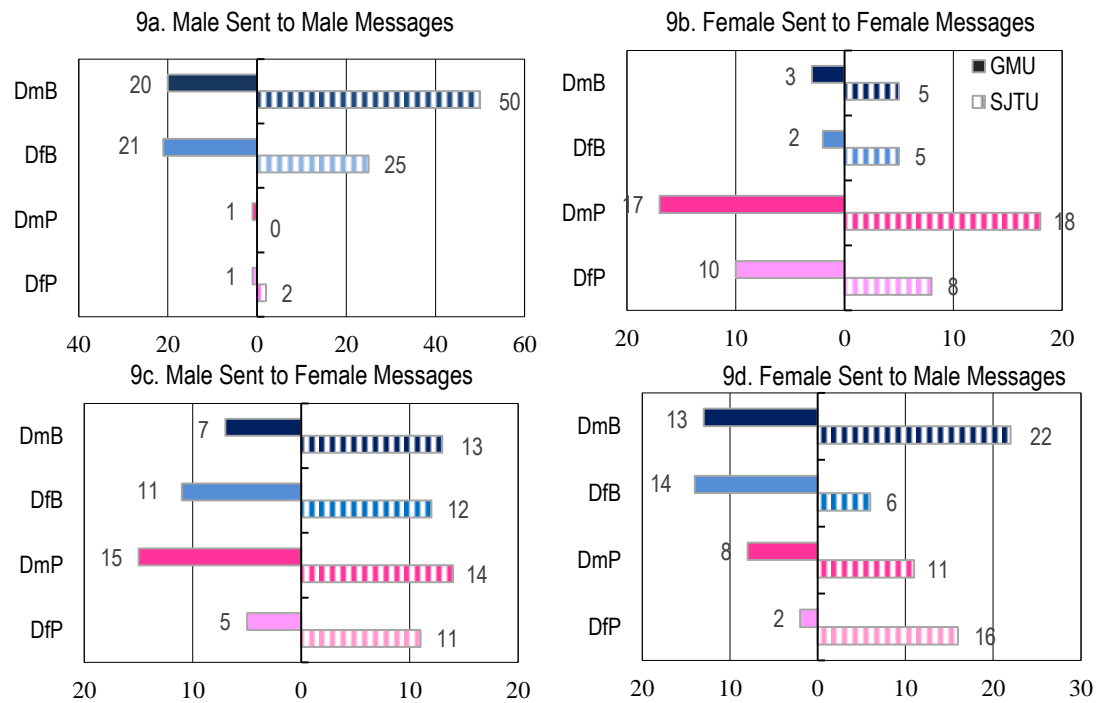


Figure 9. Communication by Matching Type and University

Figure 9 partitions the distribution of communication styles into the four possible pairings (Male-Male, Female-Female, Male-Female, Female-Male) and two cultures

(GMU, SJTU). As Figure 9a shows, only 4.7% of cases in GMU and 2.6% of cases in SJTU with Male-Male matching chose the communication style with Pink *E-meaning*. Slightly more Blue *E-meaning* messages are sent by females, but still much less than Pink *E-meaning* messages. 15.6 % of cases in GMU and 27.7 % of cases in SJTU Female-Female players chose communication styles with Blue *E-meaning* (see Figure 9b). Comparing part c and part d of Figure 9, one can observe some gender differences in communication styles in the mixed-gender environment, which we will address in our next result.

In order to investigate more about the personal differences of communication styles, we adopt a statistical classification procedure proposed by El-Gamal and Grether (1995). With this procedure, we are able to group all messages collected from each individual to investigate her communication styles. Based on the observation that some subjects always send the same style of message regardless of the gender of the counterpart, while the others vary the *E-meaning* or *A-meaning* according to their counterpart's gender, we assume in our statistical model that players in the game are comprised exclusively of five behavioral types chosen randomly from separate distributions. Each individual is assigned a positive prior probability for four “dogmatic” types and one sophisticated type.

Specifically, we restrict our attention to the following types of players: *dogmatic DmB* players, who always send a *DmB* message; *dogmatic DfB* players, who always send a *DfB* message; *dogmatic DmP* players, who always send a *DmP* message; *dogmatic DfP* players, who always send a *DfP* message; and *sophisticated* players, who send messages

with a focal point color as *E-meaning* if a focal point exists (same-gender pairings) but send messages with different *E-meanings* and the same *A-meaning* of the least popular communication style among the counterparts' population if there is no focal point (mixed-gender pairings).

For each subject i , we calculate a sequence of communication style statistics $x_1^i, \dots, x_{T_i}^i$, for each trial based on the classification of her messages and the counterpart's gender. $x_{cs^h, \tau}^i \in \{x_{dDmB, \tau}^i, x_{dDfB, \tau}^i, x_{dDmP, \tau}^i, x_{dDfP, \tau}^i, x_{soph, \tau}^i\}$, where $x_{cs^h, \tau}^i$ equals 1 if the subject i 's decision on trial τ agrees with rule h , and zero otherwise. We observe T_i decisions for a subject i . Now we define the sufficient statistic, $X_{cs^h}^i = \sum_{\tau=1}^{T_i} x_{cs^h, \tau}^i$ (the number of decisions that agree with rule h). Following El-Gamal and Grether (1995), we assume different subjects may use different rules, and the error rate is the same for all subjects and all tasks. We introduce the possibility that subjects make errors with probability ε . This allows each of our decision rules to give a positive probability (likelihood) to all possible patterns of behavior. The likelihood function for subject i is:

Equation 1

$$f^{cs, i}(x_1^i, \dots, x_{T_i}^i) = (1 - \frac{\varepsilon}{2})^{X_{cs}^i} \times (\frac{\varepsilon}{2})^{(T_i - X_{cs}^i)}$$

If there are I subjects, each indexed by i , then a natural way to estimate the communication style is to apply maximum likelihood using:

Equation 2

$$(\hat{cs}, \hat{\varepsilon}) = \underset{cs, \varepsilon}{argmax} \prod_{i=1}^I \prod_{h=1}^H (f^{cs^h, i}(x_1^i, \dots, x_{T_i}^i))^{\delta_{ih}}$$

where $\delta_{ih} \in \{0,1\}$ equals 1 if subject i is using rule h , and zero otherwise. In summary, we first calculate the maximum likelihood $f^{cs^h,i}$ of each communication rule cs^h for each individual i ; we choose the communication rule to maximize $f^{cs,i}$ for each individual i ; we multiply the obtained likelihood over individual $i \in \{1, \dots, I\}$, and maximize the outcome by choosing $(\hat{cs}, \hat{\varepsilon})$.

Turning to estimation results, Figure 10 summarizes the outcome of the El-Gamal and Grether (1995) procedure when applied to our data. Each bin bar represents the percentage of 128 subjects fit into the communication styles.

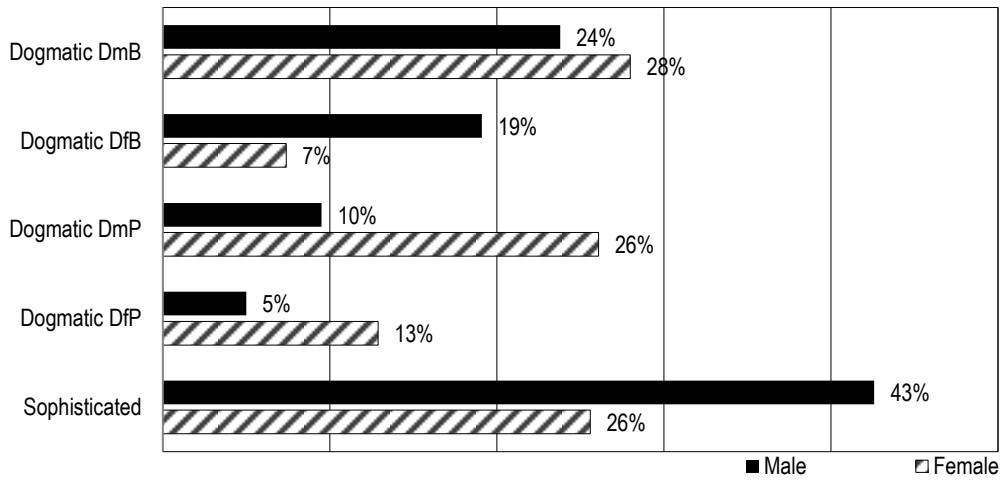


Figure 10. Communication Styles Classification by El-Gamal and Grether (1995) Method (By Gender)

RESULT 10. While a large proportion of players use sophisticated communication styles, males are more likely to be dogmatic senders who always send Blue messages, and females are more likely to be dogmatic senders who always send demanding messages.

Figure 10 shows that 43% of male players are sophisticated senders and 26% of females are sophisticated senders. Furthermore, while 43% of male players are either *dogmatic DmB* senders or *DfB* senders, only 15% of male players are *dogmatic Pink* message senders. 54% of females are *dogmatic demanding* senders while only 20% of females are dogmatic deferring senders.

Section IV. Conclusion

This paper provides evidence that independent from intention signaling, players also signal the strength of desire to have their intentions followed, which we call *attitudes*. Furthermore, people respond well to both *intentions* and *attitudes* when making their decisions. Overall, the use of *intentions* and *attitudes* in natural language communication significantly improves coordination from one-dimensional intention signaling. We also compare natural language communication with two-dimensional restricted communication, and no significant difference of coordination success has been found.

Use of a predetermined restricted message space might have methodological advantages, for example by allowing for clean tests of theoretical predictions and simplifying data analysis. However, consistent with the existing literature (Charness and Dufwenberg 2009; Cooper and Kühn 2014; Cason and Mui 2015), we find that restricted signaling has less rich structure than free-form communication.

We conclude by discussing three future research ideas that extend our study. The classification procedure we adopt includes a 2×2 message space: two *E-meanings* and two *A-meanings*. However, a likely feature of natural language is that *attitudes* are more

finely grained. For instance, two demanding messages may differ in that one is more demanding than the other. If so, players may be able to take advantage of a common understanding of more subtle distinctions in attitude to achieve coordination. It would be interesting in our setting to implement classification procedures which include more attitude categories. We speculate that doing so would reduce differences between restricted (predetermined) communication with two-dimensional meanings and free-form communication.

Our model departs from other intention signaling papers (e.g. Farrell 1987, 1988; Rabin 1991, 1994) principally in the nature of the multidimensional language space. We intentionally focused on a pure coordination game so that social preferences do not complicate our analysis. In this environment the use and effect of communication is more readily discerned. Of course, many natural environments of interest cannot be modeled as pure coordination games, and social preferences do play a ubiquitous role in determining economic choices. An important example is deception when players may benefit from misleading other participants. The role of multi-meaning messages in environments that include the possibility of strategic signaling, especially when messages can be “deceptive” or “partially deceptive,” is an important open question for future investigation.

In relation to restricted language, natural language communication might better enhance group identity. This might then promote lie-aversion or guilt-aversion, and might also improve one’s understanding of the connections between one’s decisions and one’s

outcomes.¹⁹ In our experiment, group-identity likely plays a rather small role given that one-shot simultaneous communication occurs between anonymously paired subjects, with neither feedback nor face-to-face interaction. Understanding how rich natural language enhances group-identity would be particularly useful.

¹⁹ See Frey and Bohnet (1995), Valley et al. (1998), Mohlin and Johannesson (2008), Cooper and Kühn (2015 *Working Paper*).

CHAPTER THREE: ADVICE AND SOCIAL LEARNING: AN EXPERIMENTAL INVESTIGATION²⁰

Section I. Introduction

Social learning is inherent to human existence and has a strong impact on shaping our economic and social decisions. This social learning often takes the form of friendly advice. For example, tourists researching travel to new countries can consult friends for accommodation and attraction recommendations. College graduates deciding on whether to take a particular job can ask directly for their friends' and professors' advice and suggestions. However, individuals can also learn by directly observing the experiences of others which can be done without asking for advice. For instance, the same tourists seeking to research their next trip can also make decisions based on the trips they see their friends post on Facebook without needing to ask a single question. College graduates can also evaluate jobs by observing the career outcomes of others that have previously worked in the same position. Word-of-mouth advice and historical observation of other's experiences are two key mechanisms through which social learning can take place. Historical observation in this context is defined as observing the actions and outcomes of past cohorts while word-of-mouth advice can be thought of as

²⁰ We would like to thank National Science Foundation (Grant No. 203599) for the support of this research.

directly communicating with individuals that have direct first-hand experience in the subject of interest.

While both historical observation and word-of-mouth advice are crucial channels of social learning, recent studies show that advice is a more effective force in shaping the decisions that people make and tends to push those decisions in the direction of the predictions of rational theory (e.g. Schotter 2003, Schotter and Sopher 2002, 2006, 2007, Chaudhuri, Schotter and Sopher 2009, Celen, Kariv and Schoter 2010). Researchers find it truly puzzling that subjects choose to follow the advice of other more frequently than merely just copying their predecessor's actions. As a result, individuals achieve a better social outcome despite the fact that advice is generally as informative as historical observation. They name this phenomenon "Advice Puzzle" or "Advice Paradox."

This paper investigates why advice often promotes social learning better than observing the history of past actions and corresponding outcomes. Inspired by Manski (2004), I hypothesize that advice is more influential and effective than historical observation because it provides information about past generations' expectations of counterfactuals. In other words, advice may in certain contexts help to address the selection problem in dynamic social learning processes which history cannot solve.

The selection problem is no less of an impediment to social learning. The problem is that only the outcomes of chosen actions are observable; one cannot observe the outcomes that would have occurred if persons had selected other actions. The logical impossibility of observing counterfactual outcomes has long been recognized to pose a fundamental difficulty for empirical research in the social sciences. As Manski (2004)

noted, “econometricians have long studied how identification problems limit the conclusions that can be drawn in empirical research. Decision makers attempting to interpret the experiences of past cohorts face much the same identification problem as do empirical researchers.”

The impossibility of observing counterfactuals is not only a challenge for empirical economists but also an obstacle for everyone trying to learn from the experience of past cohorts. This problem is especially prominent if one can only learn by observing the history. However, advice is endowed with the potential to alleviate this bias simply because advice can contain the information about expected counterfactual outcomes. To test this I designed an experiment using intergenerational games. In the experiment, players can learn from past cohorts by observing the history of decisions and outcomes, receiving Recommendation Only advice or Full advice (including both recommendation and written justification statements). I hypothesize that the written justification statements contain not only chosen actions and corresponding outcomes, but also counterfactuals related to expected outcomes of the actions not chosen. Consequently, the treatment with Full Advice enhances social learning more and achieves better social outcome.

Section II. Literature Review

In the last decade, several experiments have been conducted to compare the effects between observing the history and receiving the advice on economic and social outcomes in various environments. Subjects engage in what are called "intergenerational games." In these games a sequence of non-overlapping "generations" of players play a

stage game for a finite number of periods and are then replaced by other players who continue the game in their role for an identical length of time. Players in generation t are either observed by their successors in generation $t + 1$ or advise them on how they should behave.

In the majority of the intergenerational games, advice was designed to be composed of two parts: (i) A recommended action, which is a suggestion of action to choose, and (ii) a free-form statement offering a justification for the proposed action. For example, *Schotter and Sopher (2007)* conducted an intergenerational Ultimatum Game. In this paper, subjects play an Ultimatum Game with a \$10 endowment where each generation plays once and only once before it is retired. Between generations, subjects can learn either by advice, history or both of them. Their results suggest that advice was followed in a very direct way and had a significant impact on behavior. *Schotter and Sopher (2006)* investigated the development of conventions of trust in intergenerational games using the trust game of *Berg et al. (1995)* as the decision problem. They find that, in general, advice receivers tend to follow advice, and the advice they receive is trust-decreasing. *Schotter and Sopher (2003)* investigate intergenerational Battle of Sexes Games (BOSG) and found when a stage-game equilibrium state has been reached, regardless which one, subjects overwhelmingly tell their successors to adhere to it. Furthermore, their results show that word-of-mouth social learning is a far stronger force in the creation of social conventions than learning the lessons of history. Chaudhuri, Schotter and Sopher (2009) conducted an intergenerational Minimum Effort Game, which is a coordination game with weak strategic complementarities and Pareto-ranked

equilibria. They find that coordination is most likely result in the public advice treatment compared to history, private history and private history plus advice.

There are some experiments where advice contains a recommended action. In *Celen, Kariv and Schotter (2010)*, the experiments are designed so that both action and advice are identically informative in equilibrium. They introduced social learning to Celen and Kariv (2005)'s experiment. Basically, in this game eight subjects sequentially choose A or B. Decision A was the correct choice if the sum of eight private signals (uniform distributed over $[-10, 10]$) was positive, Decision B was correctly if the sum of eight private signals was negative. They find that despite the informational equivalence of advice and actions, subjects appear to be more willing to follow the advice given to them by their predecessor than to copy their action, and that the presence of advice increases subjects' welfare.

These intergenerational games have proven repeatedly that the value of advice is high. Generally speaking, people follow more closely to advice than history, and their behavior converts to more rational direction as a result of following advice. However, when economists test this idea, they found that surprisingly subjects do not realize the value of advice, at least the value is not reflected on the market. Nyarko and Schotter (2007) examine the market for advice in an investment game, where they decide on whether to choose a safe option or an investment option. Subjects can jointly discover the probability of high and low payoff in the investment option with social learning. They find subjects bid significantly higher for data than they do for advice.

In broad brush-stroke, advice is found to be more effective in shaping the decisions that people make and tends to push those decisions in the direction of the predictions of rational theory in various economic decision-making problems. The question left unanswered is: why? The experiment we are going to show in this paper, which investigates whether advice contains expectation of counterfactuals and the value of advice is improved because of it, is another step towards answering this question.

Section III. The Experiment: Design and Procedure

The general features of our intergenerational social learning are as follows: Subjects once recruited are ordered into generations. Each generation plays the game once and only once independently or with an opponent (depends on the game). After their participation in the game, subjects in any generation t are replaced by a next generation, $t + 1$. These features are largely consistent with the design of previous literature on intergenerational games (e.g. Schotter and Sopher 2007). As shown in Table 7, we have three treatments: History treatment, Recommendation Only Advice treatment, and Full Advice treatment.

Table 7. Treatments of Social Learning Experiment

| Treatment | History | Recommendation Only Advice | Full Advice (Recommendation+ Justification) |
|-------------|--|--|---|
| Description | Subject observes a complete history of the actions chosen by all of their previous predecessors. | Subjects receive advice from their immediate predecessors. The advice can only include a recommended action. | Subjects receive advice from their immediate predecessors. The advice can include a recommended action as well as written justification statements. |

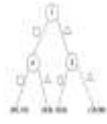


In the history treatment, subjects in generation t are able to observe a complete history of what the predecessors chose from generation 1 to $t-1$. In the Recommendation Only Advice treatment, subjects in generation t are able to give advice to their immediate successors (i.e. subjects in generation $t+1$) by suggesting a recommended action. In the Full Advice treatment, the subjects in generation t are able to their successors advice which contains two parts: (i) a recommended action; (ii) a free-form statement offering a justification for the recommended action. The payoffs to any subject in the experiment are equal to the payoffs earned during her lifetime plus a discounted payoff, which is 50% of the payoffs achieved by her immediate successor in the game.

Subjects engage in four separate intergenerational games: a Job Search Game, a Battle of Sexes Game, an Ultimatum Game and a Trust Game as defined by Berg et al. (1995). In the Job Search Game, subjects play basic search paradigms of job search tasks paradigm (see Lippman and McCall 1976). Subjects face two alternatives at each period: SEARCH or STOP. Every time they would choose SEARCH, the computer would then offer a wage by typing back “ w is my offer,” where w would be a random selected wage derived from a uniform wage distribution with a mean of 100 and a range from 0 to 200. Each search cost c points, c varied from trial to trial (always known by the searcher). Once subjects were satisfied by the wage offered after n times search, they could choose STOP and their final payoffs would be their accepted wage w minus $n*c$ points. In the Battle of Sexes Game, the game has two pure-strategy equilibria. In one (Triangle,

Triangle), Player 1 does relatively well and receives a payoff of 150, whereas player 2 does less well and receives a payoff of 50. In the other equilibrium, (Square, Square), just the opposite is true. In the disequilibrium all payoffs are zero. In the Ultimatum Game, subjects are randomly assigned to the role of Sender and Receiver. The Sender was initially allocated 100 units, which she can divide into two amounts, x and $100-x$. The amount x is proposed to the Receiver as her portion which the Receiver could either accept or reject. If the Receiver accepts the proposal, the payoffs would be x for the Receiver and $100-x$ for the Sender. If the Receiver rejected the proposal, each subject's payoff would be zero. In the Trust Game, subjects are randomly assigned to the role of Sender and Returner. The Sender is initially allocates 100 units. The task of the Sender is to send some amount $x \in [0, 100]$ to Returner. The amount x is then multiplied by 3 and given to Returner. The Returner could then either keep it all or send back any amount $y \in [0, 3x]$ to Sender. Payoffs are $\pi_S = 100 - x + y$ for the Sender and $\pi_R = 3y - x$ for the Returner.

The experiment had 8 generations. In each generation a subject would play one of the four games independently or with a different opponent (depends on which game they are playing). Consider the following table:

Table 8. Rotation Scheme For Subjects

| Games with Opponent | | | | | Game without Opponent | |
|---------------------|--------|---|---|--|-----------------------|-----------|
| | | Battle of Sexes Game | Ultimatum Game | Trust Game | Job Search Game | |
| | |  |  |  | | |
| Generation 1 | Row | Player 1 | Player 2 | Player 3 | Generation 1 | Player 4 |
| | Column | Player 8 | Player 7 | Player 6 | Generation 2 | Player 5 |
| Generation 2 | Row | Player 2 | Player 3 | Player 4 | Generation 3 | Player 8 |
| | Column | Player 6 | Player 5 | Player 1 | Generation 4 | Player 7 |
| Generation 3 | Row | Player 3 | Player 4 | Player 5 | Generation 5 | Player 1 |
| | Column | Player 7 | Player 8 | Player 2 | Generation 6 | Player 6 |
| Generation 4 | Row | Player 4 | Player 1 | Player 8 | Generation 7 | Player 3 |
| | Column | Player 5 | Player 6 | Player 7 | Generation 8 | Player 2 |
| Generation 5 | Row | Player 9 | Player 10 | Player 11 | Generation 1 | Player 12 |
| | Column | Player 16 | Player 15 | Player 14 | Generation 2 | Player 13 |
| Generation 6 | Row | Player 10 | Player 11 | Player 12 | Generation 3 | Player 16 |
| | Column | Player 14 | Player 13 | Player 9 | Generation 4 | Player 15 |
| Generation 7 | Row | Player 11 | Player 12 | Player 13 | Generation 5 | Player 9 |
| | Column | Player 15 | Player 16 | Player 10 | Generation 6 | Player 14 |
| Generation 8 | Row | Player 12 | Player 9 | Player 16 | Generation 7 | Player 11 |
| | Column | Player 13 | Player 14 | Player 15 | Generation 8 | Player 10 |

In this table we see 16 players performing our experiment in 8 generations. In Period 1, Player 1 and 8 play the Battle of Sexes Game while Players 2 and 7 play the Ultimatum Game, Players 3 and 6 play Trust Game and Player 4 and 5 play Job Search Game. When they have finished their respective games, we rotate them in the next period so that in Period 2, Player 2 and 6 Play the Battle of Sexes Game, Player 3 and 5 play Ultimatum Game, Player 4 and 1 play Trust Game and Player 8 and 7 play Job Search Game. Then the same type of rotation is carried out from Period 3 to the end of the

experiment. The rotation scheme is designed that subjects experience different opponent in different games and play and only play one generation in each of the games. The intergenerational game is intensively time-consuming since a subject cannot make decision until all of his predecessors have finished their decision-making. This design saves us and subjects time since at every given period, one-fourth of the subjects play the Battle of Sexes Game, one-fourth play the Ultimatum Game, one-fourth play Trust Game and one-fourth play the Job Search Game. Schotter and Sopher (2003, 2006 and 2007) also adopted a similar rotation scheme, the difference is our subjects play 4 different games instead of three.

Section IV. Hypothesis

Decision makers may consider all the outcomes of all the possible actions before they make decisions. However, people who wish to learn from these decision makers' experience cannot observe the outcomes that would have occurred if the decision makers had selected other actions. Advice has the potential to adjust this problem. Predecessors can provide information regarding their expected outcomes if they chose something. This information can fundamentally improve the social learning efficiency. In Hypothesis 1, we hypothesize that predecessors do talk about counterfactuals. In Hypothesis 2 and 3, we hypothesize that the advice, which includes counterfactuals, is more likely to be followed, and the successors make more rational decisions consequently.

HYPOTHESIS 1. Written justification statements include players' expectation of counterfactual outcomes.

HYPOTHESIS 2. The advice with written justification statements including counterfactual outcomes is more likely to be followed.

HYPOTHESIS 3. Players make decisions further in the direction of the predictions of rational theory when they receive written justification statements including counterfactual outcomes than when they receive written justification statements without counterfactual outcomes.

If the first three hypotheses are true, we can derive two more aggregate hypotheses: Hypothesis 4 and 5. If people do include counterfactuals in their intergenerational advice and it works, then the Full Advice treatment will be both more influential (advice better followed) and efficient (pushing the behavior towards more rational decision making outcome) than both the history treatment and Recommendation Only treatment.

HYPOTHESIS 4. The advice in the Full Advice treatment is followed more closely than Recommendation Only Advice and History treatments.

HYPOTHESIS 5. The decisions in the Full Advice treatment are pushed more in the direction of the predictions of rational theory than Recommendation Only advice and History treatments.

Section IV. Discussion

We extend the paradigm to better understand how different channels of social learning impact the efficiency of learning and how this then influences economic

outcomes. Extensive literature reveals that economic decisions vary according to the nature of social learning in an economic environment (Schotter 2003, Schotter and Sopher 2003, 2006, 2007).

We hypothesize that in advice, people do talk about counterfactuals. Furthermore the advice, which includes counterfactuals, is more likely to be followed. Consequently, the successors make more rational decisions.

We are also interested to know whether communication styles will differ between genders and across cultures. The reason is that we live in a shrinking world where people are increasingly interested in fostering cross-cultural and between-gender professional partnerships. Understanding how to foster these new types of economic relationships effectively, and to ensure communication is productive, is perhaps of first-order importance in our natural environment.

As noted by Gibbons,²¹ “The spirit of cheap talk is that anything could be said, but formalizing this would require M (i.e. the set of messages can be sent) to be a very large set.” The investigation reported in the three chapters of my dissertation shed light on the fundamental question of how rich a message space should be in order to closely capture key features of communication. This dissertation provides a small step towards understanding the rich structure of language and how it impacts decision making. More work needs to be done both theoretically and experimentally in unraveling the rich structure of language in bargaining, coordination, information-transmission and other social and economic environments.

²¹ See Page 212, *Game Theory for Applied Economists*, 1992 by Princeton University press.

APPENDIX

APPENDIX A - PROOFS

PROOF OF PROPOSITION 1: $\forall \mathcal{E}_k \in \mathcal{E}$, by definition it is necessarily true that $\pi_i(s'_i, s_j^{\mathcal{E}_k}) \leq \pi_i(s_i^{\mathcal{E}_k}, s_j^{\mathcal{E}_k})$, $\forall s'_i \in S_i$ and $s'_i \neq s_i^{\mathcal{E}_k}, \forall i, j \in \{1, 2\}$, $j \neq i$. Given $m_i \in Q(\mathcal{E}_k)$ and the assumption of Credibility assumption (Assumption 1), it follows that Player i honors the E-meaning of her message by playing \mathcal{E}_k .

PROOF OF PROPOSITION 2: Let $r_{ly}^j, r_{lw}^j, r_{ky}^j, r_{kw}^j$ be player i 's belief about the likelihood they will receive messages $m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_y)$, $m_j \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_w)$, $m_j \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_y)$, $m_j \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$ respectively. Given the completeness, we have $r_{ly}^j + r_{lw}^j + r_{ky}^j + r_{kw}^j = 1$. Suppose \mathcal{A}_w indicates stronger desire to have the message followed than \mathcal{A}_y . For a pure coordination game with equal payoff for two pure-strategy, $\pi_i(\mathcal{E}_l) = \pi_j(\mathcal{E}_l) = \pi_i(\mathcal{E}_k) = \pi_j(\mathcal{E}_k) > \pi_i(\mathcal{E}_m) = \pi_j(\mathcal{E}_m)$. Suppose $\pi_i(\mathcal{E}_l) = X$ and $\pi_i(\mathcal{E}_m) = Y$, the expected payoff functions for four messaging strategies can be simplified as below:

$$E[\Pi_i(m_i \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_w), \varphi_i(m_i, \cdot))] = \pi_i(\mathcal{E}_l) \times (r_{ly}^j + r_{lw}^j + r_{ky}^j) + \pi_i(\mathcal{E}_m) \times r_{kw}^j = (X-Y)(r_{ly}^j + r_{lw}^j + r_{ky}^j) + Y = (X-Y)(1 - r_{kw}^j) + Y$$

$$E[\Pi_i(m_i \in Q(\mathcal{E}_l) \cap D(\mathcal{A}_y), \varphi_i(m_i, \cdot))] = \pi_i(\mathcal{E}_l) \times (r_{ly}^j + r_{lw}^j) + \pi_i(\mathcal{E}_k) \times r_{kw}^j + \pi_i(\mathcal{E}_m) \times r_{ky}^j = \pi_i(\mathcal{E}_l) \times (r_{ly}^j + r_{lw}^j + r_{kw}^j) + \pi_i(\mathcal{E}_m) \times r_{ky}^j = (X-Y)(r_{ly}^j + r_{lw}^j + r_{kw}^j) + Y = (X-Y)(1 - r_{ky}^j) + Y$$

$$E[\Pi_i(m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w), \varphi_i(m_i, \cdot))] = \pi_i(\mathcal{E}_k) \times (r_{ky}^j + r_{kw}^j + r_{ly}^j) + \pi_i(\mathcal{E}_m) \times r_{lw}^j = (X-Y)(r_{ky}^j + r_{kw}^j + r_{ly}^j) + Y = (X-Y)(1 - r_{lw}^j) + Y$$

$$E \left[\Pi_i \left(m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_y), \varphi_i(m_i, \cdot) \right) \right] = \pi_i(\mathcal{E}_k) \times (r_{ky}^j + r_{kw}^j) + \pi_i(\mathcal{E}_l) \times r_{lw}^j + \pi_i(\mathcal{E}_m) \times r_{ly}^j = \pi_i(\mathcal{E}_k) \times (r_{ky}^j + r_{kw}^j + r_{lw}^j) + \pi_i(\mathcal{E}_m) \times r_{ly}^j = (X-Y)(r_{ky}^j + r_{kw}^j + r_{lw}^j) + Y = (X-Y)(1 - r_{ly}^j) + Y$$

Without loss of generality, we assume $r_{lw}^j = \min\{r_{ly}^j, r_{lw}^j, r_{ky}^j, r_{kw}^j\}$. Then comparing the four expected utility functions, we have $E[\Pi_i(m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w), \varphi_i(m_i, \cdot))]$ with the highest expected payoff, thus $m_i \in Q(\mathcal{E}_k) \cap D(\mathcal{A}_w)$ is the best ex-ante messaging strategy.

APPENDIX B – INSTRUCTIONS

INSTRUCTIONS 1²²

Welcome to Our Experiment!

You have been assigned a participant number and it is on the note card in front of you. This number will be used throughout the study. Please inform us if you do not know or cannot read your participant number.

All payments will be made to you in cash privately at the end of the study today. In addition to \$5 for showing up on time, you can earn additional money based on your decisions and the decisions of your randomly matched counterparts.

No communication with other participants is allowed during this experiment. If you have any questions please raise your hand, and the experimenter will assist you. Today's experiment will last about 1 hour.

Your anonymity in this study is assured. Your name will never be collected or connected to any decision you make here today.

Please read these instructions carefully!

Instructions

There are **3 rounds** in this study. Each round consists of **1 decision**. Each round, your and your counterpart's payoff will be determined by your joint decisions. At the beginning of each round, you will be randomly matched with a new counterpart. At the end of the experiment, one round will be randomly selected to determine

²² Instructions 1 are used for communication Treatment, Free-form Communication Treatment, Restricted Communication with One-dimensional or Two-dimensional Meaning(s) Treatments in U.S.

your and your counterpart's payoff. Please notice: (i) You are matched with a **different player** every round; (ii) You will not know your counterparts' decisions until the end of the experiment; (iii) Every round has a 1/3 chance of being chosen as the payoff round. So, **it is in your best interest to treat each round as if it is the one that determines your payoff.**

At the beginning of each round, you will be told the gender of your counterpart for that round. (*Free-from Communication Treatment*: Then you and your counterpart each have one chance to **write a message to each other**. You will not be able to read your counterpart's message to you before you write your message. The experimenter will deliver your message to your counterpart, and your counterpart's message to you after both messages are written.) (*Restrictive Communication Treatments with One-dimensional or Two-dimensional Meanings*: Then you and your counterpart each have one chance to **send a message to each other. You can only choose one of the messages on the square message sheet by checking the box in front of it**. You will not be able to read your counterpart's message to you before you write your message. The experimenter will deliver your message to your counterpart, and your counterpart's message to you after both messages are written.) Afterwards you and your counterpart play a game. As shown in the graph, there are two options available for each player: Option **Pink** and Option **Blue**. **If both you and your counterpart choose Option Blue**, then both you and your counterpart earn **\$10**; **if both you and your counterpart choose Option Pink**, then both you and your counterpart earn **\$10**; **if you and your counterpart choose different options**, i.e. you choose Option Pink while your counterpart chooses Option Blue or you choose Option Blue while your counterpart chooses Option Pink, then both you and your counterpart earn **\$0**. Note that Pink is a color often preferred by females, and Blue is a color often preferred by males.

In summary, when you and your counterpart choose the **same option**, both of you earn **\$10**; if you and your counterpart choose **different options**, then you both earn **\$0**.

MESSAGE SHEET 1²³

□ I am choosing Pink.

□ I am choosing Blue.

²³ Message Sheet 1 is used for Restricted Communication with One-dimensional Meaning treatment.

MESSAGE SHEET 2²⁴

□ I am choosing Pink no
matter what.

□ I am choosing the color your
message tells me to. If your
message says the same, I am
choosing Blue.

□ I am choosing Blue no
matter what.

□

□ I am choosing the color your
message tells me to. If your
message says the same, I am
choosing Pink.

²⁴ Message Sheet 2 is used for Restricted Communication with Two-dimensional Meaning treatment.

ANSWER SHEET 1²⁵

Round #

Your randomly matched counterpart's gender is:

☐ Female

☐ Male

Your participate number:

Please choose one option:

☐ Option Blue

☐ Option Pink

²⁵ Answer sheet 1 is used for all the treatments conducted in U.S.

INSTRUCTIONS 2²⁶

欢迎参加实验！

在您的面前有一张卡片是您被分配到的参与号码。这个号码将被在整个实验中一直使用。请您举手示意如果您不知道或是无法阅读您的参与号码。

所有您赚得的报酬都会在实验结束后以现金形式保密性地支付。除了 10 元的准时出席费之外，根据您和您的随机分到的对家的决策，您可能赚得更多。

在实验过程中是**不允许交流**的。如果您有任何疑问，请举手示意，实验员会来协助。整个实验会持续大概 1 个小时。

您参与的匿名性是被确保的。您的名字不会被记录或与您的实验决策挂钩。

请您仔细阅读以下的实验说明。

实验说明

本实验共有**三轮**。每一轮包括一个**决策**。在每一轮中，您和你的随机分配到的对家的共同决策会决定你们的报酬。在每一轮开始之前，您会被随机非配到一个新的对家。在三轮实验结束之后，有一轮会被随机抽取来决定您和您当轮对家的报酬。请注意：（i）在每一轮中，您和**不同的对家**配对；（ii）直到实验结束您的支付报酬之前，您不会知道关于您任何对家的决策；（iii）每一轮都有 1/3 的机会被选为决定报酬的那一轮实验。所以，**您利益最大化的选择是把每一轮都当成最后可能被选定的那一轮实验来认真对待。**

²⁶ Instructions 2 are used for No communication Treatment, Free-form Communication Treatment conducted in China.

在每一轮开始时，您会被告知您的对家的性别。*(Free-form Communication Treatment:* 然后您和您的对家有一个机会给互相一个的通信。在您在纸上写下您的通信之前，您不会阅读到对家的信。也就是说，实验员会将您的纸条，和对家的纸条在双方都完成之后递送。)*)* 之后，您和您的对家玩一个游戏。如我们提供的彩图所示，在游戏中，您有两个选择：粉红色或是蓝色。如果您和您的对家都选择蓝色，那么您和您的对家都得到 40 元人民币；如果您和您的对家都选择粉红色，那么您和您的对家都得到 40 元人民币；如果您和您的对家选择不同颜色，比如说，您选粉红色但是您的对家选蓝色，又或是您选蓝色，您的对家选粉红色，那么您和您的对家都不在游戏中得到任何报酬。请注意：女性经常偏好粉红色；男性经常偏好蓝色。

归纳一下，当您和您的对家选择同一颜色，您们都得到 40 元。但是如果您和您的对家选不同颜色，那么你们都不在游戏中得到任何报酬。

ENGLISH TRANSLATION OF INSTRUCTIONS 2

Welcome to Our Experiment!

You have been assigned a participant number and it is on the note card in front of you. This number will be used throughout the study. Please inform us if you do not know or cannot read your participant number.

All payments will be made to you in cash privately at the end of the study today. In addition to 10 RMB for showing up on time, you can earn additional money based on your decisions and the decisions of your randomly matched counterparts.

No communication with other participants is allowed during this experiment. If you have any questions please raise your hand, and the experimenter will assist you. Today's experiment will last about 1 hour.

Your anonymity in this study is assured. Your name will never be collected or connected to any decision you make here today.

Please read these instructions carefully!

Instructions

There are **3 rounds** in this study. Each round consists of **1 decision**. Each round, your and your counterpart's payoff will be determined by your joint decisions. At

the beginning of each round, you will be randomly matched with a new counterpart. At the end of the experiment, one round will be randomly selected to determine your and your counterpart's payoff. Please notice: (i) You are matched with **a different player** every round; (ii) You will not know your counterparts' decisions until the end of the experiment; (iii) Every round has a 1/3 chance of being chosen as the payoff round. So, **it is in your best interest to treat each round as if it is the one that determines your payoff.**

At the beginning of each round, you will be told the gender of your counterpart for that round. (*Free-from Communication Treatment:* Then you and your counterpart each have one chance to **write a message to each other**. You will not be able to read your counterpart's message to you before you write your message. The experimenter will deliver your message to your counterpart, and your counterpart's message to you after both messages are written.) Afterwards you and your counterpart play a game. As shown in the graph, there are two options available for each player: Option **Pink** and Option **Blue**. **If both you and your counterpart choose Option Blue**, then both you and your counterpart earn **40 RMB**; **if both you and your counterpart choose Option Pink**, then both you and your counterpart earn **40 RMB**; **if you and your counterpart choose different options**, i.e. you choose Option Pink while your counterpart chooses Option Blue or you choose Option Blue while your counterpart chooses Option Pink, then both you and your counterpart earn **0 RMB**. Note that Pink is a color often preferred by females, and Blue is a color often preferred by males.

In summary, when you and your counterpart choose the **same option**, both of you earn **40 RMB**; if you and your counterpart choose **different options**, then you both earn **0 RMB**.

ANSWER SHEET 2²⁷

第 # 轮

你在本轮被随机分配的对家类别是：

☐ 女性

☐ 男性

您的参与号码：

请您选择一个选项：

☐ 蓝色

☐ 粉红色

²⁷ Answer sheet 2 is used for all the treatments conducted in China. The English translation of Answer sheet 2 is literally Answer sheet 1.

INSTRUCTIONS 3²⁸

Instructions

Welcome to this experiment. In addition to \$5 for showing up on time, you will be paid in cash based on your decisions in this experiment. Please read the instructions carefully. No communication with other participants is allowed during this experiment. Please raise your hand if you have any questions, and the experimenter will assist you.

Summary:

A excel file of messages written by participants in a previous experiment is open on your computer. The instructions for that previous experiment are also on your desk. You will not play this game today, but it is important you understand this game. Please take a few minutes to read that game's instructions now.

Task 1:

Your task 1 today is to determine the communication style of each message. For each message you can choose one of the following four options: ***Deferring Pink, Demanding Pink, Deferring Blue, Demanding Blue***. ***If the message does not reasonably fit into any of those options, then you can choose "None of these"***.

You should choose "***Deferring Pink***" if you believe the message writer prefers to choose pink, but is ultimately deferring the choice to her or his counterpart.

You should choose "***Demanding Pink***" if you believe the message writer prefers to choose pink, and requests that her or his counterpart also chooses pink.

You should choose "***Deferring Blue***" if you believe the message writer prefers to choose blue, but is ultimately deferring the choice to her or his counterpart.

²⁸ Instructions 3 are used for the Evaluation Sessions in U.S.

You should choose "***Demanding Blue***" if you believe the message writer prefers to choose blue, and requests that her or his counterpart also chooses blue.

Your earnings from task 1:

Three messages will be randomly chosen. The amount you earn in Task 1 depends on the way you and the other people in this room categorized those three messages. For each message, we will determine whether your categorization is the same as the most popular categorization of the people in this room. If none of your categorizations are the same as the most popular categorization, then you earn \$1. If one of your categorizations is the same, then you earn \$3. If two are the same then you earn \$7, and if all three of your categorizations match the most popular categorizations then you earn \$15.

Task 2:

Your second task is to guess whether the writer of each message is male or female.

Your earnings from Task 2:

Three messages will be randomly chosen. If all of your guesses are wrong you earn \$1; if one of your guesses is right you earn \$3; if two are right you earn \$7 and if all three are right you earn \$15.

Total earnings:

You can earn up to \$15 from the task 1 and up to \$15 from the task 2 in addition to the \$5 for showing up today.

INSTRUCTIONS 4²⁹

欢迎参加实验！

感谢你的准时出席。你已经获得了 5 元的准时出席费。如果你认真阅读以下实验说明，可能赚得比这多得多的钱。你赚得的所有报酬我们会在实验结束后马上以现金形式支付。实验过程中不允许交流。如果您有任何问题，请举手示意，实验员会来协助。

实验说明：

一份 excel 文件已经在你面前的电脑上打开。文件中我们输入了 300 多条由之前相关实验的参与者互相交流时写下的信息。这个相关实验的实验说明也在你的桌上。你今天不会参与这个相关实验相同或类似的实验，但是你对这个相关实验是否理解可能决定了你今天的任务完成情况。请用 3 分钟快速地阅读一下这个相关实验的说明。

任务 1：

你的第一个任务是要决定每一条信息的交流类别。有 5 个交流类别可以选择：“可妥协地建议粉红色”，“强烈要求对方选择粉红色”，“可妥协地建议蓝色”，“强烈要求对方选择蓝色”，只有一条信息完全无法被分进这 4 个类别中的任何一个时，你可以选择“不属于以上类别”。

具体来说，你应该选择“可妥协地建议粉红”如果你认为写下信息的人更愿意选择粉红色，但是最终愿意妥协于对方的意见。

你应该选择“强烈要求对方选择粉红色”如果你认为写下信息的人更愿意选择粉红色，并且也强烈要求对方也选择粉红色。

你应该选择“可妥协地建议蓝色”如果你认为写下信息的人更愿意选择蓝色，但是最终愿意妥协于对方的意见。

你应该选择“强烈要求对方选择蓝色”如果你认为写下信息的人更愿意选择蓝色，并且也强烈要求对方也选择蓝色。

任务 1 的报酬

我们随机抽取了三条信息放在信封中。实验最后，我们会把你对信息的分类和在场的其他参与者比较。如果你的归类和在场的大部分人相同，则算对。三条信息当中，如果全不对，那么你只得 5 元；如果对一条，得 10 元；如果对两条，得 20 元；如果全对，得 35 元。

任务 2：

你的第二个任务是猜写下信息的之前相关实验的参与者是男生还是女生。

任务 2 的报酬

我们会用与任务 1 相同的三条随机抽取的信息。三条信息中，如果全不对，那么你只得 5 元；如果对一条，得 10 元；如果对两条，得 20 元；如果全对，得 35 元。

总报酬

²⁹ Instructions 4 are used for the Evaluation Sessions in China.

你的总报酬是 5 元的准时出席费，加上任务 1 的报酬，再加上任务 2 的报酬。

ENGLISH TRANSLATION OF INSTRUCTIONS 4

Instructions

Welcome to this experiment. In addition to 5 RMB for showing up on time, you will be paid in cash based on your decisions in this experiment. Please read the instructions carefully. No communication with other participants is allowed during this experiment. Please raise your hand if you have any questions, and the experimenter will assist you.

Summary:

A excel file of messages written by participants in a previous experiment is open on your computer. The instructions for that previous experiment are also on your desk. You will not play this game today, but it is important you understand this game. Please take a few minutes to read that game's instructions now.

Task 1:

Your task 1 today is to determine the communication style of each message. For each message you can choose one of the following four options: ***Deferring Pink, Demanding Pink, Deferring Blue, Demanding Blue***. ***If the message does not reasonably fit into any of those options, then you can choose "None of these"***.

You should choose "***Deferring Pink***" if you believe the message writer prefers to choose pink, but is ultimately deferring the choice to her or his counterpart.

You should choose "***Demanding Pink***" if you believe the message writer prefers to choose pink, and requests that her or his counterpart also chooses pink.

You should choose "***Deferring Blue***" if you believe the message writer prefers to choose blue, but is ultimately deferring the choice to her or his counterpart.

You should choose "***Demanding Blue***" if you believe the message writer prefers to choose blue, and requests that her or his counterpart also chooses blue.

Your earnings from task 1:

Three messages will be randomly chosen. The amount you earn in Task 1 depends on the way you and the other people in this room categorized those three messages. For each message, we will determine whether your categorization is the same as the most popular categorization of the people in this room. If none of your categorizations are the same as the most popular categorization, then you earn 5 RMB. If one of your categorizations is the same, then you earn 10 RMB. If two are the same then you earn 20 RMB, and if all three of your categorizations match the most popular categorizations then you earn 35 RMB.

Task 2:

Your second task is to guess whether the writer of each message is male or female.

Your earnings from Task 2:

Three messages will be randomly chosen. If all of your guesses are wrong you earn 5 RMB; if one of your guesses is right you earn 10 RMB; if two are right you earn 20 RMB and if all three are right you earn 35 RMB.

Total earnings:

You can earn up to 35 RMB from the task 1 and up to 35 RMB from the task 2 in addition to the 5 RMB for showing up today.

REFERENCES

- Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica*, 50, 1431–1451.
- Battaglini, M., and U. Makarov (2011). Cheap Talk with Multiple Audiences: An Experimental Analysis. *Games and Economic Behavior*, 83, 147–164.
- Binger B.R., Hoffman, E., Libecap, G.D. (1990) “An Experimentric Study of the Cournot Theory of Firm Behavior,” *Working paper, University of Arizona*.
- Blume, A., Ortmann, A. (2007) “The Effects of Costless Pre-play Communication: Experimental Evidence from Games with Pareto-ranked Equilibria,” *Journal of Economic Theory*, 132: 274-290.
- Brandts, J., Cooper, D. (2007). “It’s What You Say, Not What You Pay: An Experimental Study of Manager-employee Relationships in Overcoming Coordination Failure,” *Journal of the European Economic Association*, 5(6): 1223-1268.
- Cason, T., Mui, V. (2015) "Rich communication, social motivations, and coordinated resistance against divide-and-conquer: A laboratory investigation." *European Journal of Political Economy* 37: 146-159.
- Cason, T., Sheremeta, R., Zhang, J. (2012) “Communication and Efficiency in Competitive Coordination Games,” *Games and Economic Behavior*, 76: 26-43.
- Chakraborty, A., and R. Harbaugh (2010). Persuasion by Cheap Talk. *American Economic Review*, 100, 2361–382.
- Charness, G. (2000), “Self-Serving Cheap Talk: A Test of Aumann’s Conjecture,” *Games and Economic Behavior* 33, 177-194.

- Charness, G., Dufwenberg, M. (2006), "Promises and Partnership," *Econometrica*, 74 (6): 1579-1601.
- Charness, G., Dufwenberg, M. (2010), "Bare Promises: An Experiment," *Economic Letters*, 107: 281-283.
- Charness, G., Grosskopf, B. (2004), "What Makes Cheap Talk Effective? Experimental Evidence," *Economics Letters*, 83: 383-389.
- Chen, Y., Jeon, G., Kim, Y. (2014) "A Day without a Search Engine: An Experimental Study of Online and Offline Searches," *Experimental Economics*, 17:512-536.
- Choi, S., Lee, J. (2014) "Communication, coordination, and networks." *Journal of the European Economic Association* 12.1: 223-247.
- Clark, K., Kay, S., Sefton, M. (2001), "When Are Nash Equilibria Self-enforcing? An Experimental Analysis," *International Journal of Game Theory*, 29:495-515.
- Cooper, R., Dejong, D.V., Forsythe, R., and Ross, T.W. (1989), "Communication in the Battle of the Sexes Games: Some Experimental Results," *RAND Journal of Economics* 20(4), 568-587.
- Cooper, R., Dejong, D.V., Forsythe, R., and Ross, T.W. (1992), "Communication in Coordination Games," *Quarterly Journal of Economics*, 53: 739-771.
- Cooper, D., Kagel, J. (2005) "Are Two Heads Better Than One? Team versus Individual Play in Signaling Games," *American Economic Review*, 95(3): 447-509.
- Cooper, D. Kühn, K. (2014) "Communication, Renegotiation, and the Scope for Collusion," *American Economic Journal: Microeconomics*, 6(2): 247-278.
- Cooper, D. Kühn, K. (2015) "Communication and Cooperation: A Methodological Study," Working Paper.

Crawford, V. (2003) "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions," *American Economic Review*, 93: 133-149.

Crawford, V., Sobel, J. (1982) "Strategic information transmission." *Econometrica: Journal of the Econometric Society*: 1431-1451.

Daughety, A., Forsythe, R. (1987) "Regulation and the Formation of Reputations: A Laboratory Analysis," in *Public Regulation: New Perspectives on Institutions and Policies*, E. Bailey, ed., *MIT press*.

Daughety, A., Forsythe, R. (1987) "The Effects of Industry-wide Price Regulation on Industrial Organization," *Journal of Law, Economics, & Organization*, 3(2): 397-434.

Dawes, R.M., McTavish, J., Shaklee, H. (1977) "Behavior Communication and Assumptions about Other People's Behavior in a Commons Dilemma Situation," *Personality and Social Psychology*, 35(1): 1-11.

Deck, C., Servátka, M., Tucker (2013) "An Examination of the Effect of Messages on Cooperation under Double-blind and Single-blind Payoff Procedures," *Experimental Economics*, 16:597- 607.

Demichelis, S., Weibull, J.W. (2008) "Language, Meaning, and Games: A Model of Communication, Coordination and Evolution," *American Economic Review*, 98(4):1292-1311.

Duffy, J., Feltovich, N. (2002) "Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talks," *Games and Economic Behavior*, 39:1-27.

Duffy, J., Feltovich, N. (2006) "Words, Deeds, and Lies: Strategic Behavior in Games with Multiple Signals," *Review of Economic Studies*, 73(3): 669-688.

Ellingsen, T., Johannesson, M. (2007) "Anticipated Verbal Feedback Induces Altruistic Behavior," *Evolution and Human Behavior*, 29(2): 100-105.

Ellingsen, T., Östling, R. (2010) “When Does Communication Improve Coordination,” *American Economic Review*, 100 (4): 1695-1724.

El-Gamal, M., Grether, D. (1995) “Are People Bayesian? Uncovering Behavioral Strategies,” *Journal of the American Statistical Association* 90(432): 1137-1145.

Farrell, J. (1987) “Cheap talk, Coordination, and Entry,” *RAND Journal of Economics*, 18:34-39.

Farrell, J. (1988). “Communication and Nash Equilibrium,” *Economic Letters*, 27: 209-214.

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

Frey B., Bohnet I. (1995), “Institutions Affect Fairness: An Experimental Investigation,” *Journal of Institutional and Theoretical Economics*, 151/2:286-303.

Gächter, S., Nosenzo, D., Sefton, M. (2013) “Peer Effects in Pro-social Behavior: Social Norms or Social Preferences?” *Journal of the European Economic Association*, 11(3): 548-573.

Hennig-Schmidt, H., Li, Z.-Y., Yang, C. (2008) “Why People Reject Advantageous Offers – Non-monotone Strategies in Ultimatum Bargaining,” *Journal of Economic behavior and Organization*, 65:373-384.

Houser, D., Xiao, E. (2011), “Classification of Natural Language Messages Using A Coordination Game,” *Experimental Economics*, 14(1); 1-14.

Holm, H.J. (2000) “Gender-Based Focal Points,” *Games and Economic Behavior* 32, 292-314.

Holt, C., Davis, D. (1990) “The Effects of Non-binding Price Announcements on Posted-offer Markets,” *Economics Letters*, 34 (4): 307-310.

Isaac, R.M., Ramey, V., Williams, A. (1984) "The Effects of Market Organization on Conspiracies in Restraint of Trade," *Journal of Economic Behavior & Organization*, 5(2): 191-222.

Isaac, R.M., Walker, J. (1988) "Communication and Free Riding Behavior: the Voluntary Contributions Mechanism," *Economic Inquiry* 26(4): 585-608.

Kimbrough, E., Smith, V., Wilson, B. (2008) "Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long Distance Trade," *American Economic Review*, 98(3): 1009-39.

Lai, E. K., Lim, W., Wang, J. T. Y. (2015). An experimental analysis of multidimensional cheap talk. *Games and Economic Behavior*, 91, 114-144.

Levy, G., Razin, R. (2007). On the Limits of Communication in Multidimensional Cheap Talk: A Comment. *Econometrica*, 75, 885–893.

Lundquist, T., Ellingsen, T., Gribbe, E., Johannesson, M. (2009) "The Aversion to Lying," *Journal of Economic Behavior and Organization*, 70: 81-92.

Marschak, J. (1965) "Economics of Language," *Behavioral Science*, 10 (2): 135-140.

Mohlin, E., Johannesson, M. (2008) "Communication: Content or Relationship," *Journal of Economic Behavior & Organization*, 65: 409-419.

Ong, Q., Riyanto, Y., Sheffrin, S. (2012) "How Does Voice Matter? Evidence from the Ultimatum Game," *Experimental Economics*, 15: 604-621.

Palfrey, T., Rosenthal, H., & Roy, N. (2015). How cheap talk enhances efficiency in threshold public goods games. *Games and Economic Behavior*.

Rabin, M. (1991) "Focal Points in Pre-Game Communication," Working paper, *Department of Economics, UC Berkeley*.

Rabin, M. (1994) "A Model of Pre-game Communication," *Journal of Economic Theory* 63: 370-391.

Rubinstein, A. (2000) "Economics and Language: Five Essays," *Cambridge University Press*.

Schelling, T.C. (1960) "The Strategy of Conflict," *Cambridge MA: Harvard University Press*.

Selten, R., Warglien, M. (2007), "The Emergence of Simple Languages in an Experimental Coordination Game," *PNAS*, 104 (18): 7361-7366.

Schotter, A., Sopher, B. (2007) "Advice and Behavior in Intergenerational Ultimatum Games: An Experimental Approach," *Games and Economic Behavior*, 58: 365-393.

Sutter, M., Strassmair, C. (2009) "Communication, Cooperation and Collusion in Team Tournaments - An Experimental Study," *Games and Economic Behavior*, 66: 506-525.

Valley, K., Moag, J., Bazerman, M. (1998) "A Matter of Trust: Effects of Communication on the Efficiency and Distribution of Outcomes," *Journal of Economic Behavior and Organization*, 34: 211-238.

Van Huyck, J., Battalio, R., Beil, R. (1990) "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80: 234-248.

Van Huyck, J., Battalio, R., Beil, R. (1993) "Asset Markets as an Equilibrium Selection Mechanism: Coordination Failure, Game Form Auctions, and Tacit Communication," *Games and Economic Behavior*, 5: 485-504.

Vespa, E., Wilson, A. J. (2015). Communication with Multiple Senders: An Experiment. *Quantitative Economics*.

Wittgenstein, L. (1957) "The Philosophical Investigations," *The Macmillan Company, New York*.

Xiao, E. and Houser, D. (2005). "Emotion Expression in Human Punishment Behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 102(20): 7398-7401.

Xiao, E., Houser, D. (2007) "Emotion Expression and Fairness in Economic Exchange," *Manuscript, George Mason University*.

BIOGRAPHY

Siyu Wang graduated from Chengdu No.7 High School, Sichuan, China, in 2006. She received her Bachelor of Science from Shanghai Jiao Tong University in 2010. She received her Master of Arts in Economics from George Mason University in 2013.