Cheap Talk Games: Comparing Direct and Simplified Replications

Fu-Wen Hsieh and Joseph Tao-yi Wang*

23st May, 2016

Abstract

To study strategic information transmission in organizations, we conduct a simplified version (with only 3 states) of the sender-receiver game experiment designed by Wang, Spezio, and Camerer (2010), in which an informative sender advises an uninformed receiver to take an action (to match the true state), but has incentives to exaggerate. We also have the same subjects play the original 5-state game. We find similar "overcommunication" behavior with Taiwanese subjects—messages reveal more information about the true state than what equilibrium predicts—that let us classify subjects into various level-k types. However, results from the simplified version are closer to equilibrium prediction, with more senders robustly classified as level-2.

J.E.L. classification codes: C72, C91, D83

Keywords: Sender-Receiver Game, Strategic Information Transmission, Lying, Laboratory Experiment

* Department of Economics, National Taiwan University, 1 Roosevelt Road, Section 4, Taipei 106, Taiwan. Fu-Wen Hsieh: r01323046@ntu.edu.tw; Joseph Tao-yi Wang: josephw@ntu.edu.tw. Shu-Yu Liu, Meng-Chien Su and Ally Wu provided excellent research assistance. We thank comments from Chen-Ying Huang, Pohan Fong, and Jian-Da Zhu. We thank financial support from the Ministry of Science and Technology of Taiwan (MOST), formerly National Science Council (NSC) of Taiwan (Grant# 102-2628-H-002-002-MY4). All remaining errors are our own.

1. Introduction

We acquire information from parents, friends, advisers, salespersons, doctors and so on; however, conflict of interest affects the way we cope with these information. Consumers tend to be skeptical about salespersons introducing their products because they know salespersons have incentives to lie. This is an example of a strategic sender-receiver game where senders have information advantage and their preferences are not aligned with receivers who make decisions. The amount of information transmitted is affected by how strong the conflicts of interest are between senders and receivers. Similar examples include analysts vs. investors, doctors vs. patients, candidates vs. interviewers, and so on.

Communication is also vital in organizations, as higher level managers are neither omnipotent nor omniscience, and need to gather information from their subordinates and direct them to carry out orders. When all parties are on the same page and want to work toward the same goal, full information transmission is expected and usually realized. For example, Alonso, Dessein, and Matouschek (2008) and Dessein and Santos (2006) show that communication solves the coordination problem created by specialization and decentralizing authority when organizations adopt to local information.

However, when people have different agendas and there is conflict of interests, strategic information transmission occurs, communication within an organization may not be perfect, especially when the informed party has a biased agenda. For example, project leaders may have the incentives to paint a rosy picture about the current project, even when it is best for top management to cut losses and terminate it. In fact, Rantakari (2008) shows that incentive alignment and centralized control are substitutes and organizations fall back to more centralized authority structure in the presence of conflict of interest.

Focusing on strategic information transmission, Crawford and Sobel (1982) consider a one-dimensional sender-receiver game. A sender who has full information sends a message to a receiver, and the receiver takes an action that decides payoffs of both players. Crawford and Sobel (1982) predict that information transmission decreases as the preference difference between the sender and the receiver increases. They also predict no informative equilibrium exists when conflict of interest is sufficiently large. That is, "babbling equilibrium" is the most informative equilibrium.

Experimental economists have used controlled experiments to test theoretical predictions of Crawford and Sobel (1982). These experiments usually adopt a discrete state space and vary in the number of states, messages and actions. Gneezy (2005) adopts a simple sender-receiver game with 2 states x 2 messages x 2 actions and reports that senders are more likely to lie when receiver's loss decreases or sender's profit increases. The experiments of Dickhaut et al. (1995) Cai and Wang (2006) and Wang et al. (2010) all study the change in sender and receiver behavior under different conflict of interest. Dickhaut et al. (1995) adopt a sender-receiver game with 4 states, 4 actions, and messages being single or consecutive sequence of integers from 1 to 4 to show that as preferences diverge, less information are transmitted.[1] Cai and Wang (2006) study sender-receiver games with 5 states and 9 actions (5 actions corresponding to states and 4 intermediate action), as well as messages being any combination of states.[2] Their experimental results show that senders "overcommunicate" and send informative messages to receivers, even when the equilibrium model predicts only the "babbling equilibrium" exists. Wang et al. (2010) adopt a setting where the numbers of states, messages and actions are all 5, and use eye-tracking to monitor the behavior

[1] Messages could be {1}, {2}, {3}, {4}, {1, 2}, {2, 3}, {3, 4}, {1, 2, 3}, {2, 3, 4}, {1, 2, 3, 4}.

[2] The state space is {1, 3, 5, 7, 9} and the action space is {1, 2, 3, 4, 5, 6, 7, 8, 9}. Senders could send messages like {1, 7}, {3, 5} by choosing one or more numbers in the state space.

of senders in sender-receiver games.[3] Follow-up experiments allow receivers to costly punish liars, let senders to costly keep silence or send vague messages, and introduce multiple senders or receivers (Sanchez-Pages and Vorsatz, 2007, 2009; Serra-Garcia et al., 2011; Battaglini and Makarov, 2014; Lai, Lim and Wang, 2015).[4]

The number of states (and corresponding messages and actions) would influence communication. When there are more possible states, there is more potential for interaction and deception. Hence, we expect the complexity of the information transmission process to increase with the number of states, even when the equilibrium model predicts the same most informative equilibrium.

Sender-receiver games with different parameter space and conditions are extensively studied in the past, but few research compare the results across different state space sizes. We address this by both replicating and modifying the experiment design of Wang et al. (2010). Our experiment consists of two parts. The first part is the Replicate Treatment with the same 5-state setting as in Wang et al. (2010), and the second part is a Simplified Treatment under an environment with only 3 states. In other words, the Simplified Treatment adopts similar experimental procedures, but has a smaller state space (and correspondingly reduced biases).[5] We focus on the comparison of senders' behavior in the two treatments.

Comparing the two treatments, we obtain comparative static results showing that subject behavior is closer to equilibrium prediction in the Simplified Treatment than in the Replicate Treatment. The equilibrium model fails to

---

[3] Senders see a state and send one number from {1, 2, 3, 4, 5}, and receivers choose an action from {1, 2, 3, 4, 5}.

[4] The few studies that go beyond the discrete state space are Minozzi and Woon (2016), Vespa and Wilson (2016) and Battaglini, Lai, Lim and Wang (2016).

[5] As the size of the state space decreases, the minimum amount of bias which leads to a babbling equilibrium decreases.

explain the abundant amount of low-type (L0 and L1) messages, while the level-k model can.[6] This is consistent with the literature on sender-receiver games, starting from Crawford (2003) who first analyzed sender-receiver games with a level-k model, setting the L0 sender message to be truth-telling rather than random. Cai and Wang (2006) and Wang, Spezio and Camerer (2010) both used the level-k model to explain their results. Kawagoe and Takizawa (2009) reported that the level-k analysis explains their results better than other theories.

Using maximum likelihood estimation, we classify senders into either L0, L1 or L2 types, and evaluate type classification stability by resampling senders' choices. 12-14% of senders are classified as L0 in both treatments. The proportion of L1 senders decrease from 39% to 25% when the size of the state space decreases from 5 to 3. In contrast, more senders are classified as L2 in the Simplified Treatment. Resampling of senders' choices shows that L2 senders are more stable in the Simplified Treatment. These results are consistent with the level-k model (but not cognitive hierarchy) since it predicts the more complicated Replicate Treatment would induce senders to think their opponents have lower levels of reasoning. These results also indicate that complicated communication channels lead to noisier behavior which reduces worker productivity (and hence efficiency). In contrast, simple and clear communication channels can increase organization efficiency, especially when productivity is related to strategic sophistication. Thus, organizations should design simple but sufficient channels for information transmission.

---

[6] Stahl and Wilson (1994), Nagel (1995), and Camerer et al. (2004) pioneered steps-of-reasoning models of bounded rationality. In the level-k model of Stahl and Wilson (1994) and Nagel (1995), subjects incorrectly believe their opponents have a specific level of bounded rationality, and play best response to this (naïve) belief. In contrast, the cognitive hierarchy model of Camerer et al. (2004) assumed subjects have correct but truncated beliefs about others since they cannot image the reasoning of types higher than themselves. Costa-Gomes and Crawford (2006) reported that the level-k model accounts for most of the predictable component of systematic deviations from equilibrium.

The remainder of this paper is organized as follows. We introduce the sender-receiver game in section 2. Section 3 explains the design of our experiments. Section 4 analyzes our experimental results, and section 5 concludes.

2. The Sender-Receiver Game

We conduct a simplified version with only 3 possible states of the sender-receiver game experiment designed by Wang, Spezio, and Camerer (2010). For comparison, we also replicate 20 rounds of their experiment before the simplified game. At the beginning of the experiment, we randomly assign players to be either senders or receivers. This role is kept the same throughout the experiment. In the first part of experiment, the Replicate Treatment, the sender is informed about the true state s at the beginning of each period, which is unknown to receiver and drawn from the state space $S = \{1, 2, 3, 4, 5\}$ with equal probability. Meanwhile, the sender is also informed about the bias b, which is the commonly known preference difference between the sender and the receiver. The bias is either 0, 1, or 2, and drawn from a probability distribution unknown to the subjects. Then, the sender sends a message to the receiver from the message space $M = \{1, 2, 3, 4, 5\}$. Seeing the message, the receiver chooses an action a from the action space $A = \{1, 2, 3, 4, 5\}$. Payoffs are determined by the following formulas for the receiver and the sender, respectively:

$u_R = 110 - 20\ |\mathrm{s} - \mathrm{a}|^{1.4}$ and $u_S = 110 - 20\ |\mathrm{s} + b - a|^{1.4}$.

In other words, senders earn the highest payoffs when the receiver chooses the action that equals to the true state plus the bias. In contrast, the receiver wants to choose the action that equals to the true state.

In the second part of experiment, the Simplified Treatment, the state, message, and action space change from $\{1, 2, 3, 4, 5\}$ to $\{1, 2, 3\}$. The true state is still drawn with equal probability, but now 1/3 each (instead of 1/5). The bias is shifted to be drawn from {-1, 0, 1}, though b = -1 occurs rarely and serves only as "catch" trials. In fact, the case of b = -1 is merely relabeling the states 1 to 3 as 3 to 1, and is mathematically identical to the case of b = 1.

Wang, Spezio, and Camerer (2010) derive the level-k model for the sender-receiver game, beginning with L0 senders who send truthful messages and L0 receivers credulously taking the action equal to the message. L1 senders best respond to L0 receivers, L1 receivers best respond to L1 senders and so on. This results in the behavioral predictions listed in Table 1 and Table 2. Note that Table 1 replicates Table 1 in Wang, Spezio and Camerer (2010), except for the SOPH types whose empirical best response is calculated using our data.

Take b = 1 in the Simplified Treatment as an example. Predictions are recorded in the middle panel of Table 2. L0 senders send messages 1, 2 and 3 for states 1, 2 and 3, respectively, while L0 receivers take actions 1, 2 and 3 when seeing messages 1, 2 and 3, respectively. This is recorded in the first row. Knowing L0 receivers would follow their messages, L1 senders send messages equal to their preferred action, or 2, 3 and 3 for states 1, 2 and 3, respectively. Knowing such exaggeration, L1 receivers take actions 1 and 2 when seeing messages 2 and 3, respectively, and choose action 1 when seeing the unexpected message 1. This corresponds to the second row. Anticipating L1 receivers' discounting of the message, L2 senders exaggerate further and always send message 3, while L2 receivers always take action 2 in response. This coincides with the babbling equilibrium and is listed in the third row. The fourth row specifies the empirical best response, or the optimal strategy when playing against a randomly drawn opponent from our data. This is the behavioral prediction for the sophisticated types (SOPH).

Table 1: Behavioral Predictions of the Level-k Model (Replicate)

Sender message (condition on state)	1	2	3	4	5	Receiver action (condition on message)	1	2	3	4	5
b = 0											
L0/EQ sender	1	2	3	4	5	L0/EQ receiver	1	2	3	4	5
b = 1											
L0 sender	1	2	3	4	5	L0 receiver	1	2	3	4	5
L1 sender	2	3	4	5	5	L1 receiver	1	1	2	3	4
L2 sender	3	4	5	5	5	L2 receiver	1	1	1	2	4
EQ/L3 sender	4	5	5	5	5	EQ/L3 receiver	1	1	1	1	4
SOPH sender	3	4	5	5	5	SOPH receiver	1	1	2	3	4
b = 2											
L0 sender	1	2	3	4	5	L0 receiver	1	2	3	4	5
L1 sender	3	4	5	5	5	L1 receiver	1	1	1	2	4
L2 sender	4	5	5	5	5	L2 receiver	1	1	1	1	4
EQ/L3 sender	5	5	5	5	5	EQ/L3 receiver	1	1	1	1	3
SOPH sender	5	5	5	5	5	SOPH receiver	2	2	2	3	4

Note: Sender messages for various level-k types are listed in the left, while receiver actions are in the right. The top panel is for the case of $b = 0$; the middle panel is for $b = 1$; the bottom panel is for $b = 2$. Each row corresponds to a particular type, and the five numbers represent the sender message (receiver action) for five possible states (messages).

Table 2: Behavioral Predictions of the Level-k Model (Simplified)

Sender message (condition on state)				Receiver action (condition on message)			
State	1	2	3	Message	1	2	3
b = 0							
L0/EQ sender	1	2	3	L0/EQ receiver	1	2	3
b = 1							
L0 sender	1	2	3	L0 receiver	1	2	3
L1 sender	2	3	3	L1 receiver	1	1	2
EQ/L2 sender	3	3	3	EQ/L2 receiver	1	1	2
SOPH sender	3	3	3	SOPH receiver	1	2	2
b = -1							
L0 sender	1	2	3	L0 receiver	1	2	3
L1 sender	1	1	2	L1 receiver	2	3	3
EQ/L2 sender	1	1	1	EQ/L2 receiver	2	3	3

Note: Sender messages for various level-k types are listed in the left, while receiver actions are in the right. The top panel is for the case of $b = 0$; the middle panel is for $b = 1$; the bottom panel is for $b = 2$. Each row corresponds to a particular type, and the five numbers represent the sender message (receiver action) for five possible states (messages).

3. Experimental Design and Procedure

We conducted a total of 8 experimental sessions at the Taiwan Social Science Experiment Laboratory (TASSEL) in National Taiwan University (NTU), consisting of 118 subjects with 10 to 20 subjects in each session. We made announcements on online forums (ptt.cc, the largest Bulletin Board System in Taiwan) and via email sent to the TASSEL subject pool. Subjects were NTU undergraduate or graduate students, which voluntarily signed up on TASSEL's online recruiting website. They were paired randomly to play 20 rounds of the Replicate Treatment and 40 rounds of the Simplified Treatment.[7] Experiments typically lasted for 2.5 to 3 hours. The exchange rate was 6 Experimental Standard Currency (ESC) for NT$1. Subjects earned between NT$623 and NT$1001 (including the NT$100 show up fee), with an average of NT$ 860.8 (approximately US$ 28.7).

All experiments were conducted in Chinese using z-Tree (Zurich Toolbox for Readymade Economic Experiments) developed by Fischbacher (2007). Subjects were informed about the game via paper instruction that were read out loud. Sample instructions are provided in the Appendix.

In the Replicate Treatment (first part), there are 3 practice rounds and 20 real rounds. Subjects were randomly assigned to the same role as sender or receiver for all 23 periods. States were drawn with equal probability, while biases 0, 1, 2 were drawn with probabilities 0.2, 0.4 and 0.4, respectively. Subjects were informed about the payoffs by both on-screen payoff tables (Figure 1) and tables in the experimental instructions (Table 3). Since we did not eye-track our subjects, we did not follow Wang, Spezio, and Camerer (2010) to add a small random number to the payoffs to make it uncertain (so subjects would look at the payoff table).

[7] In session 3, the Simplified Treatment had only 35 rounds due to a z-Tree crash.

In the Simplified Treatment (second part), there were 36 periods with bias equal to 0 or 1. Subjects were assigned the same role as in the Replicate Treatment. States were drawn with equal probabilities, while biases 0 and 1 were drawn with probabilities 0.25 and 0.75, respectively. In addition, there were 4 “catch” trials to attracting subject attention. Two rounds of “catch” trials had a bias of -1. Another 2 rounds of “catch” trials switched the role of the two players (so senders are now receivers and vice versa). Finally, the sender was asked to predict receiver action after sending the message, and earned extra 3 ESC if their prediction equals the action.

Figure 1 : Screenshot for the Replicate Treatment

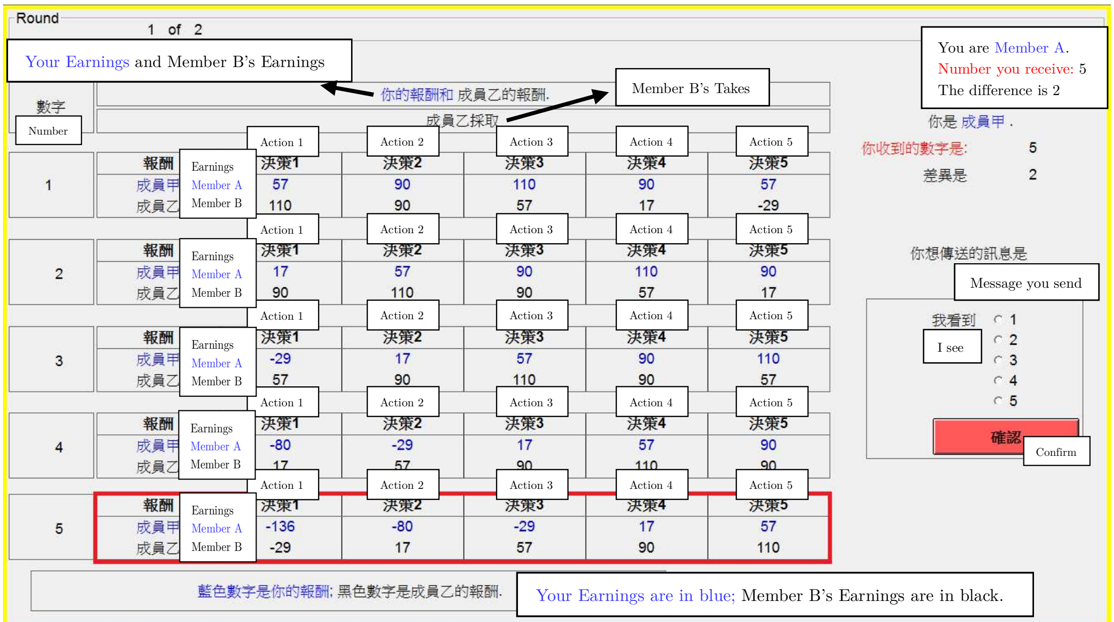


Table 3: Payoff Table in the Experiment Instructions

State – Action	0	±1	±2	±3	±4		
Receiver Payoff	110	90	57	17	-29		

State + Bias – Action	0	±1	±2	±3	±4	±5	±6
Sender Payoff	110	90	57	17	-29	-80	-136

4. Experimental Results

4.1 Sender Message and Receiver Action

We report results after dropping the two types of "catch" trials in the Simplified treatment. Results including these trials (Tables S2 and S3 in the Appendix) are similar to results without the "catch" trials. When the bias is 0, senders tell the truth and receivers follow their advice regardless of treatment. In the Replicate Treatment, only 8 out of 245 observations (4 of 245 observations) do we see sender sending message not equal to the true state (receiver taking action not equal to the message). Meanwhile, in the Simplified Treatment, none of the senders send messages different from the true state, and receivers take actions not equal to message in only 2 out of 542 observations.

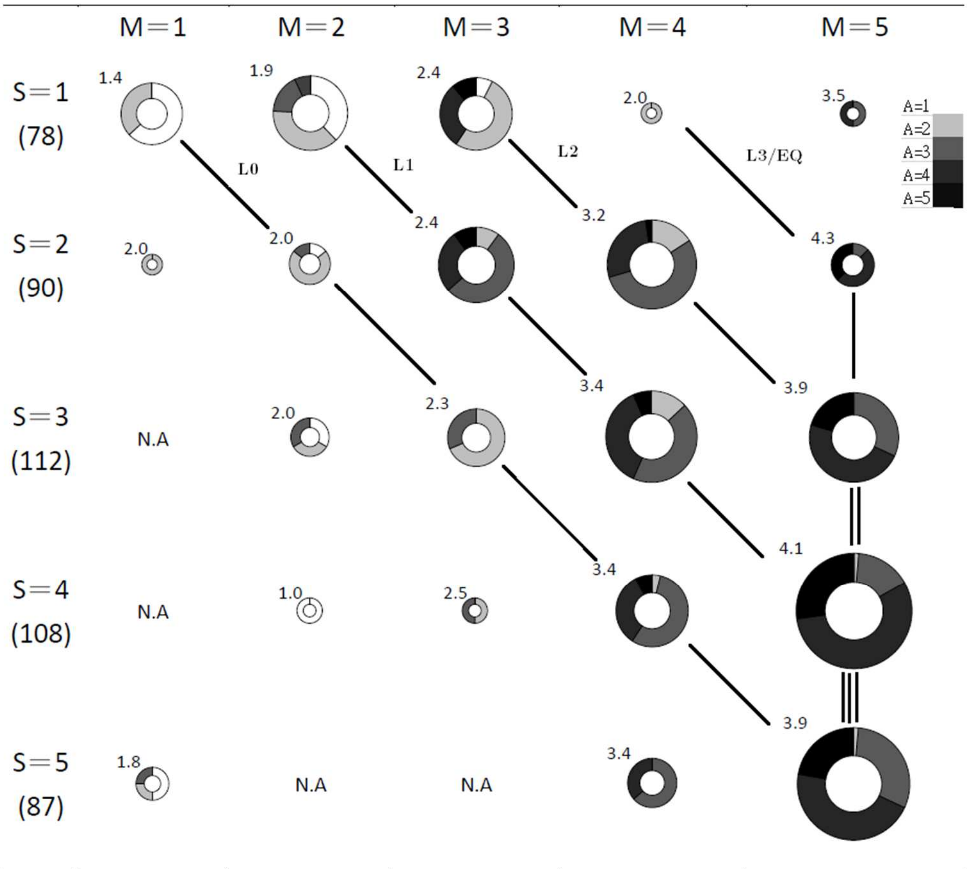
4.1.1 Replicate Treatment

Figure 2 and 3 report results for $b = 1$ and 2 in the Replicate Treatment. The size of the donut charts in Figure 2 and 3 are scaled by the frequency of occurrence for senders sending message conditional on a given state. The fractions in each donut chart represents actions taken by receivers for each message and state. The darker the color, the higher the action. The number near each donut chart is the average action taken by the receivers. Note that the receiver only knows the message, not the true state, so the average action for different states given a certain message are similar. The average action increases as the message increases.

As the level-k model predicts, charts locate above the diagonal are larger than those below. When $b = 1$, charts are largest when they are 1 or 2 steps (L1 or L2) above the diagonal (L0). Note that since the matrix is bounded, senders are predicted to all send message 5 when the state is 5. When $b = 2$, messages that are two steps above the diagonal (L0) indicate L1, and more states

(3 to 5) are predicted to exhibit message 5. Messages are now closer to the babbling equilibrium (all sending 5) than $b = 1$. This is partly predicted by the level-k model, but the frequency of L1 choices seems to be lower. These aggregate results are consistent with Wang, Spezio, and Camerer (2010).

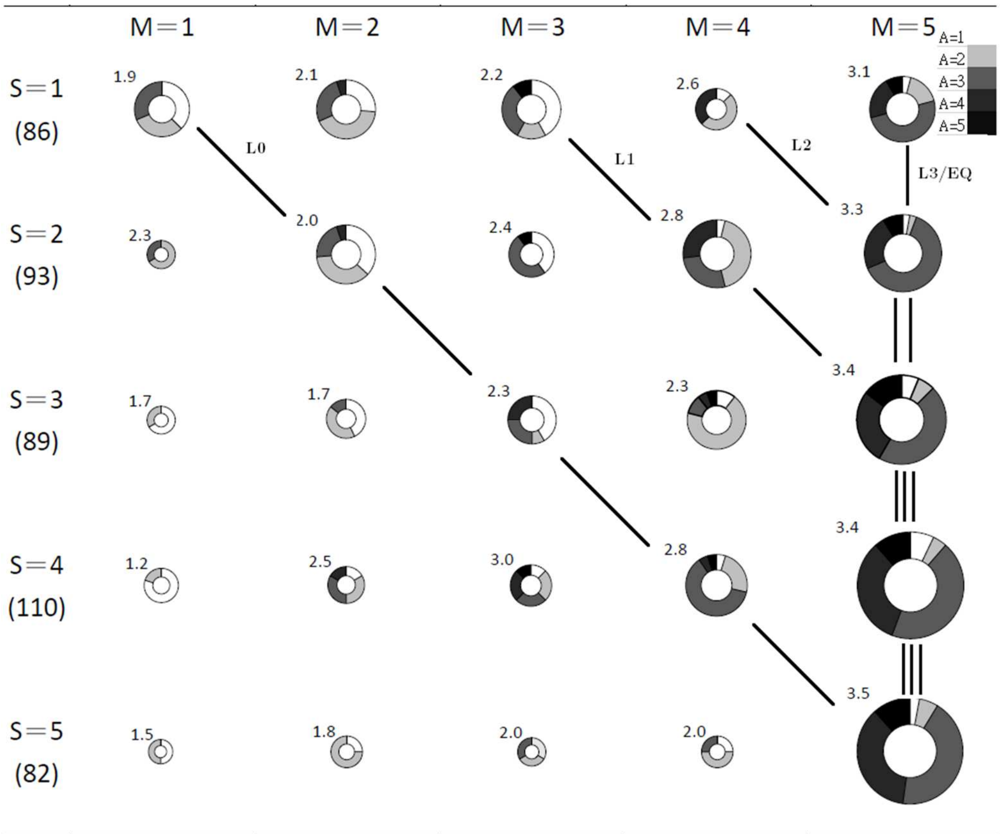
Figure 2: Message and Action under bias = 1, Replicate Treatment



Note: Donut size scaled by frequency of message occurrence. Fractions in each donut represents actions taken for each message and state. The darker the color, the higher the action. The number near each donut is the average receiver action taken.

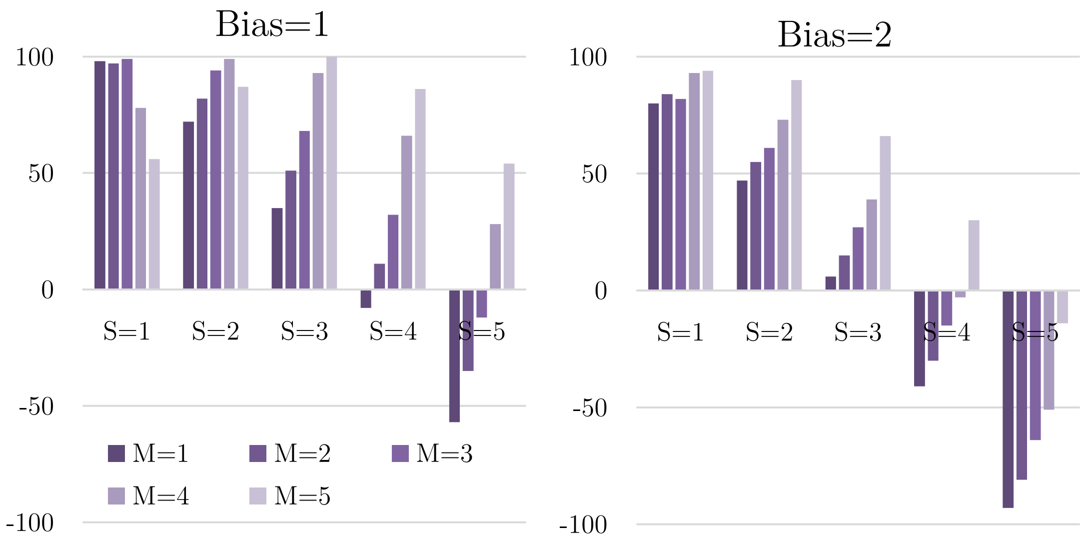
Most senders and receivers do well in the experiment. Figure 4 reports senders' expected payoff E(π) for sending different messages, and Figure 5 shows receivers' expected payoff E(π) for taking different actions. For most states, senders choose the message that yield the highest or second-highest payoff. The only two exceptions occur at $s = 1, 2$ for $b = 2$, when the difference between expected payoff for each message is small. On the other hand, receivers take the action that yield the highest or second-highest payoff for most messages. Since subjects rarely send message 5 when the true state is 1 or 2, this increases the likelihood of having higher states when seeing a message of 5, which induces receivers to choose higher actions. When the message is 5, the average action taken by receivers is 3.4 in our experiment, close to that of Wang, Spezio, and Camerer (2010) and higher than the babbling equilibrium.

Figure 3: Message and Action under bias = 2, Replicate Treatment



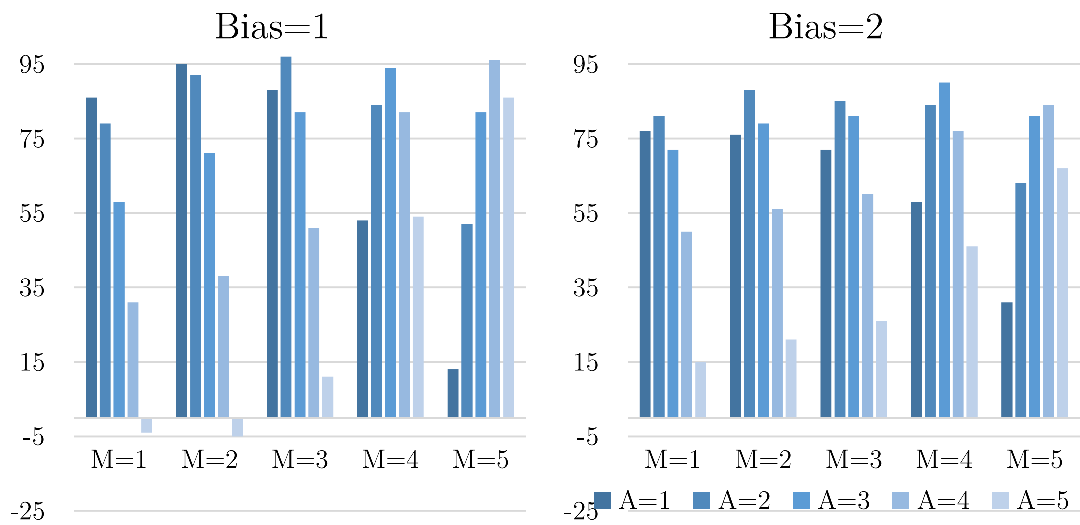
Note: Donut size scaled by frequency of message occurrence. Fractions in each donut represents actions taken for each message and state. The darker the color, the higher the action. The number near each donut is the average receiver action taken.

Figure 4: Sender's Expected Payoff, Replicate Treatment



Note: Bars represent sender's expected payoff sending message M for each state S.

Figure 5: Receiver's Expected Payoff, Replicate Treatment

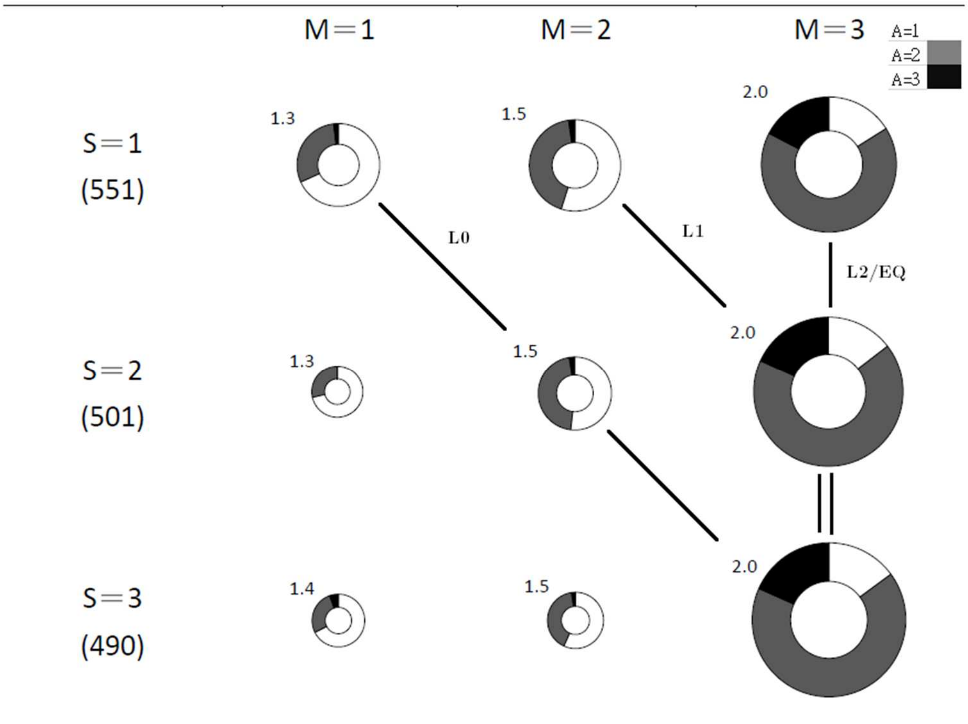


Note: Bars represent receiver's expected payoff taking action A for each message M.

4.1.2 Simplified Treatment

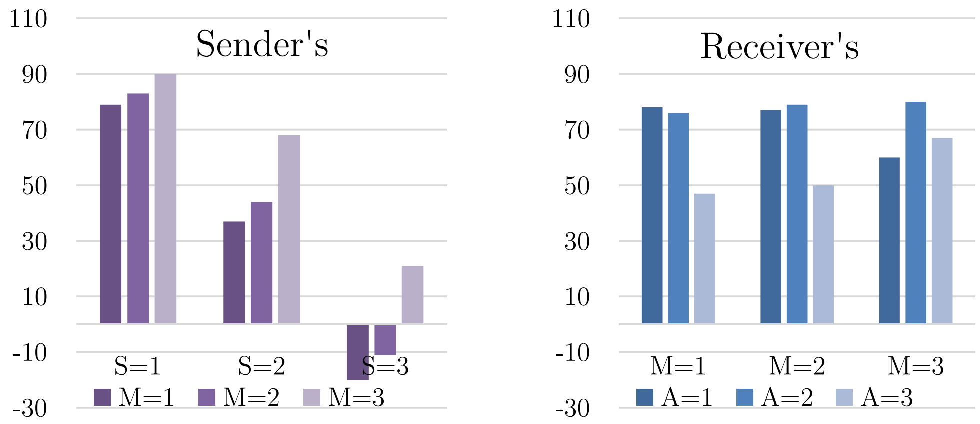
In Figure 6, the size of donut charts are re-scaled because the largest donut chart increases from 70 to 409, and the message space is reduced from 5 to 3. The results are close to the babbling equilibrium (where all messages are 3), and similar to that of $b = 2$ in the Replicate Treatment. Senders sometimes do not send the equilibrium message of 3 when the true state is 1, but split between 1 and 2 instead. This reveals some information regarding the true state. Seeing a message of 3, receivers on average take action 2, coinciding with the babbling equilibrium prediction. Figure 7 shows that senders earn the most when always sending message 3. However, sending other messages yields payoffs close to optimal when the state is 1. This low cost of sending other messages, together with overcommunication, explains their abundance. Seeing message 3, receivers earn the most by taking action 2.

Figure 4: Final choices under bias = 1, Simplified Treatment



Note: Donut size scaled by frequency of message occurrence. Fractions in each donut represents actions taken for each message and state. The darker the color, the higher the action. The number near each donut is the average receiver action taken.

Figure 7: Player's Expected Payoff, Simplified Treatment, Bias = 1



Note: Bars represent expected payoff choosing message M (action A) for state S (message M).

Table 4: Sender Message Sophistication, by Subject

	Low(Simplified)	Mixed(Simplified)	High(Simplified)	Total
Low(Replicate)	**15**	6	5	26
Mixed(Replicate)	4	**5**	9	18
High(Replicate)	1	1	**13**	15
Total	20	12	27	59

Note: Cells show number of sender subjects with the corresponding message sophistication in the Replicate (row) and Simplified (column) Treatment.

4.1.3 Message Sophistication

In the aggregate data, we observe more messages consistent with L2 or higher sent in the Simplified Treatment, so we examine individual level data. We exclude the bias and state combinations which messages of L1 and L2 coincide and split messages into either High (L2 or above) or Low (L1 or below). We then deem a sender's message sophistication as High or Low when the sender sends High or Low messages more than 3/4 of the time. Otherwise, the sender is Mixed.

Table 4 compare senders across the two Treatment. More than half of the senders maintain the same message sophistication across treatments. In fact, 15 senders are Low in both, 5 are Mixed, and 13 are High. 20 out of 59 senders increase their message sophistication in the Simplified Treatment, while only 6 senders decrease their message sophistication.

4.2 Information Transmission

Table 5 shows comparative static results for the correlation between states, messages, and actions. Predictions of the equilibrium model are in parentheses. The results are similar to that reported by Wang, Spezio, and Camerer (2010) and confirm the prediction made by Crawford and Sobel (1982): All correlations decrease as bias increases, even though the empirical correlations do not exactly match theory. This is because subjects tend to follow the level-k model instead of equilibrium. As a result, we observe in the Replicate Treatment overcommunication when b is large, adherence to the truth-telling equilibrium when $b = 0$, and heterogeneous choices when $b = 1$. What is more, Individual subjects may also exhibit noisy behavior which lowers aggregate correlation.

The equilibrium model predicts a babbling equilibrium when the bias is 2, in which senders always send the same message of 5 and receivers take actions according to their prior belief about the true state. Therefore, we expect zero correlation between states, messages, and actions in equilibrium. Instead, the correlations of our data are not close to 0, consistent with "overcommunication" found by Cai and Wang (2006). In the Simplified Treatment, the empirical correlations given a bias of 0 (1) are similar to those of bias equal to 0 (2) in the Replicate Treatment (which share the same equilibrium prediction).

Table 6 shows that subject average payoffs are lower than equilibrium prediction, though standard deviations are larger than predicted. This shows a potential Pareto improvement for the subjects. Furthermore, the difference between average and predicted payoffs is larger for senders than for receivers. This is likely due to the concave payoff function and unreachable sender ideal actions (above 5 or 3), making mistakes more costly for senders than for receivers.

Table S4 and S5 in the Appendix split the data into the first and second half to examine potential learning effects. The results show that the potential learning effect is quiet small. In fact, the correlations do not always converge to the prediction of the equilibrium model.

Table 5: Information Transmission: **S**tates, **M**essages and **A**ctions

Treatment	Bias	Corr(s, m)	Corr(m, a)	Corr(s, a)
Replicate	0	0.99(1)	0.99(1)	0.99(1)
	1	0.68(0.71)	0.74(1)	0.53(0.71)
	2	0.39(0)	0.50(0)	0.24(0)
Simplified	0	1.00(1)	1.00(1)	1.00(1)
	1	0.24(0)	0.43(0)	0.12(0)

Note: Information transmission measured by correlation between states, messages and actions. Equilibrium prediction in parenthesis. The equilibrium correlation between States and Actions should be 0.71, instead of 0.65 as stated in Wang, Spezio, and Camerer (2010).

Table 6: Average Sender and Receiver Payoffs

Bias	u_S	Predicted u_S	u_R	Predicted u_R
Replicate Treatment				
0	109.10(4)	110(0)	109.10(4)	110(0)
1	81.48(30)	87.40(17)	88.65(23)	91.40(19)
2	40.58(55)	49.00(50)	78.27(31)	80.80(21)
Simplified Treatment				
0	99.89(2)	100(0)	99.89(2)	100(0)
1	55.77(39)	63.67(33)	75.88(24)	80.80(14)

Note: Standard deviation in parenthesis.

4.3 Level-k Type Classification

We estimate the level-k type of each sender, assuming that senders are of a certain level-k type throughout the experiment and choose each message with probability $\frac{e^{\lambda\mathbb{E}[\pi(m|s)]}}{\sum_{\mu\in M} e^{\lambda\mathbb{E}[\pi(\mu|s)]}}$ where $\mathrm{E}[\pi(m \mid s)]$ is the expect payoff of sending message m given state s. Maximum Likelihood Estimations are conducted using this logit error structure.

We classify senders as L0, L1 or L2 using either Replicate or Simplified

Treatment data, or both. We do not identify L3 and SOPH types because the SOPH type coincides with L2 in the Simplified Treatment, as shown in Table 1 and 2. Moreover, in the Replicate Treatment, the SOPH type acts like L2 when the bias is 1, but acts like L3 when the bias is 2. Hence, there is separation between L2 and SOPH type for only one case (b = 2, s = 1). Likewise, there are only 2 cases that tell L3 from SOPH (b = 1, s = 1, 2), while L2 and L3 differ in only 3 cases.[8] However, we have only 20 rounds in our Replicate Treatment, instead of 45 rounds in Wang et al. (2010). So, each case occurs 20x0.4x0.2=1.6 times on average, making it difficult to distinguish L2, L3, and SOPH types.[9]

Figure 8 shows the main results of our type classification using a logit error structure and dropping the data of bias equal to 0. In the Simplified Treatment, 12% of senders are classified as L0, 25% as L1, and 63% as L2. In the Replicated Treatment, the proportion of L0 type is unchanged, while that of L1 increases from 25% to 39%. Table 7 compares type classification across treatments. Most senders who are classified as L2 in the Replicate Treatment are also classified as L2 in the Simplified Treatment. Sender who are not classified as L2 in the Replicate Treatment, are classified as a higher type in the Simplified Treatment more than half of the time. In fact, predicting subjects' level-k type in the Simplified Treatment with their level-k type in the Replicate Treatment using OLS regression yields a coefficient of 0.508 (s.e. 0.112) and an intercept of 0.829 (s. e. 0.171).

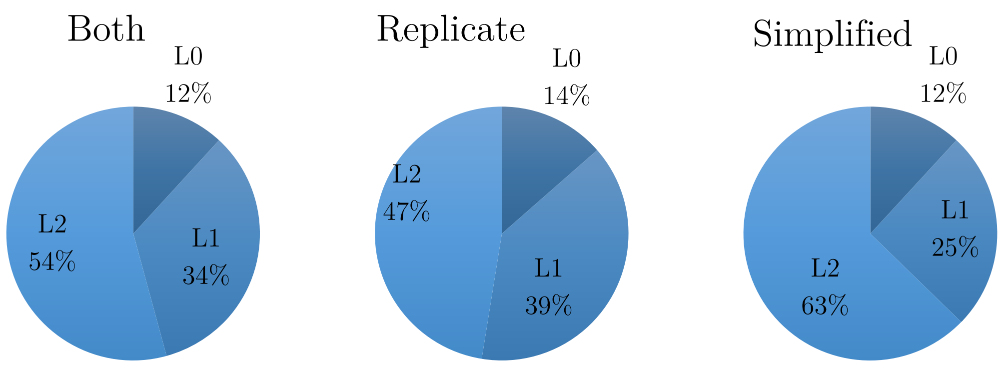
Table 8 reports some robustness checks. To begin with, we compare the classification results using data from the Replicate Treatment, the Simplified Treatment, and both. 78% of the senders are identified as the same type throughout all three classifications. Using a spike-logit error structure as in

[8] The 3 cases are the 1 case where L2 and SOPH differ and 2 cases that L3 and SOPH differ.

[9] Attempts to identify L3 types find only 7 (12%) L3 senders (previously classified as L2).

Costa-Gomes and Crawford (2006) in the Replicate Treatment results in only 10% of the senders being classified as another type. Thirdly, including $b = 0$ data only results in a small number of senders being identified as a different type. Finally, we resample each subject 1000 times and perform classification. The top panel of Table S6 and the bottom panel of Table S7 in the Appendix report the number of times the resampling type coincide with the original.[10] We find the Simplified Treatment not only induces more L2 senders, but more stable classification of these L2 senders: 84% of them are robustly classified as L2 more than 95% of the time, compared to 40%. What is more, comparing across panels within Table S6 and S7 show that this is not due to more rounds in the Simplified Treatment: Resampling 36 rounds instead of 20 rounds increases robust L2 senders by only 8-14%.

Figure 8: Level-k Classification (Replicate, Simplified or Both)



Note: Pie chart indicates the distribution of level-k types classified using data from the Replicate Treatment, Simplified Treatment, or both.

Table 7: Change of Level-k Classification

	L0(Simplified)	L1(Simplified)	L2(Simplified)	Total

[10] In the Replicate Treatment, we draw 20 rounds with replacement. In the Simplified Treatment, we draw 36 rounds as we exclude the 4 "catch" trials where bias is -1 or roles are reversed. In session 3, for the Simplified Treatment we also draw 36 rounds (with replacement) from the 31 rounds of data we obtained before z-Tree crashed.

L0(Replicated)	**3**	5	0	8
L1(Replicated)	2	**8**	13	23
L2(Replicated)	2	2	**24**	28
Total	7	15	37	59

Note: Each cell shows the number of sender subjects with the corresponding level-k type in the Replicate (row) and Simplified (column) Treatment.

Table 8: Robustness of Level-k Classification

% of subjects identified as the same level-k type	
Both vs. Replicate vs. Simplified	78%
Replicate: logit vs. spike-logit error	90%
Both: b=0 excluded vs. included	97%
Replicate: b=0 excluded vs. included	93%
Simplified: b=0 excluded vs. included	100%

5. Conclusion

We conduct a simplified version of Wang et al. (2010) and compare the results with a replicate one. The results of the replication are similar to that of Wang et al. (2010), while subject behavior in the Simplified Treatment is closer to the babbling equilibrium. In addition, more senders are classified as L2 in the Simplified Treatment, and such classifications are more robust. Similarly, more senders increase their message sophistication, moving to High (L2 or above) messages in the Simplified Treatment, likely because it is a simpler game.

Since the Replicate Treatment is relatively complicated, some senders might think their opponents have lower level of thinking and send less sophisticated messages. This has important theoretical implications, since in models of bounded rationality, the level of subject is decided by two factors: the maximum level one could infer, as considered in Camerer et al. (2004), and one's belief of the opponent sophistication, as in Costa-Gomes and Crawford (2006). In most applications, one cannot easily separate these two factors. In our data, the increase in sender levels is likely due to the change in beliefs about opponent levels in different environments.

In applications, overcommunication can play an important role in organization design. In particular, it mitigates conflict of interest, and substitutes for incentive alignment. Hence, it also substitute for centralized control in environments where incentive alignment and centralized authority are substitutes (Rantakari, 2008). Failure to take into account overcommunication could lead to excessive headquarter intervention or insufficient decentralization, reducing productivity and efficiency.

Reference

Alonso, Ricardo, Wouter Dessein and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review, 98(1): 145–179.*

Battaglini, Marco, Ernest K. Lai, Wooyoung Lim and Joseph Tao-yi Wang. 2016. "The Informational Theory of Legislative Committees: An Experimental Analysis." Working paper.

Battaglini, Marco and Uliana Makarov. 2014. "Cheap talk with multiple audiences: an experimental analysis." *Games and Economic Behavior*, 83: 147–164

Cai, Hongbin and Joseph Tao-yi Wang. 2006. "Overcommunication in Strategic Information Transmission Games." *Games and Economic Behavior*, 56(1), 7-36.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics*, 119(3): 861–98.

Costa-Gomes, Miguel A. and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 96(5): 1737-1768.

Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and

Boundedly Rational Misrepresentation of Intentions." *American Economic Review*, 93(1): 133–49.

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

Dessein, Wouter and Tano Santos. 2006. "Adaptive Organizations." *Journal of Political Economy*, 114(5): 956-995.

Dickhaut, John W. Kevin A. McCabe, and Arijit Mukherji. 1995. "An experimental study of strategic information transmission" *Economic Theory*, 6(6): 389-403.

Gneezy, Uri. 2005. "Deception: The Role of Consequences." *American Economic Review*, 95(1): 384–94.

Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10, 171–178.

Kawagoe, Toshiji and Hirokazu Takizawa. 2009. "Equilibrium refinement vs. level-k analysis: An experimental study of cheap-talk games with private information." *Games and Economic Behavior*, 66(1): 238-255.

Lai, Ernest K., Wooyoung Lim and Joseph Tao-yi Wang. 2015. "An Experimental Analysis of Multidimensional Cheap Talk." *Games and Economic Behavior*, 91, 114-144.

Minozzi, William and Jonathan Woon. 2016. "Competition, preference uncertainty, and jamming: A strategic communication experiment." *Games and Economic Behavior*, 96: 97–114.

Nagel Rosemarie. 1995. "Unraveling in Guessing Games: An experimental study." *American Economic Review*, 85(5): 1313-1326.

Rantakari, Heikki. 2008. "Governing Adaptation." *Review of Economic Studies*, 75(4): 1257-1285.

Serra-Garcia, Marta, Eric van Damme and Jan Potters. 2011. "Hiding an inconvenient truth: Lies and Vagueness." *Games and Economic Behavior*, 73(1): 244-261.

Sanchez-Pages, Santiago, and Marc Vorsatz. 2007. "An Experimental Study of Truth-Telling in a Sender-Receiver Game." *Games and Economic Behavior*, 61(1): 86-112.

Sanchez-Pages, Santiago, and Marc Vorsatz. 2009. "Enjoy the silence: an experiment on truth-telling." *Experimental Economics*, 12(2): 220-241.

Stahl, D. O., and Wilson, P. W. 1994. "Experimental evidence on players' models of other players." *Journal of Economic Behavior & Organization*, 25(3), 309–327.

Vespa Emanuel and Alistair J. Wilson. 2016. "Communication with Multiple Senders: An experiment." *Quantitative Economics*, 7(1): 1-36.

Wang, Joseph Tao-yi, Michael Spezio and Colin F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation To Understand Truth Telling and Deception in Sender-Receiver Games" *American Economic Review*, 100(3): 984-1007.

Online Appendix:

Table S1: Number of (Normal) Trials

Bias	0	1	2
Replicate Treatment	245	475	460
Simplified Treatment	542	1542	118

Note: Reverse-role "catch" trials excluded.

Table S2: Aggregate Sender Message With/without "Catch" Trials

	Without "Catch" Trials			With "Catch" Trials		
	M=1	M=2	M=3	M=1	M=2	M=3
Simplified Treatment ($b = 0$)						
S=1	178	0	0	179	0	0
S=2	0	181	0	0	182	0
S=3	0	0	183	0	0	187
Simplified Treatment ($b = 1$)						
S=1	107	133	311	121	145	348
S=2	35	81	385	42	97	434
S=3	37	44	409	49	62	450

Table S3: Aggregate Receiver Action With/without "Catch" Trials

	Without Catch Trials			With "Catch" Trials		
	A=1	A=2	A=3	A=1	A=2	A=3
Simplified Treatment ($b = 0$)						
M=1	177	1	0	188	1	0
M=2	0	180	1	0	186	1
M=3	0	0	183	0	0	196
Simplified Treatment ($b = 1$)						
	Without Catch Trials			With "Catch" Trials		
	A=1	A=2	A=3	A=1	A=2	A=3
M=1	123	52	4	143	64	5
M=2	140	112	6	167	130	7
M=3	167	739	199	180	818	234

Table S4: Information Transmission: Learning Results

Rounds	Bias	Corr(S, M)	Corr(M, A)	Corr(S, A)
Replicate Treatment				
1-10	0	0.984	0.998	0.982
11-20		0.998	0.988	0.994
1-10	1	**0.644**	0.761	0.521
11-20		**0.720**	0.714	0.542
1-10	2	0.381	0.488	**0.192**
11-20		0.404	0.524	**0.290**
Simplified Treatment				
Rounds	Bias	Corr(S, M)	Corr(M, A)	Corr(S, A)
1-20	0	1.000	0.995	0.995
21-40		1.000	1.000	1.000
1-20	1	0.260	0.439	0.127
21-40		0.219	0.404	0.104

Table S5: Average Payoffs of Senders and Receivers: Learning Results

Rounds	Bias	u_S	(std)	u_R	(std)
Replicate Treatment					
1-10	0	108.67	5.01	108.67	5.01
11-20		109.52	3.07	109.52	3.07
1-10	1	81.07	29.30	89.19	25.03
11-20		81.86	30.54	88.17	21.36
1-10	2	41.19	56.02	76.69	31.31
11-20		39.89	54.56	80.05	30.12
Simplified Treatment					
Rounds	Bias	u_S	(std)	u_R	(std)
1-20	0	99.79	2.49	99.79	2.49
21-40		100.00	0.00	100.00	0.00
1-20	1	53.23	40.52	75.19	24.90
21-40		58.12	38.01	76.52	22.84

Table S6: Level-k Robustness: Replicate Treatment

Level-k Type	All	L0	L1	L2
Resample 20 rounds				
95%	41%	38%	35%	46%
90%	7%	0%	9%	7%
80%	19%	25%	22%	14%
60%	19%	38%	9%	21%
<60%	15%	0%	26%	11%
Total(N)	100%(59)	100%(8)	100%(23)	100%(28)
Resample 36 rounds				
95%	51%	50%	48%	54%
90%	8%	0%	9%	11%
80%	14%	38%	9%	11%
60%	17%	13%	13%	21%
<60%	10%	0%	22%	4%
Total(N)	100%(59)	100%(8)	100%(23)	100%(28)

Table S7: Level-k Robustness: Simplified Treatment

Level-k Type	All	L0	L1	L2
Resample 20 rounds				
95%	51%	29%	13%	70%
90%	14%	0%	27%	11%
80%	10%	29%	7%	8%
60%	17%	14%	40%	8%
<60%	8%	29%	13%	3%
Total(N)	100%(59)	100%(7)	100%(15)	100%(37)
Resample 36 rounds				
95%	66%	29%	40%	84%
90%	5%	0%	7%	5%
80%	14%	43%	33%	0%
60%	8%	14%	7%	8%
<60%	7%	14%	13%	3%
Total(N)	100%(59)	100%(7)	100%(15)	100%(37)

Note: Total number of senders in parentheses.

Experimental Instructions (English Translation)

TASSEL Experimental Instruction p.1

Experimental Payment

At the end of the experiment, you will receive a show-up fee of NT$100, plus the amount of NT$ converted from the experimental standard currency (ESC) you receive throughout the experiment. The ESC you receive, which differs from person to person, is determined by your decision, other participants' decision, and chance. Each participant will be paid privately, and you are under no obligation to tell others how much you earned. **Note: The exchange rate between ESC and NT$ is 6:1. (6 ESC = NT$1.)**

Part 1

This is an experiment on group decisions among two individuals. Part 1 consists of 3 practice rounds and 20 paid rounds. Each group consists of two members, Member A and B. At the beginning of the experiment, you will be randomly assigned to be either A or B. Once decided, your role remains the same throughout the experiment. However, at the beginning of each round, the computer will randomly rematch participants to form new groups; thus, members in your group may not be the same each round.

1. Seeing the Number
   At the beginning of each round, member A will see a number (n) drawn from $\{1, 2, 3, 4, 5\}$ with equal probability. Only member A can see this number. In addition, the computer will assign a preference difference (d) between member A and B to be either 0, 1, or 2. The preference difference is public information that is known to both members, and determines payoffs (as shown below).
2. Sending a Message
   After seeing the number and preference difference, member A sends the message, "The number I see is ___ (1, 2, 3, 4, 5)". Note that "the number member A sees" need not be "the number member A sends."
3. Taking Action
   After seeing the message and preference difference, member B takes an action (a). The action is a number chosen from $\{1, 2, 3, 4, 5\}$.
4. Payoff Determination
   Depending on the preference difference, different payoff tables will show up on the left side of your decision screen, based on the following formula:

$n - a$	0	± 1	± 2	± 3	± 4		
B's payoff	110	90	57	17	-29		

$n + d - a$	0	± 1	± 2	± 3	± 4	± 5	± 6
A's payoff	110	90	57	17	-29	-80	-136

Note: In some cases (Example 1), member A's highest possible payoff is not 110.

Example 1: The number is 5 and preference difference is 2. If B takes Action 1,
$n - a = 4$, and member B's payoff is -29 ESC.
$n + d - a = 6$, and member A's payoff is -136 ESC.

If B takes Action 3,
$n - a = 2$, and member B's payoff is 57 ESC.
$n + d - a = 4$, and member A's payoff is -29 ESC.

If B takes Action 5,
$n - a = 0$, and member B's payoff is 110 ESC.
$n + d - a = 2$, and member A's payoff is 57 ESC.

Example 2: The number is 2 and preference difference is 0. If B takes Action 1,
$n - a = 1$, and member B's payoff is 90 ESC.
$n + d - a = 1$, and member A's payoff is 90 ESC.

If B takes Action 2,
$n - a = 0$, and member B's payoff is 110 ESC.
$n + d - a = 0$, and member A's payoff is 110 ESC.

If B takes Action 3,
$n - a =$ -1, and member B's payoff is 90 ESC.
$n + d - a =$ -1, and member A's payoff is 90 ESC.

TASSEL Experimental Instruction p.2

Part 2

Part 2 is also an experiment on group decisions among two individuals, and consists of 40 paid rounds. Each group consists of two members, Member A and B. At the beginning of the experiment, you will be randomly assigned to be either A or B. Once decided, your role remains the same throughout the experiment, **except for a few special rounds (so beware!)**. However, at the beginning of each round, the computer will randomly rematch participants to form new groups; thus, members in your group may not be the same each round.

1. Seeing the Number
   At the beginning of each round, member A will see a number (n) drawn from {1, 2, 3} with equal probability. Only member A can see this number. In addition, the computer will assign a preference difference (d) between member A and B to be either -1, 0, or 1. The preference difference is public information that is known to both members.

2. Sending a Message
   After seeing the number and preference difference, member A sends the message, "The number I see is ___ (1, 2, 3)".

3. Taking Action
   After seeing the message and preference difference, member B takes an action (a). The action is a number chosen from {1, 2, 3}.
   In the meantime, member A is asked to predict the action of member B.

4. Payoff Determination

$n - a$	0	± 1	± 2	
B's payoff	100	70	21	

$n + d - a$	0	± 1	± 2	± 3
A's payoff	100	70	21	-40

Member A earns an extra 3 ESC if s/he correctly predicts member B's action.

Experimental Instructions (in Chinese)

TASSEL 實驗說明 p.1

實驗報酬

本實驗結束後，你將得到定額車馬費新台幣 100 元，以及你在實驗中獲得的「法幣」所兌換成之新台幣。（「法幣」為本實驗的實驗貨幣單位。）你在實驗中能獲得的「法幣」會根據你所做的決策、別人的決策，以及隨機亂數決定，每個人都不同。每個人都會個別獨自領取報酬，你沒有義務告訴其他人你的報酬多寡。**請注意：本實驗中「法幣」與新台幣兌換匯率為 6:1。**(法幣 6 元=新台幣 1 元)

第一部分

第一部分為兩人一組的共同決策實驗，共有三個練習回合與二十個正式回合。每組有成員甲和成員乙兩人。在實驗一開始時，電腦會隨機決定你是成員甲還是成員乙。一旦決定之後，你的成員身份不會再變動。然而，每回合一開始時，電腦會將所有人打散重新隨機分組，因此，每次你遇到的成員並不一定相同。

1. 看到數字

每回合一開始，成員甲將會看到一個隨機抽取的數字，數字可能為{1、2、3、4、5}其中一個，每一個被抽到的機率皆相同，此數字只有成員甲可以看見。此外，電腦會指定甲、乙的差異為{0、1、2}其中之一，差異為公開資訊，成員甲跟乙都可以看見。差異會影響報酬，將會在報酬決定部份詳細說明。

2. 傳遞訊息

成員甲看到數字以及兩人差異後，選擇傳送訊息：「我看到____ (1、2、3、4、5)」。成員甲「看到的數字」與「他傳遞的訊息」不必然相同。

3. 做出決策

成員乙看到訊息、兩人差異後做出決策，決策為{1、2、3、4、5}其中之一。

4. 報酬決定

在實驗中，對應不同的差異，總共會有三種報酬表，將會在你做決策時出現在畫面的左側，以下為報酬表的公式：

數字－決策	0	±1	±2	±3	±4		
乙的報酬	110	90	57	17	－29		
數字＋差異－決策	0	±1	±2	±3	±4	±5	±6
甲的報酬	110	90	57	17	－29	－80	－136

請注意：成員甲的最大可能報酬在某些情況下並不是 110，如範例一。

範例一：數字為 5，兩人差異為 2，若乙選擇決策 1，

數字－決策 ＝ 4，乙的報酬為－29 法幣。

數字＋差異－決策 ＝ 6，甲的報酬為－136 法幣。

若乙選擇決策 3，

數字－決策 ＝ 2，乙的報酬為 57 法幣。

數字＋差異－決策 ＝ 4，甲的報酬為－29 法幣。

若乙選擇決策 5，

數字－決策 ＝ 0，乙的報酬為 110 法幣。

數字＋差異－決策 ＝ 2，甲的報酬為 57 法幣。

範例二：數字為 2，兩人差異為 0，若乙選擇決策 1，

數字－決策 ＝ 1，乙的報酬為 90 法幣。

數字＋差異－決策 ＝ 1，甲的報酬為 90 法幣。

若乙選擇決策 2，

數字－決策 ＝ 0，乙的報酬為 110 法幣。

數字＋差異－決策 ＝ 0，甲的報酬為 110 法幣。

若乙選擇決策 3，

數字－決策 ＝ －1，乙的報酬為 90 法幣。

數字＋差異－決策 ＝ －1，甲的報酬為 90 法幣。

TASSEL 實驗說明 p.2

第二部分

第二部份也是兩人一組的共同決策實驗，共有四十個回合。每組有成員甲和成員乙兩人。在實驗一開始時，電腦會隨機決定你是成員甲還是成員乙。一旦決定之後，**在大部分回合你的成員身份不會再變動，僅有少數例外，請留意。**另外，每回合一開始時，電腦會將所有人打散重新隨機分組，因此，每次你遇到的成員並不一定相同。

1. 看到數字

每回合一開始，成員甲將會看到一個隨機抽取的數字，數字可能為{1、2、3}其中一個，每一個被抽到的機率皆相同，此數字只有成員甲可以看見。此外，電腦會指定甲、乙的差異為{－1、0、1}其中一個，差異為公開資訊，成員甲跟乙都可以看見。

2. 傳遞訊息

成員甲看到數字以及兩人差異後，選擇傳送訊息：「我看到____ (1、2、3)」。

3. 做出決策

成員乙看到訊息、兩人差異後做出決策，決策為{1, 2, 3}其中之一。

在成員甲傳送訊息後，將會請成員甲預測成員乙的決策。

4. 報酬決定

數字－決策	0	±1	±2	
乙的報酬	100	70	21	
數字＋差異－決策	0	±1	±2	±3
甲的報酬	100	70	21	－40

成員甲成功預測乙的決策則可獲得額外的 3 法幣。