An Experimental Study On the Impact of Communication in Outguessing Games*

Sean Inoue
Working Paper

June 14, 2019

Abstract

Communication introduces new ways in which level-k may matter in outguessing games. Prior to playing an outguessing game, a sender sends a message to a receiver stating that the sender will play a specific action. It is predicted that the message causes players to behave according to the basic model of level-k presented in Crawford (2003): Level-0 senders are truthful and level-0 receivers believe level-0 senders. Level-k senders best respond to level-(k − 1) receivers. Level-k receivers believe that level-k senders are truthful. Subjects play five periods of the game, anonymously and randomly playing against the field in each period. This design is utilized to analyze how experience impacts a subject's level. The experiment finds that level-0 play is common in the first period of play, but vanishes almost completely by the last period, indicating that subjects do not have a stable preference for truth-telling. Play is mostly focused on levels 0, 1, and 2. The number of unidentifiable players rises over time, indicating that players play more complex strategies as they grow experienced.

JEL: C91
Keywords: *cheap-talk, level-k, experimental economics*

*Thank you to Andreas Blume, Martin Dufwenberg, and Charles Noussair for much appreciated feedback. Thanks doubly to Charles Noussair and the Experimental Science Laboratory at the University of Arizona for providing the funds and resources to carry out the experiment.

1 Introduction

In competitive settings, players sometimes use cheap talk as a part of their strategy. In the title match of the 2007 USA Rock Paper Scissors championship, prior to the first throw, David Borne said “let’s roll” to his opponent Jamie Langridge. The commentators noted that this statement from Borne was calling for his opponent to play Rock. Borne continued to make statements before throws during that match. Throughout the 2006 World Series of Poker Main Event, the eventual winner Jamie Gold would frequently talk to his opponents during hands. Gold would sometimes tell the complete truth about his privately held cards as a part of his strategy. In the book “Caro’s Book of Poker Tells” by Mike Caro, one section is devoted to the poker adage “weak means strong” while another section is devoted to “strong means weak.” These sections emphasize that a poker player acts like they have a weak hand when they frequently hold strong cards, while players who act like they have strong hands typically hold weak cards.

When communication is feasible, Crawford (2003) proposes that players anchor on truth. This experiment tests this effect. It is predicted that with communication, subjects play according to the level-k model presented in Crawford (2003): Level-0 senders send truthful messages and level-0 receivers believe messages. Level-k senders best respond to level-(k $-$ 1) receivers, and level-k receivers believe that a level-k sender is telling the truth about their action. Typically it is assumed by many papers in the literature that a majority of players will be level-1 or level-2 players. The experimental design allows for levels 0 through 5 to be measured, assuming that subjects are not on level 6 or above. Additionally, in each period, all senders will be randomly paired against all receivers. This means that in each period, each subject will make between five and eight choices, which provides high confidence on whether subjects play a specific level.

The goal of this paper is to show the impact of communication in outguessing games. This is accomplished in two ways. First, play with communication is compared to play without communication. This paper will compare the patterns of play to assess if communication has an impact. Secondly, level-k theory will be tested. This will provide a deeper exploration into the exact impact of pre-play communication.

Many experimental studies examine level-k models in games involving communication. Cai & Wang (2007), in their experiment testing Crawford & Sobel (1982), use level-k with an honest level 0 as an explanation for subject play. Sánchez-Pagés and Vorsatz (2007), and Holm and Kawagoe (2010) examine communication in conjunction with games that are similar to matching pennies, but in a private-information setting. Kawagoe & Takizawa

(2009) provide additional evidence that supports level-k as a viable explanation of behavior in communication games. These papers examine cases of private information, whereas this project focuses on level-k in a complete information setting.

Other papers involving communication center around the idea that players are biased toward truth-telling. In this paper, this bias equates to seeing a plethora of level-0 play. Blume et al. (2001) and Cai & Wang (2007) note that subjects overcommunicate. Rode (2006) shows that subjects sometimes communicate truthfully against their benefit. Charness & Dufwenberg (2005), Gneezy (2005), Sutter (2009), and Hurkens & Kartik (2009) examine this bias in more detail and attempt to analyze why and how subjects may be truth-biased.

Two theory papers inspired this project. These papers are inspired by seminal papers by Farrell (1987, 1988), which analyze equilibria where, due to the use of pre-play communication, actions can be correctly inferred by players. Ellingsen and Östling (2010) examine all 2x2 games and use a level-k model to identify on all classes of games where communication can help or hurt coordination on Nash equilibria. Crawford (2003) uses the level-k model to explain the events of D-Day, which he models as a zero-sum game between the Axis and Allies. The primary contribution of this paper is an experiment that can specifically test whether people use level-k, with truth-telling being a level 0 and where there is no other reasonable behavioral explanation for observing that type of play. Additionally, tests can determine if level-k correlates with individual characteristics and if players' levels change over time.

This paper contributes to the growing literature on the general methodology of level-k and the measurement of level-k in a setting with communication. Arad and Rubinstein (2012) propose the 11-20 game as a test of level-k. Georganas et al. (2015) show that level-k is not the same for the same subjects across two families of games. Heap et al. (2014) use a wide variety of framing effects to test level-k hypotheses in hide-and-seek and coordination games, which they do not find evidence to support.

Many other papers analyze games in a laboratory setting and then explain behavior using level-k. These include Stahl and Wilson (1995), Nagel (1995), Ho et al. (1998), Costa-Gomes et al. (2001), Camerer et al. (2004), Crawford and Iriberri (2007), and many others.

2 Experimental Setup and Predictions

In this experiment, two players, a sender and a receiver, will play the simultaneous move game in Figure 1. The sender has a strict preference for the state matching, while the receiver has a strict preference for a specific mismatching of the state. In the communication

		Receiver					
		A	B	C	D	E	F
Sender	A	$1,0$	$0,1$	$0,0$	$0,0$	$0,0$	$0,0$
	B	$0,0$	$1,0$	$0,1$	$0,0$	$0,0$	$0,0$
	C	$0,0$	$0,0$	$1,0$	$0,1$	$0,0$	$0,0$
	D	$0,0$	$0,0$	$0,0$	$1,0$	$0,1$	$0,0$
	E	$0,0$	$0,0$	$0,0$	$0,0$	$1,0$	$0,1$
	F	$0,1$	$0,0$	$0,0$	$0,0$	$0,0$	$1,0$

Table 1: The Basic Game Form

treatment, the sender must send a message m to the receiver,[1] where m states "I will take action X[2]" for some $X \in \mathbf{A} = \{A, B, C, D, E, F\}$.

2.1 Predictions Under No Communication

This game has a unique mixed-strategy Nash equilibrium in which each player randomizes uniformly over all pure strategies. To see that this is the only Nash equilibrium, first note that there are no pure-strategy equilibria, meaning that only mixed strategies must be considered. Additionally, due to the cyclical nature of this game, the following notation will be used: Let the actions have the cyclic order (A, B, C, D, E, F). Note that, for an arbitrary action $X \in \mathbf{A}$, $X(N)$ is the action X shifted N actions forward in the cycle. This means, for example, that $X(6) = X$.

Suppose the sender plays a mixed strategy involving exactly $N = 2$ pure strategies. Denote those strategies X and Y. Because the best response to a pure strategy is unique, each pure strategy has a different best response, and each best response has the same payoff, the sender makes the receiver indifferent by placing equal probability on X and Y. Mixing uniformly between X and Y makes the receiver indifferent between two pure strategies $X(1)$ and $Y(1)$. However, for any receiver strategies $X(1)$ and $Y(1)$, the only way to make the sender indifferent between two best responses is to mix uniformly between $X(1)$ and $Y(1)$, which makes the sender indifferent between pure strategies $X(1)$ and $Y(1)$. If a mixed-

[1]The experiment uses symbols in place of letters to mitigate unconscious bias towards certain letters. In the text, using letters is more convenient and illustrative.

[2]The instructions state "I will take action X," while the Z-tree program states "I will select action X." Both statements have similar strength of meaning, so it is not expected to impact results, but it is noted here for complete disclosure.

strategy Nash equilibrium exists, both players must best respond to each other, which implies that $X(1) = Y$ and $Y(1) = X$ hold. This is a contradiction, since $X(2) = X(N) \neq X$.

This same logic applies to mixtures involving $N < 6$ pure strategies. A mixed-strategy equilibrium can only exist when $X(N) = X$, which only holds for N=6. That mixed strategy is a uniform mixture across all possible strategies.

Prediction 1: With no communication, subjects will play according to the unique mixed-strategy Nash equilibrium: A uniform distribution of actions will be observed.

2.2 Predictions with Pre-play Communication

With communication, the sender must send a message to the receiver. This message is a statement that the sender intends to play a specific action. After the message is sent, both players simultaneously decide on an action to take. Note that because the equilibrium of the no-communication game is unique, because communication does not change the set of payoffs, and because this is a game of complete information, communication does not change the set of Nash equilibrium outcomes.

However, communication may have an impact on players who are level-k thinkers. In particular, consider a level-k model in which a level-0 sender is assumed to always be telling the truth and a level-0 receiver is assumed to always believe the sender's message. A level-1 sender then best responds to a level-0 receiver, while a level-1 receiver believes that a level-1 sender tells the truth. This is a model used in a multitude of papers that originates from Crawford (2003). Under this level-k model, suppose that the sender plays the message "I will take action A." Play would then proceed according to table 2 depending on the level of the players playing the game. It is commonly observed in other level-k studies (Cai & Wang (2007), Georganas et al. (2015)) that the majority of the players are level-1 or level-2, with few level-0 and few level-3 players.

Prediction 2: With communication, suppose that message "I will take action X" is sent. A majority of senders will play actions $X(1)$ and $X(2)$, and a majority of receivers will play actions $X(2)$ and $X(3)$.

If it is true that subjects are level-k thinkers, subjects should see higher average payoffs than the Nash equilibrium payoff, which is $\frac{1}{6}$. In this game, subjects are rewarded for figuring out exactly what the other player is doing. It follows that without communication, subjects should not see higher average payoffs than the Nash equilibrium.

Prediction 3: With communication, subjects' payoffs will be higher than the Nash equilibrium payoff. Without communication, subjects' payoffs will not be higher than the Nash equilibrium payoff.

The communication level-k model is unique in that play is heavily influenced with communication. It is possible that without communication, a player playing an outguessing game can behave in a similar manner. Define the no-communication level-k model as follows: Suppose that a level-0 sender is naturally drawn to some specific action, for instance, A. If receivers believe that senders are drawn to action A, a level-0 receiver should always play action B. If senders then believe that receivers believe that senders are drawn to action A, a level-1 sender should play action B. This logic identically replicates the pattern of play produced with communication, illustrated in Table 2. Note that the no-communication level-k model is fairly robust to different initial level-0 sender beliefs: As long as players in the game believe that a level-0 sender plays A more frequently than any other actions, a level-0 receiver's unique best response is B.

In this setting, players should not play according to the no-communication level-k model when there is no communication. With communication, a focal belief can be formed. Without communication, any level-0 assumption is arbitrary, and could only happen if there was something distinct about one specific action. It is predicted that communication is the force that causes this theory of level-k to be observed.

Prediction 4: Without communication, play will not be consistent with the no-communication level-k model.

Additionally, little is known about how cognitive ability correlates with level hierarchies, and about how level-k play evolves over time. Subjects' Cognative Reflection Test (CRT) score is measured. I will then test whether CRT score or experience has an influence on level of play.

Prediction 5: As CRT score increases, subjects will play higher levels.

Prediction 6: Subjects' level will increase as they play more periods of the game.

3 Experimental Design

Subjects completed the task at the Experimental Science Laboratory at the University of Arizona. The experiment was coded in z-Tree [16]. In the experiment, actions were labeled $\{\#, \%, \hat{}, +, ^*, (\}$ in place of $\{A, B, C, D, E, F\}$ respectively. Subjects participated in a total of eight sessions–four for each treatment. The no-communication treatment had a total of 54 subjects–27 senders and 27 receivers. In this treatment, senders were called "row players" and receivers were called "column players." The communication treatment also had a total of 54 subjects–27 senders and 27 receivers. Each session had between 10 and 16 people. Subjects played five periods, with roles fixed as either the sender or the receiver across all

Table 2: How level-k players best respond to the message "I will take action A" assuming that a level-0 Sender is truthful and that a level-0 Receiver believes a level-0 Sender

Level	Sender Action	Receiver Action
Level-0	A	B
Level-1	B	C
Level-2	C	D
Level-3	D	E
Level-4	E	F
Level-5	F	A
Level-6	A	B
⋮	⋮	⋮

five. Within every period, each sender played the stage game with each receiver in the room. Matching was done randomly and anonymously. Subjects were paid for one play of the stage game chosen randomly at the end of the experiment. Each subject earned $8 per ECU, along with a show-up fee of $6. Additionally, in the communication treatment, subjects were asked three Cognitive Reflection Test (CRT) questions, for which they earned $1 for each correct answer. The experiment took approximately 30 minutes for the no-communication treatment and approximately 40 minutes for the communication treatment.

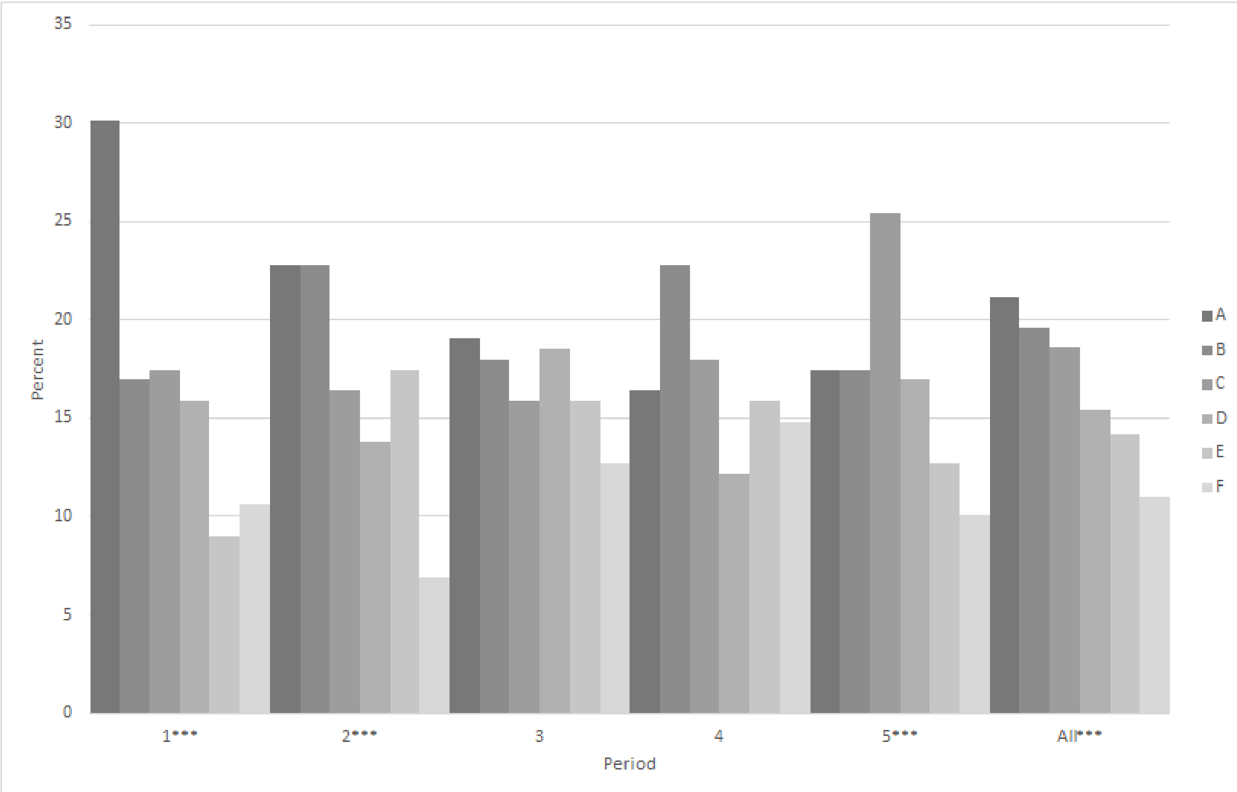
4 Results

4.1 No Communication

Without communication, Figure 1 displays the total actions that senders took, while Figure 2 displays the total actions that receivers took. Also displayed are the percentage of actions in each of the five periods, as well as the percentage of actions across all periods. It is noteworthy that the distribution does not resemble a uniformly random distribution. For senders, periods three and four do not reject the null hypothesis. For receivers, periods four and five do not reject the null hypothesis. When pooling the last three periods together, senders are different from a uniform distribution at the 5% level using a Pearson's Chi-squared ($\chi^2 = 13.33$), while the null hypothesis cannot be rejected for receivers ($\chi^2 = 8.88$).

Overall, there appears to be some adjusting and learning that points in the direction of random play, but this is not uniformly true. In particular, both senders and receivers play

Figure 1: Sender Actions No Communication

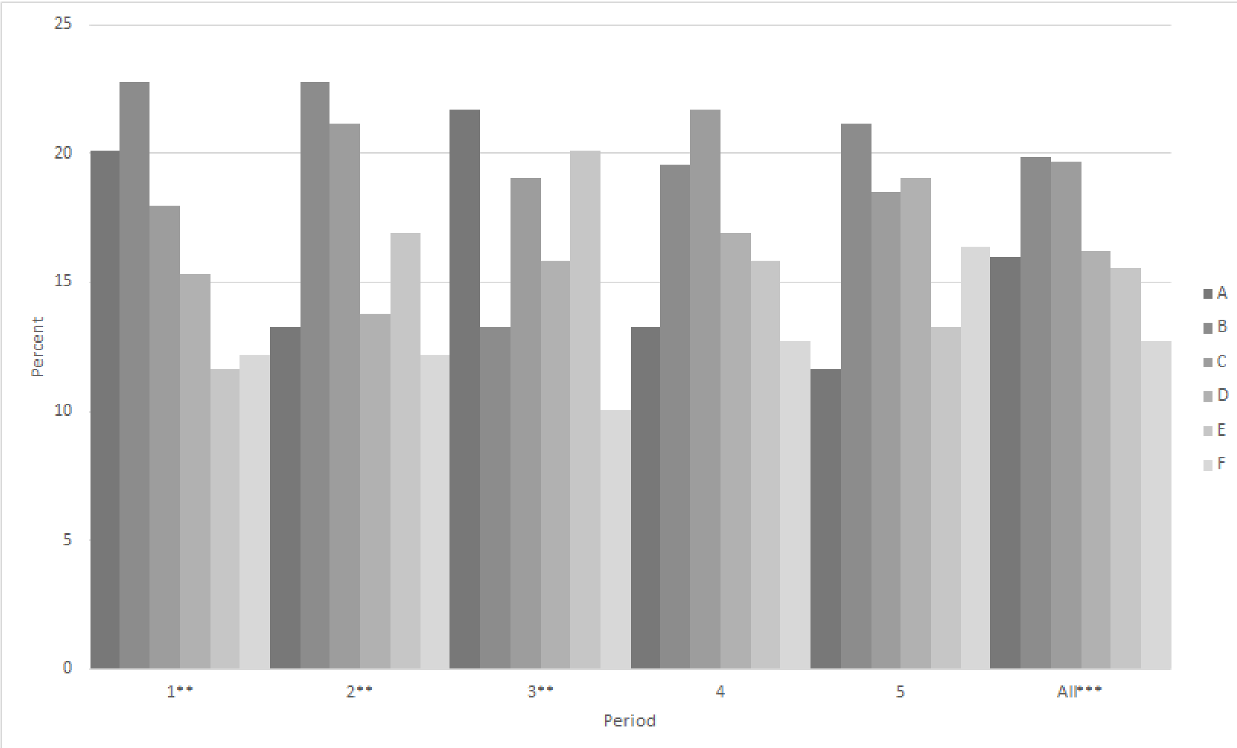


*,**, and *** indicate a significant difference to a uniform distribution of actions at the 10%, 5%, and 1% levels using a Pearson's chi-squared test.

action F (called '(' in the experiment) less frequently than other actions. This then leads to a follow up question: Does this biasing of actions impact observed play in experiments?

In this game, if a player has a belief that their opponent plays an action a higher percentage of the time than any other action, the best response is a pure strategy. If the distribution of actions has an impact on the way subjects choose actions in the no-message treatment, subjects should play a pure strategy in response. The distributions of senders and receivers who take the same action over 50% of the time in each period are displayed in Tables 6 and 7. In any given period, over half of subjects are not playing the same action more than 50% of the time. Additionally, very few subjects, if any, seem to be taking advantage of this bias. If senders note that receivers are biased toward picking B, senders should pick B more often in subsequent periods, which is not observed. If receivers know that senders are biased toward picking A, receivers should pick B more, which is also not observed. Thus,

Figure 2: Receiver Actions No Communication



*, **, and *** indicate a significant difference to a uniform distribution of actions at the 10%, 5%, and 1% levels using a Pearson's chi-squared test.

the distribution not resembling the Nash equilibrium appears to be caused by unconscious bias.

Result 1: Prediction 1 is not supported. Subjects do not randomize over all actions uniformly without communication. This is attributed to an unconscious bias away from action F.

4.2 Communication

With communication it is predicted that play proceeds according to Crawford (2003): level-0 senders match the action with the message, level-k receivers believe level-k senders and respond accordingly, and level-k senders best respond to level-(k − 1) receivers. The appendix displays the distribution of messages, the overall distribution of actions, and the overall distribution of levels for senders and receivers.

To accurately interpret overall play, level-k is measured at an individual level using a 50%

Table 3: Communication Sender Levels Using > 50% Threshold

	Period 1	2	3	4	5
0	10/27	4/27	5/27	4/27	3/27
1	8/27	13/27	6/27	8/27	4/27
2	3/27	3/27	7/27	6/27	6/27
3	0/27	0/27	0/27	1/27	1/27
4	1/27	0/27	0/27	0/27	0/27
5	0/27	4/27	2/27	1/27	2/27
No Level	5/27	3/27	7/27	7/27	11/27

Table 4: Communication Receiver Levels Using > 50% Threshold

	Period 1	2	3	4	5
0	16/27	10/27	6/27	5/27	5/27
1	6/27	11/27	12/27	10/27	8/27
2	0/27	0/27	0/27	3/27	3/27
3	1/27	0/27	0/27	0/27	0/27
4	1/27	1/27	1/27	1/27	0/27
5	1/27	1/27	1/27	1/27	0/27
No Level	2/27	4/27	7/27	7/27	11/27

threshold: If, in a given period, over half of a player's actions are consistent with a level of play, that player is classified as that level. This threshold has a high degree of accountability in this game, since for any message there is only a single action that is consistent with a specific level.

The number of subjects classified in each level in each period of play is displayed in Tables 3 and 4. In the first period, over half of receivers and over a third of senders are consistent with level-0 play. Additionally, many senders are consistent with level-1 play. In the second period, level-0 play diminishes greatly for both senders and receivers. In period two, almost half of senders play consistently with level 1, while receivers mostly play consistently with level 0 and level 1. By period five, over a third of senders and receivers

Table 5: Among No Level Subjects, How Many Subjects Randomize

	Period				
	1	2	3	4	5
Sender	1/5	2/3	5/7	6/7	7/11
Receiver	1/2	3/4	5/7	3/7	4/11

do not play consistently with a specific level. Additionally, subjects are spread out between levels 0, 1, and 2.

One main takeaway from the data is that, surprisingly, level-0 prevails in the first period of play but then diminishes significantly as subjects gain experience. This strongly indicates that subjects are initially biased to be honest. This bias changes with experience, however, and level-k appears to explain more (but not all) of the data as subjects play more. Additionally, receivers are primarily level-0 and level-1, while senders are more split between levels 0, 1, and 2. This could be due to differences between how senders take actions and how receivers take actions.

Another surprising result is the emergence of play that is not classifiable as any level. As subjects become more experienced, subjects tend to differentiate their play more. It begs the question of whether some subjects are becoming sophisticated over time by playing truly randomly or whether subjects are varying the level they are playing on within the same period. To test this, Table 5 lists the number of subjects in each period that are "randomizing." In any given period, subjects whose play is consistent with at most three consecutive levels are considered to be not random, while all other subjects are considered to be random. For example, this means that in the last period, 7 out of 11 receivers only played actions consistent with three consecutive levels, while 4 out of 11 receivers did not play actions that were consistent with at most three consecutive levels. Analyzing this data, it is evident that there is a split between subjects who play mixtures of levels and subjects who play randomly.

One notable oddity in the data is the presence of players who are on level 5. While there are not many of these players, it is worth discussing why such players may exist. A player who is altruistic toward honest players would be incentivized to play level 5. If a portion of players want someone to get a payoff but also want to reward honesty, this explains why a fraction of the subject pool would continue to play an action that is highly unlikely to get them a payment.

Overall, neither prediction fully explains the data. In the first period, there is a large bias toward honesty, but the remainder of the subjects are mostly classified on one level or another. As subjects play more, the honesty bias largely disappears, but many subjects start either to randomize or to mix between levels. This suggests that the way in which communication influences thinking is more complex than the literature gives it credit for. Overall, players do anchor their play on senders potentially sending honest messages. However, the distribution of levels in the last period of the game is far different than the literature would predict.

Result 2: Predction 2 is not supported. In the first period, many subjects' actions are influenced by a bias for truth-telling. This bias goes away over time. Level-k explains a majority of the data across all periods, but the distribution of levels is different than the literature suggests.

Table 6: Senders Who Play Same Action Using > 50% Threshold No Communication

	Period				
Action	1	2	3	4	5
A	5/27	4/27	2/27	1/27	2/27
B	3/27	4/27	2/27	3/27	2/27
C	1/27	1/27	1/27	1/27	4/27
D	1/27	1/27	1/27	0/27	1/27
E	0/27	2/27	0/27	1/27	1/27
F	1/27	0/27	1/27	1/27	0/27
Mixture	15/27	20/27	20/27	19/27	15/27

Table 7: Receivers Who Play Same Action Using > 50% Threshold No Communication

	Period				
Action	1	2	3	4	5
A	3/27	0/27	3/27	1/27	0/27
B	2/27	3/27	0/27	2/27	2/27
C	1/27	1/27	1/27	2/27	1/27
D	0/27	0/27	0/27	0/27	1/27
E	0/27	1/27	1/27	1/27	1/27
F	0/27	1/27	0/27	0/27	1/27
Mixture	21/27	20/27	20/27	20/27	19/27

Table 8: Average Earnings

		Period					
		1	2	3	4	5	Overall
Communication	Sender	0.294***	0.262***	0.278***	0.246***	0.246***	0.265***
		(n = 187, .389)	(n = 187, .441)	(n = 187, .449)	(n = 187, .432)	(n = 187, .432)	(n = 935, .441)
	Receiver	0.267***	0.337***	0.224**	0.257***	0.193	0.256***
		(n = 187, .444)	(n = 187, .474)	(n = 187, .419)	(n = 187, .438)	(n = 187, .395)	(n = 935, .436)
No Communication	Sender	0.185	0.196	0.148	0.180	0.196	0.181
		(n = 189, .389)	(n = 189, .398)	(n = 189, .356)	(n = 189, .385)	(n = 189, .398)	(n = 945, .385)
	Receiver	0.148	0.180	0.148	0.153	0.138	0.153
		(n = 189, .356)	(n = 189, .385)	(n = 189, .356)	(n = 189, .361)	(n = 189, .345)	(n = 945, .360)

*,**, and *** indicate a significant difference from the Nash equilibrium payoff of 0.166 at the 10%, 5%, and 1% levels using a one-tailed t-test. Sample size and standard errors are in parenthesis.

Looking at Table 8, it is clear that subjects do better with communication than without communication. With communication, in all periods for senders and in all but the last period for receivers, the payoff is statistically different from the Nash equilibrium payoff at the 1% level using a one-tailed t-test. Without communication, in all periods for all roles payoffs are not statistically different from the Nash equilibrium payoff using a one-tailed t-test. This indicates that communication is allowing players to coordinate their actions more frequently than in the absence of communication.

Result 3: Prediction 3 is supported. Almost all individual periods of play earn significantly more with communication than the Nash equilibrium, while without communication subjects never earn significantly more than the Nash equilibrium.

To see that play without communication cannot be explained by the no-communication level-k model, displayed in Tables 6 and 7 are the number of subjects that play the same action more than 50% of the time in a given period. If play proceeded according to the no-communication level-k model, subjects should always play the same action. Allowing for some error, a majority of the subjects do not play the same action in each period. This means that play cannot be explained by this particular model of level-k.

Table 9: OLS Regression of Period and CRT on Level Using Level 0 to Level 3 Only

Variable	Subject Type	
	Sender	Receiver
Intercept	0.232 (0.194)	0.121 (0.137)
CRT	0.652*** (0.167)	0.518*** (0.123)
Period2	0.262 (0.228)	0.107 (0.175)
Period3	0.408* (0.234)	0.229 (0.183)
Period4	0.498** (0.231)	0.394** (0.184)
Period5	0.659** (0.252)	0.398** (0.190)
Degrees of Freedom	86	90
Adjusted R^2	0.179	0.2

*,**, and *** indicate p values that are significant at the 10%, 5%, and 1% levels. Standard errors are in parenthesis.

Result 4: Prediction 4 is supported. The distribution of actions without communication cannot be explained by the no-communication level-k model. This implies that communication fundamentally changes the way that subjects take actions.

To analyze the impact of CRT scores and experience, I utilize the following linear regression, separated for senders and receivers:

$$Level_{it} = \alpha + CRT_i + Period2 + Period3 + Period4 + Period5 + \epsilon_{it}$$

where $Level_{it} \in \{0, 1, 2, 3\}$ is the level of player i in period t, CRT is a dummy variable that is 1 if the CRT score of individual i is 2 or 3 and is 0 if the CRT score of individual i is 0 or 1, and Period$t$ is the tth period in the game. Because of the presence of players with no level, and because play above level-3 is largely absent, the sample is restricted to periods in which play was on levels 0 through 3. This means that the data set resembles an unbalanced panel. Note that individual fixed effects are not accounted for in this model, as they are highly correlated with the CRT score.

The results of this regression are in Table 9. The CRT variable is significant at the 1% level for both senders and receivers. Additionally, periods 4 and 5 are significant at the 5% level for senders and receivers, while period 3 is significant at the 10% level for senders. Overall, the regression results indicate that later periods have an impact on the levels that subjects choose, and that CRT is heavily linked to the level of play.

Result 5: Predictions 5 and 6 are supported. Both CRT and time have a strongly positive impact on the selected level of play when restricting the sample to levels 0-3.

4.3 Discussion of Increase in No-Level Play

Table 10: OLS Regression of Period and CRT on No Level

Variable	Subject Type	
	Sender	Receiver
Intercept	0.274^{***} (0.096)	0.135 (0.090)
CRT	-0.133^{*} (0.077)	-0.10341 (0.072)
Period2	-0.074 (0.115)	0.074 (0.112)
Period3	0.074 (0.115)	0.185 (0.112)
Period4	0.074 (0.115)	0.185 (0.112)
Period5	0.222^{*} (0.115)	0.333^{***} (0.112)
Degrees of Freedom	129	129
Adjusted R^2	0.038	0.052

*,**, and *** indicate p values that are significant at the 10%, 5%, and 1% levels. Standard errors are in parenthesis.

One observation about the data that is surprising is that over time there appears to be an increase in play that is not classified as any particular level. To test the impact of experience on whether a subject is classified as no level, the following regression is utilized:

$$NoLevel_{it} = \alpha + CRT_i + Period2 + Period3 + Period4 + Period5 + \epsilon_{it}$$

where $NoLevel_{it}$ is a dummy variable that equals 1 when a subject is classified as 'No Level', CRT is a dummy variable that is 1 if the CRT score of individual i is 2 or 3 and is 0 if the CRT score of individual i is 0 or 1, and Period$t$ is the tth period in the game. The results of the regression are displayed in Table 10. The last period of play has a significant impact for both players on whether a subject is a level or not, while CRT has a small significant negative impact for the Sender and no significant impact for the receiver. Although time does seem to impact whether subjects are classifiable as some level, it is worth noting that the adjusted R^2 is incredibly small.

5 Concluding Remarks

The analysis in this paper shows that in outguessing games, communication fundamentally impacts how players make decisions. Additionally, it utilizes a unique experimental approach to allow for an accurate measurement of level-k and to study the effects of experience. Overall, it is clear that level-k explains some of the behavior in this experiment, but not all of it. Additionally, it is shown that another independent measure of cognitive ability, the CRT score, is strongly linked to higher levels of play.

The results of this paper demonstrate two key issues regarding level-k. Firstly, play tends to be very honest and the beginning of the experiment, and that honesty disappears quickly as subjects gain experience. This indicates that for initial plays of a game with communication, Gneezy (2005) may be correct in assessing that senders may initially feel guilty about lying. It also seems to be the case, as is demonstrated in other communication experiments, that receivers are more heavily truth biased. This bias disappears more slowly for the receivers than for the senders. It is mildly surprising that players behave this way in a game with no mildly equitable outcome, which indicates that this is a strong bias that should be taken seriously across all communication games.

Secondly, play changes over time. Players play more complex strategies in the sense that play is less likely to fall into the bucket of a specific level over time. Because there are six actions, a 50% threshold on being a specific level is a strong indicator of a level, so it is fascinating that players diversify their level of play more over time. This is counter to the intuition level-k provides, as that theory would predict cycles of play where levels continuously increase.

References

[1] Arad, A. and Rubinstein, A., 2012. The 11-20 money request game: A level-k reasoning study. American Economic Review, 102(7), pp.3561-73.

[2] Blume, A., DeJong, D.V., Kim, Y.G. and Sprinkle, G.B., 2001. Evolution of communication with partial common interest. Games and Economic Behavior, 37(1), pp.79-120.

[3] Cai, H., and Wang, J., 2006. Overcommunication in strategic information transmission games. Games and Economic Behavior 56.1: 7-36.

[4] Camerer, C.F., Ho, T.H. and Chong, J.K., 2004. A cognitive hierarchy model of games. The Quarterly Journal of Economics, 119(3), pp.861-898.

[5] Caro, M., 2003. *Caro's book of poker tells*. Cardoza Publishing.

[6] Charness, G. and Dufwenberg, M., 2006. Promises and partnership. Econometrica, 74(6), pp.1579-1601.

[7] Costa-Gomes, M., Crawford, V.P. and Broseta, B., 2001. Cognition and behavior in normal-form games: An experimental study. Econometrica, 69(5), pp.1193-1235.

[8] Crawford, V.P., 2003. Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. American Economic Review, 93(1), pp.133-149.

[9] Crawford, V.P., Costa-Gomes, M.A. and Iriberri, N., 2013. Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. Journal of Economic Literature, 51(1), pp.5-62.

[10] Crawford, V.P. and Iriberri, N., 2007. Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?. Econometrica, 75(6), pp.1721-1770.

[11] Crawford, V. and Sobel, J., 1982. Strategic information transmission, Econometrica, vol. 50, pp. 1431–51.

[12] Ellingsen, T. and Östling, R., 2010. When does communication improve coordination?. American Economic Review, 100(4), pp.1695-1724.

[13] Farrell, J., 1987. Cheap talk, coordination, and entry. The RAND Journal of Economics, pp.34-39.

[14] Farrell, J., 1988. Communication, coordination and Nash equilibrium. Economics Letters, 27(3), pp.209-214.

[15] Fehr, E., Klein, A., Schmidt, K., 2007, Fairness and Contract Design. Econometrica 75, 121–154.

[16] Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10(2), 171-178.

[17] Georganas, S., Healy, P.J. and Weber, R.A., 2015. On the persistence of strategic sophistication. Journal of Economic Theory, 159, pp.369-400.

[18] Gneezy, U., 2005. Deception: The role of consequences. American Economic Review, 95(1), pp.384-394.

[19] Hargreaves Heap, S., Rojo Arjona, D. and Sugden, R., 2014. How Portable Is Level-0 Behavior? A Test of Level-k Theory in Games With Non-Neutral Frames. Econometrica, 82(3), pp.1133-1151.

[20] Ho, T.H., Camerer, C. and Weigelt, K., 1998. Iterated dominance and iterated best response in experimental "p-beauty contests". The American Economic Review, 88(4), pp.947-969.

[21] Hurkens, S. and Kartik, N., 2009. Would I lie to you? On social preferences and lying aversion. Experimental Economics, 12(2), pp.180-192.

[22] Kawagoe, T. and Takizawa, H., 2009. Equilibrium refinement vs. level-k analysis: An experimental study of cheap-talk games with private information. Games and Economic Behavior, 66(1), pp.238-255.

[23] Nagel, R., 1995. Unraveling in guessing games: An experimental study. The American Economic Review, 85(5), pp.1313-1326.

[24] Rode, J., 2010. Truth and trust in communication: Experiments on the effect of a competitive context. Games and Economic Behavior, 68(1), pp.325-338.

[25] Sánchez-Pagés, S. and Vorsatz, M., 2007. An experimental study of truth-telling in a sender–receiver game. Games and Economic Behavior, 61(1), pp.86-112.

[26] Stahl, D.O. and Wilson, P.W., 1995. On players' models of other players: Theory and experimental evidence. Games and Economic Behavior, 10(1), pp.218-254.

[27] Sutter, M., 2009. Deception through telling the truth?! Experimental evidence from individuals and teams. The Economic Journal, 119(534), pp.47-60.

[28] Youtube.com. 2007 USARPS Title Match, Uploaded by User 'usarpsleague' on Oct. 8, 2007.

[29] Youtube.com. Live Poker Training Video: an analysis of Jamie Gold, Uploaded by User 'ReadingPokerTells' Jan 17, 2015.

Appendix A: Figures of aggregate play

Figure 3: Messages Sent in Communication Treatment

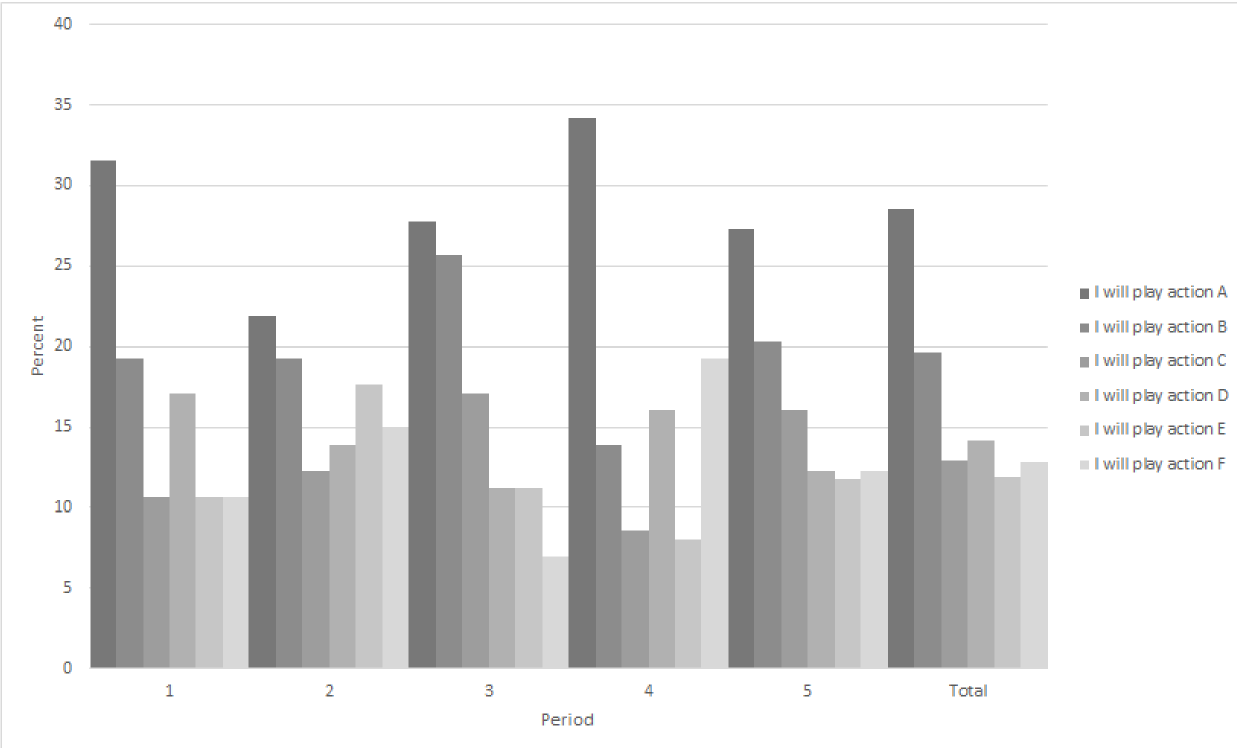


Figure 4: Sender Actions with Communication

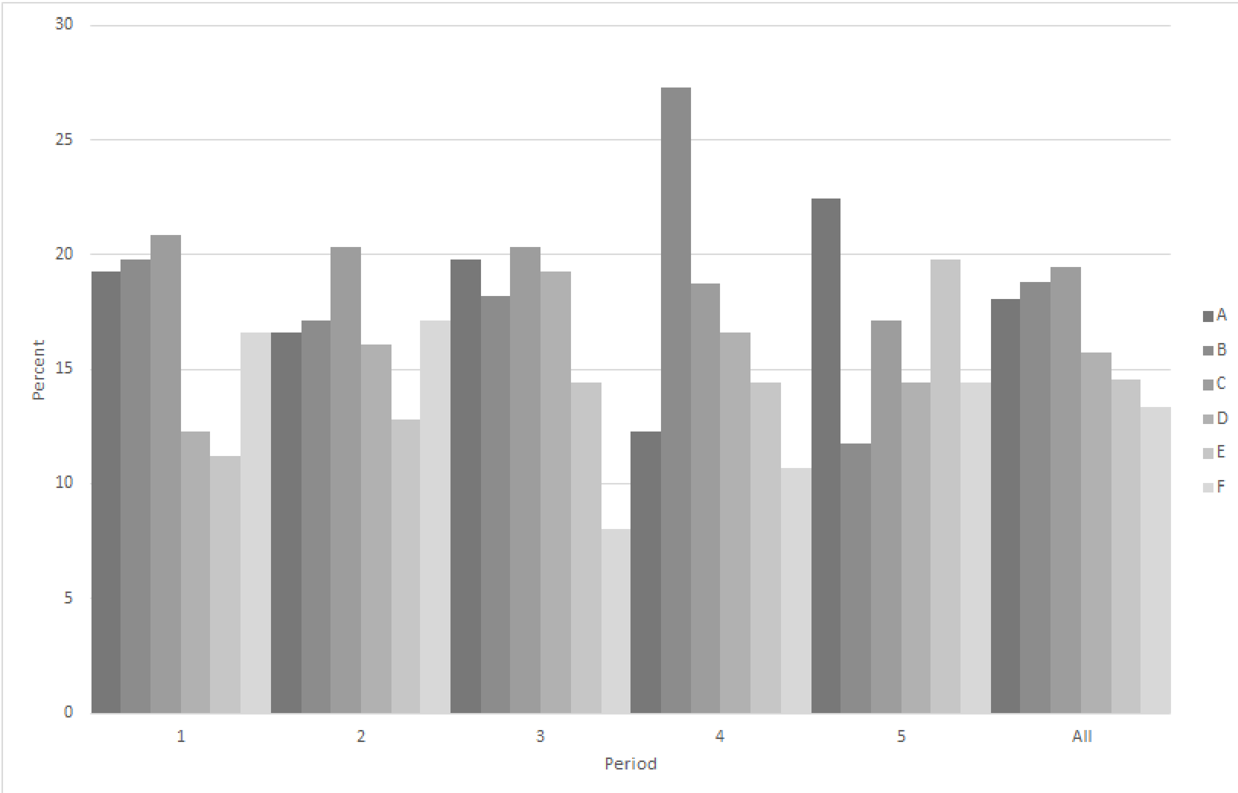


Figure 5: Receiver Actions with Communication

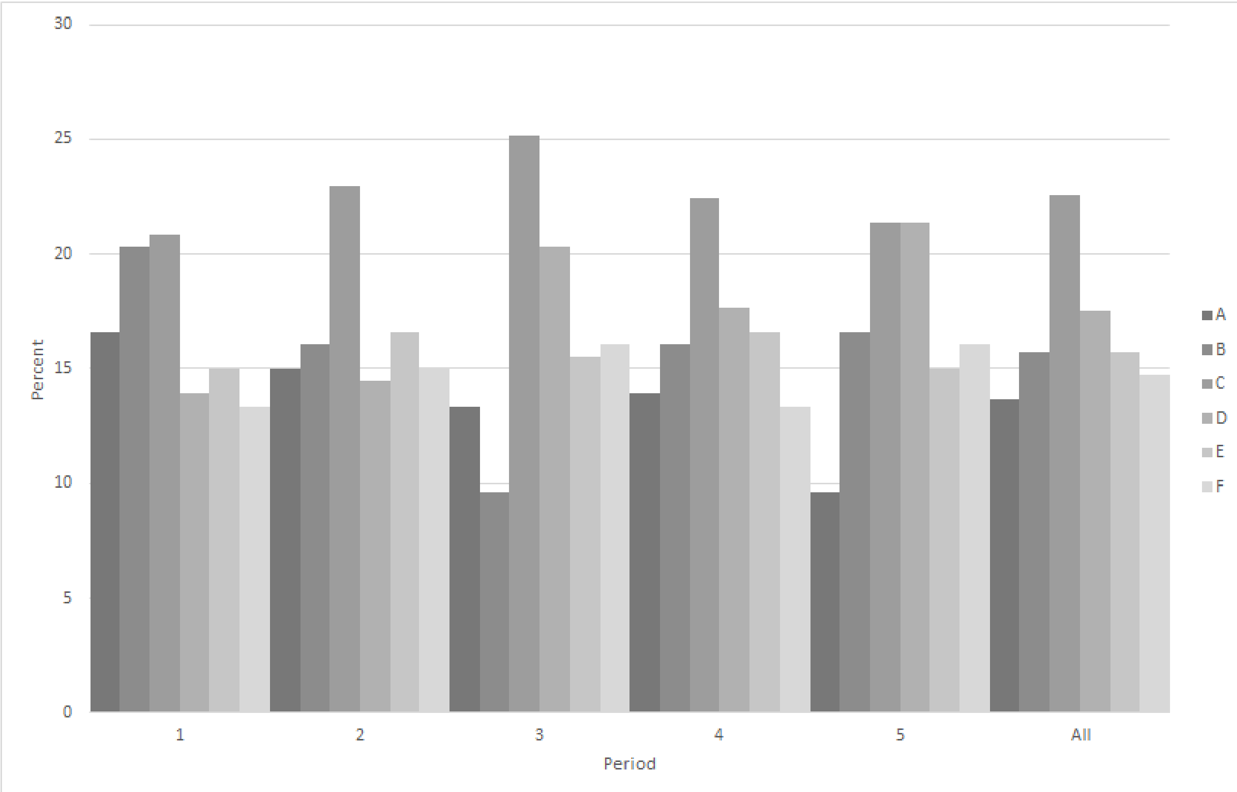


Figure 6: Sender Levels with Communication

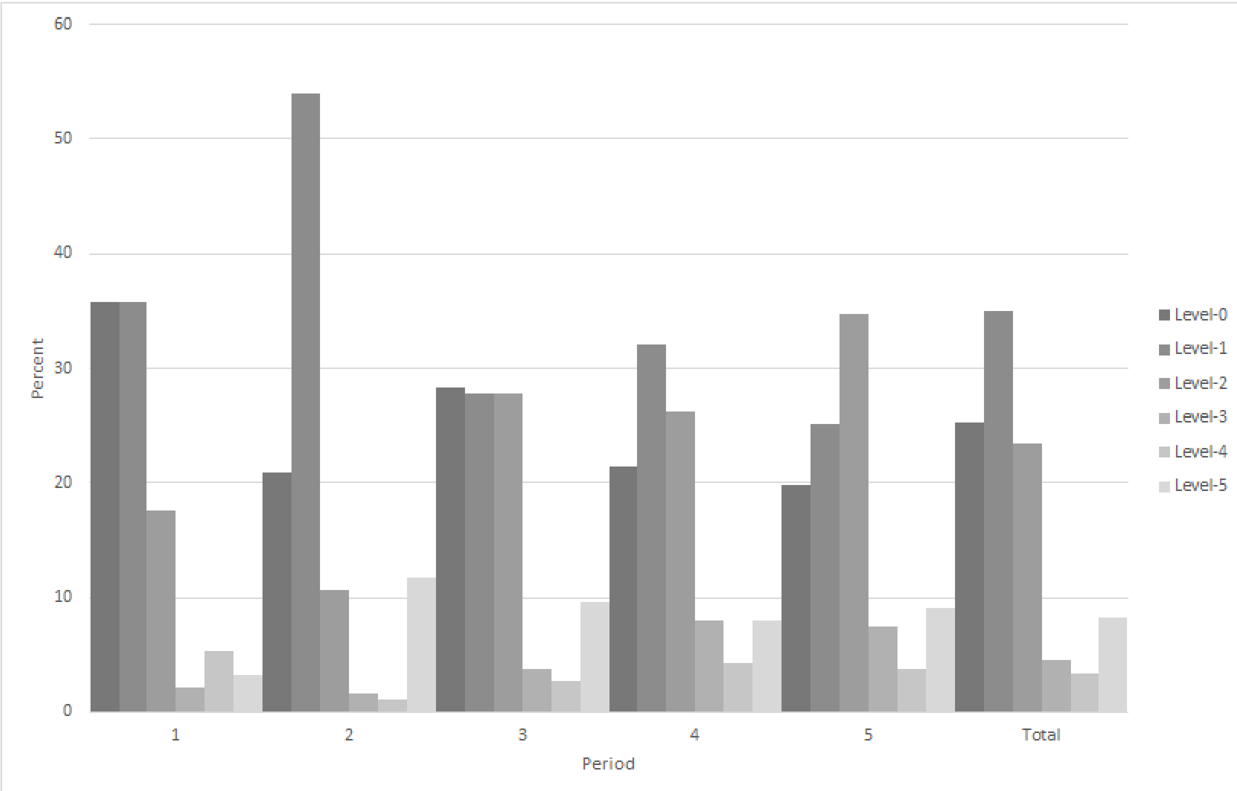
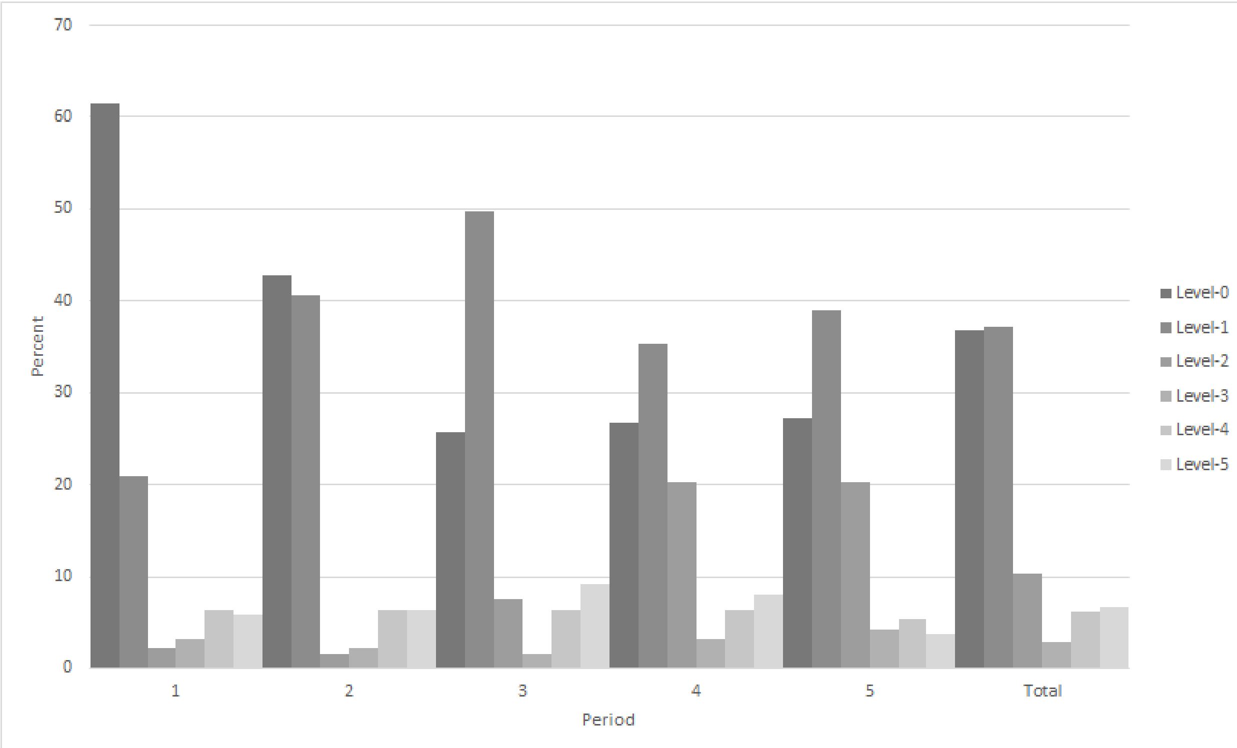


Figure 7: Receiver Levels with Communication



Appendix B: Instructions

Instructions A

Welcome! In this experiment your earnings will be based on the decisions of yourself and other players as well as chance. Please refrain from talking to others during the experiment. In addition, please silence and refrain from using any electronic devices while the experiment is ongoing. If you have any questions or concerns, please raise your hand.

In this experiment, you will be either a Sender or a Receiver. You have an equal chance of being a Sender or a Receiver. Senders will be randomly paired with Receivers in each round of the experiment. You will not see the identity of the person you are paired with, but you will see each player's decisions at the end of each period.

In each round, you and your partner will play the game depicted on the last page of the instructions. Payoffs are in Experimental Currency Units (ECUs). Sender actions and payoffs are in blue, while Receiver actions and payoffs are in red. Both players will pick their actions simultaneously. Note that both the Sender and the Receiver have 6 possible actions, labeled by the following six symbols: $\#, \%, \hat{}, +, ^*$, and (. The payoffs of each player depend on the actions of both players. The Sender's payoff is the first number in each box, while the Receiver's payoff is the second number in each box. Note that the only possible payoff a player can earn is 1 or 0. Also note that the Sender receives a payoff of 1 when the actions of both players match. The Receiver receives a payoff of 1 when the Sender plays # and the Receiver plays %, when the Sender plays % and the Receiver plays^, when the Sender plays^and the Receiver plays +, when the Sender plays + and the Receiver plays *, when the Sender plays * and the Receiver plays (, and when the Sender plays (and the Receiver plays #.

Prior to playing the game, the Sender will send one of the following messages to the Receiver: I will take action #, I will take action %, I will take action^, I will take action +, I will take action *, or I will take action (. The message will be observed by the Receiver before both players simultaneously choose their actions.

In summary, each round will proceed as follows:

First: The Sender sends a message to the Receiver. The message is one of the following: I will take action #, I will take action %, I will take action^, I will take action +, I will take action *, or I will take action (.

Second: Both the Sender and Receiver will simultaneously take an action. An action can be either $\#, \%, \hat{}, +, ^*$, and (. Payoffs for each combination of actions are listed in the figure on the next page.

You will play a total of 5 periods of this experiment. In each period, Senders will be randomly matched and play a round with every Receiver, and Receivers will be randomly matched and play a round with every Sender. In each period, each subject will decide on all actions simultaneously. This means that first, messages will be selected by Senders for every Receiver in the room. Following this, Receivers will observe the messages sent to them and both Senders and Receivers will make all decisions. The identity of each partner will be anonymous and randomized.

At the end of the experiment, each subject will be paid for one **round**, which will be randomly determined by each subject at the end of the experiment. Each subject will be paid $8 for each ECU earned, as well as a $6 appearance fee. Additional questions will be asked at the end of the experiment, which will earn $1 for each correct answer.

		Receiver					
		#	%	^	+	*	(
Sender	#	1, 0	0, 1	0, 0	0, 0	0, 0	0, 0
	%	0, 0	1, 0	0, 1	0, 0	0, 0	0, 0
	^	0, 0	0, 0	1, 0	0, 1	0, 0	0, 0
	+	0, 0	0, 0	0, 0	1, 0	0, 1	0, 0
	*	0, 0	0, 0	0, 0	0, 0	1, 0	0, 1
	(	0, 1	0, 0	0, 0	0, 0	0, 0	1, 0