

Appendix for Online Access [NOT INTENDED FOR PUBLICATION]

Appendix: Methods

Since this paper incorporates economics experiments in the laboratory, eyetracking devices, and studies the issue of deception, we expect to have readers who come from various backgrounds, such as economic theory, experimental economics, psychophysiology, and lie-detection. Therefore, we use this methodology appendix to address issues that might already be very familiar to some readers, but not to the rest. In particular, section 1 introduces video-based eyetracking to economists who are interested in learning about methods to study information acquisition, and section 2 demonstrates the relevance of eyetracking in economic experiments. Section 3 provides an argument for adding yet another paradigm (sender-receiver games) to study lie-detection, instead of adopting previous tasks such as CQT, GKT, etc. Section 4 provides the technical details of the equipment and software programs used in this study for those who are interested in replicating our results or applying this technique in future research.

A.I What is Eyetracking?

There are several ways to track a person's eyes. One of the most reliable and non-invasive way is video-based. Video-based eyetracking works by placing cameras in front of subject's eyes to capture eye images and corneal reflection of infrared sensors, and record changes up to 50-250Hz. Using eye movement images when subjects are told to fixate on certain positions on the screen, a procedure called "calibration," the experimenter can trace eye fixations and saccades on the screen and infer subject information acquisition patterns. In addition to information lookups, the eyetracker also records pupil dilation, which is correlated with arousal, pain, and cognitive

difficulty. Therefore, eyetracking provides additional data about one's decision making process, uncovering previously unobservable parameters.¹

A.II What Does Eyetracking Tell Us About the “Real World”?

Since economists are used to judging theories only by whether they predict choices accurately, it is useful to ask what direct measurement of eye fixations and pupil dilation can add. One possible inferential strategy from eyetracking is to separate competing theories that explain the same behavior. Previous studies compared offers and lookups in three-period alternating-offer bargaining (Camerer et al., 1993; Johnson et al., 2002), and in initial responses to normal-form games and two-person guessing games (Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006). In those experiments, the same choices could be caused by different decision rules, such as L1 (optimize against perceived random play) and D1 (optimize against perceived random play excluding dominated strategies) in Costa-Gomes et al. (2001), but are separated by different lookup generated by these rules.² These studies illustrate the potential for using cognitive data, besides choices, for distinguishing between competing theories or inspiring new theory.³

¹ One potential concern of adopting eyetracking is scrutiny. For example, in our experiments senders could have been more truthful simply because they were watched. Indeed, we do find many L0 and L1 types (seven out of twelve) in the display bias-partner design. But subjects could be more truthful due to the repeated game effect. Hence, such concerns should be dealt with empirically by comparing eyetracked and open box subjects. In our experiment, the hidden bias-stranger adopts random matching and contains both eye-tracked and open boxed subjects. Overall type classification results are similar to Cai and Wang (2006). Although the sub-samples of eyetracked and open box subjects do show some interesting differences, the average level of strategic thinking is comparable: None of the eyetracked subjects were EQ (L3), but there were many SOPH; none of the open box subjects were L1, but the only L0 subject was an open box. This results in lower correlation between state and message for the open box subjects, but there is still little difference in payoffs. Hence, we conclude that there is no striking difference between the two, though the sample size is small.

² For example, in the three-stage bargaining game of Camerer et al. (1993) and Johnson et al. (2002), opening offers typically fell between an equal split of the first-period surplus and the subgame perfect equilibrium prediction (assuming self-interest). These offers could be caused by limited strategic thinking (i.e., players do not always look ahead to the second and third round payoffs of the game), or by computing an equilibrium by looking ahead, adjusting for fairness concerns of other players. The failure to look at payoffs in future periods showed that the deviation of offers from equilibrium was (at least partly) due to limited strategic thinking, rather than entirely due to equilibrium adjustment for fairness (unless “fairness” means not at all responding to advantages conferred by the strategic structure). Furthermore, comparing across rounds, when players do look ahead at future round payoffs their resulting offer are

Lookup patterns and pupil dilation could be useful in the sender-receiver games, because it could potentially be used to distinguish between competing theories for overcommunication. Although our experiments are not designed to separate these theories, overcommunication of the true state is consistent with two rough accounts: guilt and cognitive difficulty. Senders may feel guilty about deceiving the receivers and potentially costing the receivers money. This is the direct cost of lying. According to this theory, senders will look at the receiver payoffs (since seeing those payoffs is the basis of guilt) and their pupils will dilate when they misrepresent the state (i.e., choose M different from S) due to emotional arousal from guilt. In this story, the guilt springs from the senders' realization that their actions cost the receivers money, which depends on seeing the receiver payoffs.

A different story is that senders find it cognitively difficult to figure out how much to misrepresent the state. For example, senders might believe that some other senders always tell the truth, and receivers might therefore believe messages are truthful. Then strategic senders have to think hard about how much to misrepresent the state to take advantage of the receivers' naïveté (as in Crawford, 2003, Kartik, Ottaviani and Squintani, 2007, Chen, 2007, and Kartik, 2008). In this story, senders do not have to pay much attention to receiver payoffs but their pupils will dilate because of the cognitive difficulty of figuring out precisely how much to exaggerate.

Ultimately, the goal is to open up the black box of human brain, and model the decision process of human behavior, which is similar to what has been done to the firm. Instead of dwelling

closer to the self-interested equilibrium prediction (see Johnson and Camerer, 2004). Thus, the lookup data can actually be used to predict choices, to some degree.

³ Another example comes from the accounting literature: James E. Hunton and McEwen (1997) asked analysts under hypothetical incentive schemes to make earnings forecast based on real firm data, and investigated factors that affect the accuracy of these forecasts. Using an eye-movement computer technology (Integrated Retinal Imaging System, IRIS), they find that analysts who employ a “directive information search strategy” make more accurate forecasts, both in the lab and in the field, even after controlling for years of experience. This indicates that eyetracking may provide an alternative measure of experience or expertise that is not simply captured by seniority. Had they not observed the eye movements, they could not have measured the difference in information search which is linked to accuracy.

on the neoclassical theory of the firm, which is merely a production function, modern economics has opened up the black box of the firm, and explicitly modeled its internal structure, such as the command hierarchy, principle-agent issues, and team production. Though there is still much to be done before we come close to what has been achieved in industrial organization, eyetracking provides a window to the soul and gives us a hint of the decision-making process inside the brain. Just as we may infer a factory's technology level by observing its inputs and wastes, we may also infer a person's reasoning process by observing the information he or she acquires (inputs) and how hard does he think (indexed by pupillary response).

A.III What Does Economics Have to Offer Regarding Lie-detection?

This study introduces an economic framework that is missing in most previous psychophysical studies on deception and lie detection. An advantage of the strategic information transmission game for studying deception is that game theory makes equilibrium predictions about how much informed agents will exaggerate what they know, when they know that other agents are fully-informed about the game's structure and the incentives to exaggerate. Even when equilibrium predictions fail, there are various behavioral models, such as level-k reasoning and quantal response equilibrium, which provide precise predictions that are testable in the lab. And while in most other deception studies,⁴ subjects are instructed to lie or give weak or poorly controlled incentives,⁵ subjects in experiments like ours choose voluntarily whether to deceive others or not (see also John

⁴ For a survey of studies on (skin-conductance) polygraph, see Theodore R. Bashore and Paul E. Rapp (1993). For lie-detection studies in psychology, see the reviews of Robert E. Kraut (1980) and Aldert Vrij (2000). For a comprehensive discussion of different cues used to detect lies, see Bella M. DePaulo et al. (2003). For individual differences in lie-detection (Secret Service, CIA and sheriffs do better), see Paul Ekman and Maureen O'Sullivan (1991) and Ekman et al. (1999). More recently studies in neuroscience using functional magnetic resonance imaging (fMRI) include Sean A. Spence et al. (2001), D. D. Langleben et al. (2002) and F. Andrew Kozel et al. (2004).

⁵ One exception is Samantha Mann et al. (2004) which used footage of real world suspect interrogation to test lie-detecting abilities of ordinary police. However, a lot of experimental control is lost in this setting. One interesting findings in this study is that counter to conventional wisdom, the more subjects relied on stereotypical cues such as gaze aversion to detect lies, the *less* accurate they were.

Dickhaut et al., 1995, Andreas Blume et al., 1998, 2001 and Cai and Wang, 2006).⁶ Senders and receivers also have clear measurable economic incentives to deceive and to detect deception.⁷

A.IV Technological Details

Eyetracking data and button responses are recorded using the mobile Eyelink II head-mounted eyetracking system (SR Research, Osgoode, Ontario). Eyetracking data are recorded at 250 Hz. The mobile Eyelink II is a pair of tiny cameras mounted on a lightweight rack facing toward the subjects' eyes, and is supported by comfortable head straps. Subjects can move their heads and a period of calibration adjusts for head movement to infer accurately where the subject is looking. Nine-point calibrations and validations are performed prior to the start of each experiment in a participant's session. Accuracy in the validations typically is better than 0.5° of visual angle. Experiments are run under Windows XP (Microsoft, Inc.) in Matlab (Mathworks, Inc., Natick, MA) using the Psychophysics Toolbox (David H. Brainard, 1997; Denis G. Pelli, 1997) and the Eyelink Toolbox (Frans W. Cornelissen et al., 2002).

Eyetracking data are analyzed for fixations using the Eyelink Data Viewer (SR Research, Hamilton, Ontario). In discriminating fixations, we set saccade velocity, acceleration, and motion

⁶ In fact, when the senders were asked after the experiment whether they considered sending a number different from the true state deception, 8 of the subjects said yes, while another 3 said no, but gave excuses such as "it's part of the game" or "the other player knows my preference difference." Only 1 subject said no without any explanation. These debriefing results also suggest that guilt has played little role in the experiment.

⁷ Most lie-detection studies have three drawbacks: (1) They do not use naturally-occurring lies (because it is then difficult to know whether people are actually lying or not). Instead, most studies create artificial lies by giving subjects true and false statements (or creating a "crime scenario") and instructing them to either lie or tell the truth, sometimes to fool a lie-detecting algorithm or subject. However, instructed deception can be different than naturally-occurring voluntary deception, and the ability to detect instructed deception might be different than detecting voluntary deception. (2) The incentives to deceive in these studies are typically weak or poorly controlled (e.g., in Spence et al. (2001) all subjects were told that they successfully fooled the investigators who tried to detect them; in Mark G. Frank and Ekman (1997), subjects were threatened with "sitting on a cold, metal chair inside a cramped, darkened room labeled ominously XXX, where they would have to endure anywhere from 10 to 40 randomly sequenced, 110-decibel startling blasts of white noise over the course of 1 hr" but never actually enforcing it.). (3) Subjects are typically not economically motivated to detect deception. Experiments using the strategic-transmission paradigm from game theory address all these drawbacks.

thresholds to 30%/sec, 9500%/sec², and 0.15°, respectively. Regions of interest (ROIs), or the boxes subject look up, are drawn on each task image using the drawing functions within the Data Viewer. Measures of gaze include Fixation Number (i.e., the total number of fixations within an ROI) and Fractional Dwell Time (i.e., the time during a given round spent on fixating a given ROI divided by the total time between image onset and response). Only those fixations beginning between 50ms following the onset of a task image and offset of the task image are considered for analysis.

All task images are presented on a CRT monitor (15.9 in x 11.9 in) operating at 85 or 100 Hz vertical refresh rate with a resolution of 1600 pixels x 1200 pixels, and at an eye-to-screen distance of approximately 24 inches, thus subtending ~36 degrees of visual angle.

References

- Bashore, Theodore R. and Rapp, Paul E.** "Are There Alternatives to Traditional Polygraph Procedures." *Psychological Bulletin*, 1993, 113(1), pp. 3-22.
- Brainard, David H.** "The Psychophysics Toolbox." *Spatial Vision*, 1997, 10, pp. 433-36.
- Camerer, Colin F.; Johnson, Eric J.; Rymon, Talia and Sen, Sankar.** "Cognition and Framing in Sequential Bargaining for Gains and Losses," K. G. Binmore, A. P. Kirman and P. Tani, *Frontiers of Game Theory*. Cambridge: MIT Press, 1993, 27-47.
- Chen, Ying.** "Perturbed Communication Games with Honest Senders and Naive Receivers," Unpublished paper, 2007.
- Cornelissen, Frans W.; Peters, Enno M. and Palmer, John.** "The Eyelink Toolbox: Eye Tracking with Matlab and the Psychophysics Toolbox." *Behavior Research Methods, Instruments & Computers*, 2002, 34, pp. 613-17.
- Costa-Gomes, Miguel; Crawford, Vincent P. and Broseta, Bruno.** "Cognition and Behavior in

- Normal-Form Games: An Experimental Study." *Econometrica*, 2001, 69(5), pp. 1193-235.
- Crawford, Vincent P.** "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review*, 2003, 93(1), pp. 133-49.
- DePaulo, Bella M.; Lindsay, James J.; Malone, Brian E.; Muhlenbruck, Laura; Charlton, Kelly and Cooper, Harris.** "Cues to Deception." *Psychological Bulletin*, 2003, 129(1), pp. 74-118.
- Ekman, Paul and O'Sullivan, Maureen.** "Who Can Catch a Liar?" *American Psychologist*, 1991, 46, pp. 913-20.
- Ekman, Paul; O'Sullivan, Maureen and Frank, Mark G.** "A Few Can Catch a Liar." *Psychological Science*, 1999, 10, pp. 263-66.
- Frank, Mark G. and Ekman, Paul.** "The Ability to Detect Deceit Generalizes Across Different Types of High-Stake Lies." *Journal of Personality and Social Psychology*, 1997, 72(6), pp. 1429-39.
- Hunton, James E. and McEwen, Ruth A.** "An Assessment of the Relation between Analysts' Earnings Forecast Accuracy, Motivational Incentives and Cognitive Information Search Strategy." *Accounting Review*, 1997, 72(4), pp. 497-515.
- Johnson, Eric J. and Camerer, Colin F.** "Thinking Backward and Forward in Games," I. Brocas and J. Castillo, *The Psychology of Economic Decisions, Vol.2: Reasons and Choices*. Oxford University Press, 2004.
- Kartik, Navin; Ottaviani, Marco and Squintani, Francesco.** "Credulity, Lies, and Costly Talk." *Journal of Economic Theory*, 2007, 136 pp. (1), pp. 749-58.
- Kartik, Navin.** "Strategic Communication with Lying Costs." *Review of Economic Studies*, 2008, forthcoming.

- Kozel, F. Andrew; Revell, Letty J.; Lorberbaum, Jeffrey P.; Shastri, Ananda; Elhai, Jon D.; Horner, Michael David; Smith, Adam; Nahas, Ziad; Bohning, Daryl E. and George, Mark S.** "A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men." *Journal of Neuropsychiatry and Clinical Neurosciences*, 2004, 16, pp. 295-305.
- Kraut, Robert E.** "Humans as Lie Detectors: Some Second Thoughts." *Journal of Communication*, 1980, 30, pp. 209-16.
- Langleben, D. D.; Schoroeder, L.; Maldjian, J. A.; Gur, R. C.; McDonald, S.; Ragland, J. D.; O'Brien, C. P. and Childress, A. R.** "Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study." *NeuroImage*, 2002, 15(3), pp. 727-32.
- Mann, Samantha; Vrij, Aldert and Bull, Ray.** Detecting True Lies: Police Officers' Ability to Detect Suspects' Lies, *Journal of applied psychology*, 2004, 89(1), pp. 137-49.
- Pelli, Denis G.** "The Videotoolbox Software for Visual Psychophysics: Transforming Numbers into Movies." *Spatial Vision*, 1997, 10, pp. 437-42.
- Spence, Sean A.; Farrow, Tom F. D.; Herford, Amy E.; Wilkinson, Iain D.; Zheng, Ying and Woodruff, Peter W. R.** "Behavioural and Functional Anatomical Correlates of Deception in Humans." *NeuroReport*, 2001, 12(13), pp. 2849-53.
- Vrij, Aldert.** *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice*. Chichester: Wiley and Sons, 2000.

Appendix: Experiment Instructions

The experiment you are participating in consists of 1 session, having 45 rounds. At the end of the last session, you will be asked to fill out a questionnaire and paid the total amount you have accumulated during the course of the sessions in addition to a \$5 show-up fee. Everybody will be paid in private after showing the record sheet. You are under no obligation to tell others how much you earned.

During the experiment all the earnings are denominated in FRANCS. Your dollar earnings are determined by the FRANC/\$ exchange rate: 200 FRANCS = \$1.

In each round, the computer program generates a secret number that is randomly drawn from the set $\{1,2,3,4,5\}$. The computer will display this secret number on member A's screen. After receiving the number, member A will send the message "The number I received is XX," to member B by staring at box XX. Hearing the message from member A, member B will then choose an action. In particular, member B can choose action 1, 2, 3, 4, or 5, using the game pad. Earnings of both members depend on the secret number and member B's action.

Member B's earnings is higher when member B's action is closer to the secret number, while member A's earnings is higher when member B's action is closer to the secret number **plus the preference difference**. The preference difference is either 0, 1 or 2, with equal chance, and will also be displayed and announced at the beginning of each round.

For example, if the preference difference is 2 and the secret number is 3, member B's earnings are higher if his or her action is closer to 3. However, member A's earnings is higher when member B's action is closer to $3 + 2 = 5$. The earning tables are provided to you for convenience.

To summarize, in each round, the computer will display the preference difference and the secret number on member A's screen. Then, member A stares at a box (on the right) containing the desired message. Member B will hear the preference difference and the message "The number I received is XX," and then choose an action. The secret number is revealed after this choice, and earnings are determined accordingly.

Practice Session: 3 Rounds

Session 1: 45 Rounds

Member B: Please make sure you record the earnings in your record sheet. **Your payments will be rounded up.** Thank you for your participation.

Appendix: Supplemental Figures and Tables

Figure S1: Sender Screen for $b=1$ and $S=4$ without payoff perturbation

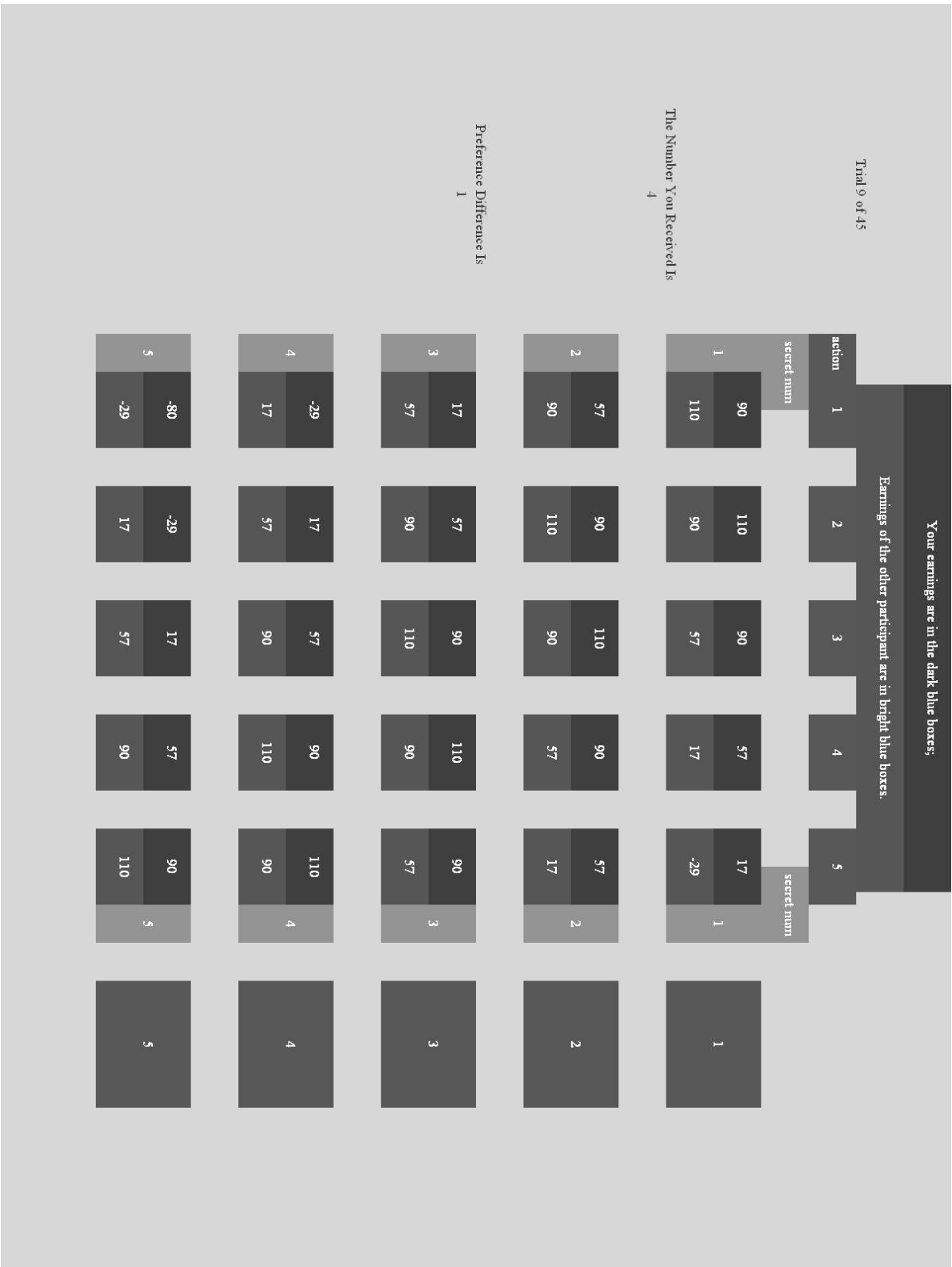


Figure S2: Raw Data Pie Charts (b=0), (Display Bias-Partner)

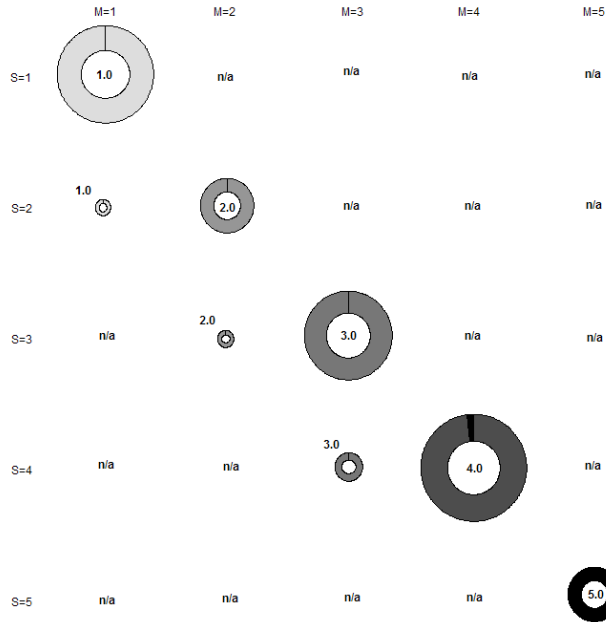


Figure S3: Raw Data Pie Chart (b=1), (Display Bias-Partner)

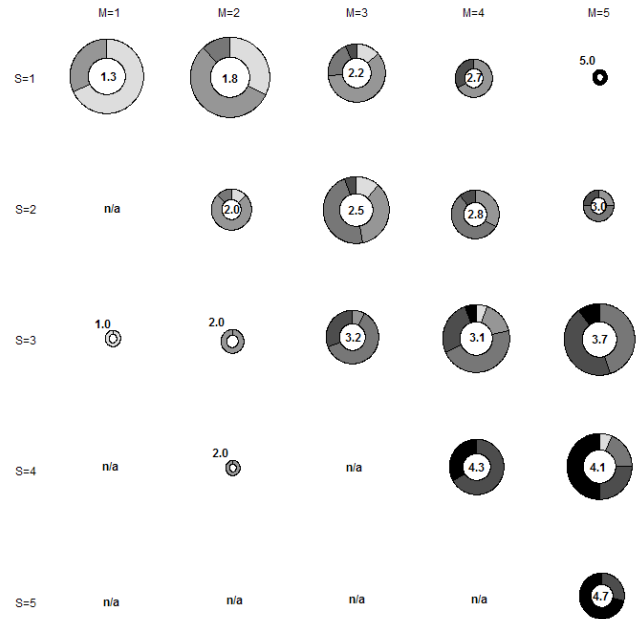


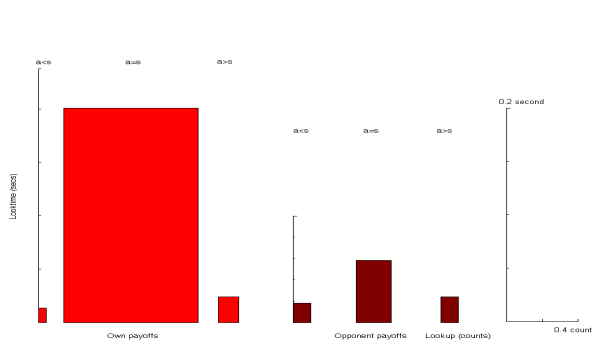
Figure S4: Raw Data Pie Chart (b=2), (Display Bias-Partner)



The true states are in rows, and senders' messages are in columns. Each cell contains the average action taken by the receivers and a pie chart break down of the actions. Actions are presented in a gray scale, ranging from white (action 1) to black (action 5). The size of the pie chart is proportional to the number of occurrences for the corresponding state and message.

Figure S5: Lookup Icon Graph for $b=0$, Type = all

Part (a): Display Bias-Partner



Part (b): Hidden Bias-Stranger

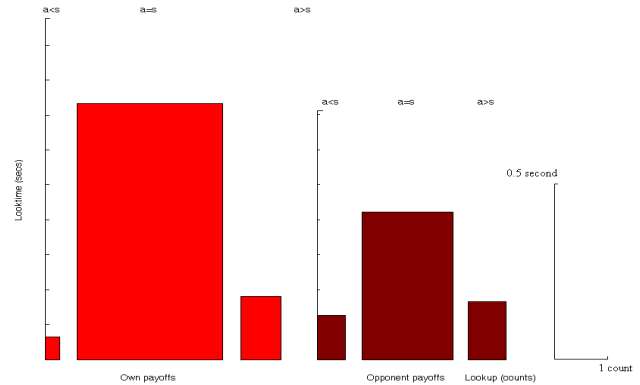
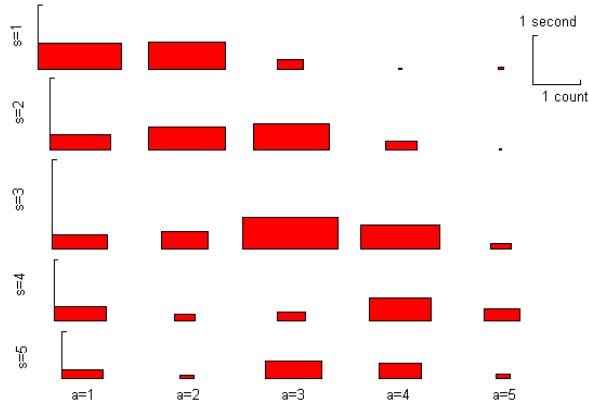


Figure S6: Lookup Icon Graph for $b=1$, Display Bias-Partner, Type = all

Part (a): Sender Payoffs



Part (b): Receiver Payoffs

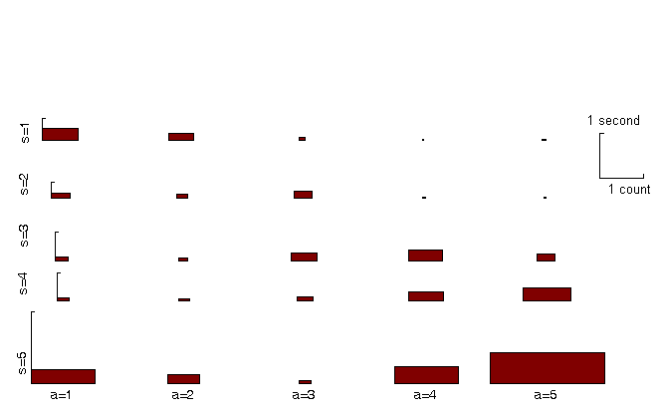
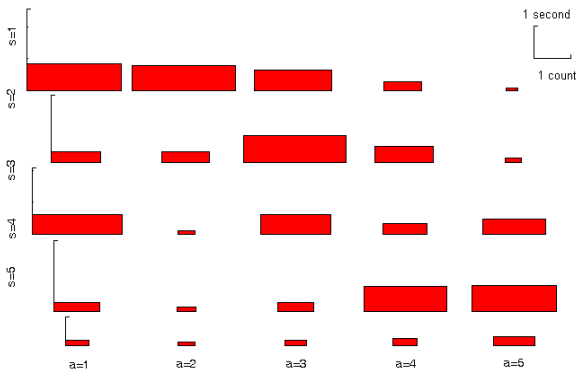
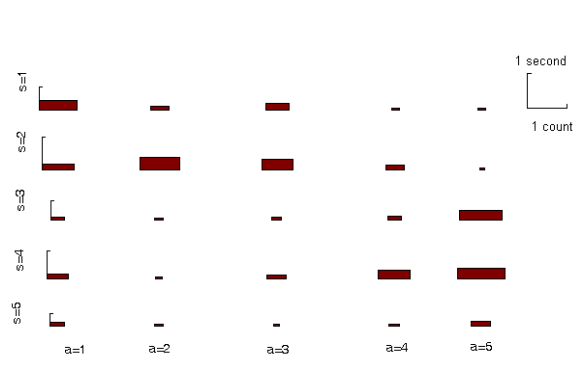


Figure S7: Lookup Icon Graph for $b=2$, Display Bias-Partner, Type = all

Part (a): Sender Payoffs



Part (b): Receiver Payoffs



Each row reports the lookup counts and time for the “true state row” corresponding to the given true state. The width of each box is scaled by the number of lookups and the height by the length of lookups (scaled by the little “ruler” in the upper right corner). The vertical bar on the first column icon represents the total lookup time summed across each row.

Figure S8: Lookup Icon Graph for b=1, Hidden Bias-Stranger, Type = all

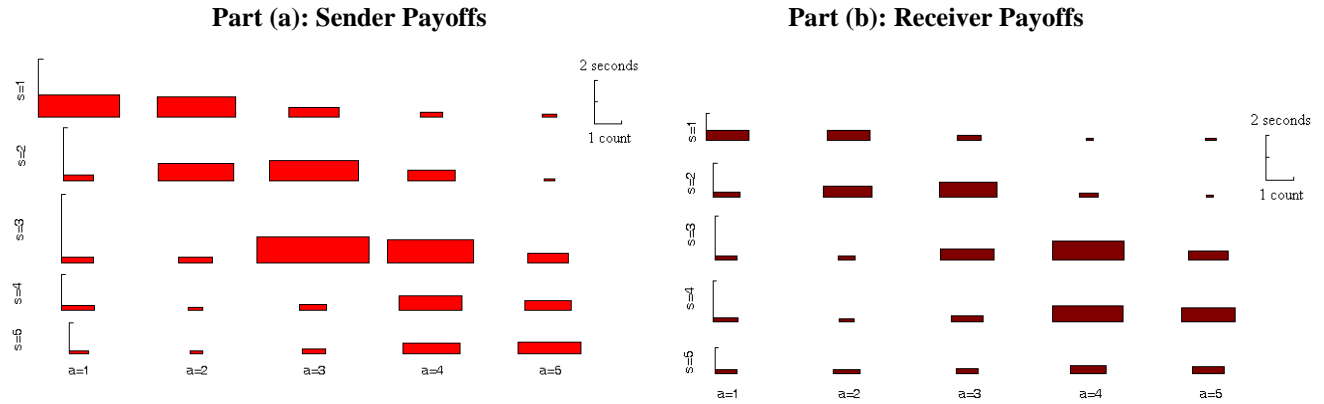
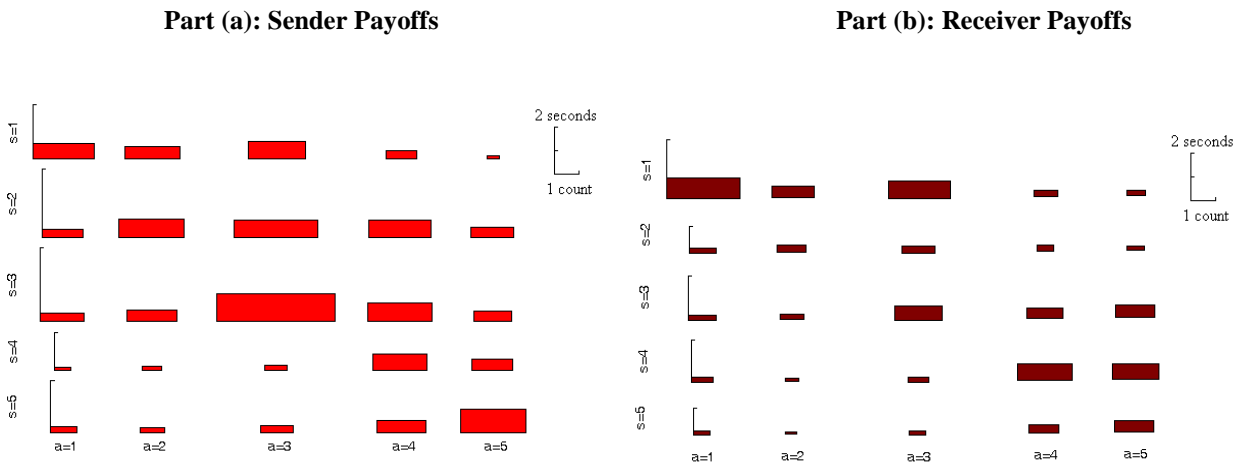


Figure S9: Lookup Icon Graph for b=2, Hidden Bias-Stranger, Type = all



Each row reports the lookup counts and time for the “true state row” corresponding to the given true state. The width of each box is scaled by the number of lookups and the height by the length of lookups (scaled by the little “ruler” in the upper right corner). The vertical bar on the first column icon represents the total lookup time summed across each row.

Table S1A: Learning – Actual Information Transmission

Display Bias-Partner					
BIAS	Rounds	Corr(S, M)	Corr(M, A)	Corr(S, A)	Predicted Corr(S, A)
0	1-15	0.880	0.833	0.732	1.000
	16-30	0.976	0.949	0.925	
	31-45	0.937	0.942	0.919	
1	1-15	0.620	0.730	0.477	0.645
	16-30	0.685	0.724	0.577	
	31-45	0.598	0.713	0.415	
2	1-15	0.384	0.584	0.372	0.000
	16-30	0.327	0.526	0.306	
	31-45	0.279	0.643	0.291	
Hidden Bias-Stranger					
BIAS	Rounds	Corr(S, M)	Corr(M, A)	Corr(S, A)	Predicted Corr(S, A)
0	1-15	0.887	0.816	0.716	1.000
	16-30	0.941	0.951	0.885	
	31-45	0.888	0.944	0.866	
1	1-15	0.602	0.730	0.436	0.645
	16-30	0.660	0.727	0.561	
	31-45	0.555	0.714	0.393	
2	1-15	0.380	0.592	0.372	0.000
	16-30	0.347	0.540	0.313	
	31-45	0.232	0.636	0.288	

Table S1B: Learning Sender and Receiver's Payoffs

Display Bias-Partner				
BIAS	Rounds	u_S (std)	u_R (std)	Predicted u_R (std)
0	1-15	96.36 (23.47)	96.48 (24.37)	
	16-30	104.63 (11.65)	104.78 (12.01)	110.00 (0.00)
	31-45	103.50 (12.46)	103.19 (12.18)	
1	1-15	79.38 (31.83)	87.04 (26.78)	
	16-30	69.19 (40.15)	87.98 (28.94)	91.40 (19.39)
	31-45	71.83 (39.05)	85.52 (27.09)	
2	1-15	46.06 (50.91)	80.63 (25.93)	
	16-30	46.74 (51.11)	81.20 (27.63)	80.80 (20.76)
	31-45	35.87 (55.73)	79.70 (29.65)	
Hidden Bias-Stranger				
BIAS	Rounds	u_S (std)	u_R (std)	Predicted u_R (std)
0	1-15	95.38 (23.56)	95.72 (24.15)	
	16-30	102.40 (15.18)	102.52 (15.53)	110.00 (0.00)
	31-45	102.00 (16.89)	101.69 (17.30)	
1	1-15	78.76 (35.63)	85.88 (28.92)	
	16-30	69.18 (39.40)	87.45 (28.61)	91.40 (19.39)
	31-45	71.40 (38.82)	84.73 (26.87)	
2	1-15	46.76 (49.84)	81.06 (26.36)	
	16-30	46.75 (50.19)	81.81 (27.15)	80.80 (20.76)
	31-45	36.22 (55.94)	79.29 (29.10)	

Table S2: Information Transmission: Correlations between S, M and A, Display Bias-Partner

Bias	r(S, M)	r(M, A)	r(S, A)	Predicted r(S, A)
0	.99	1.00	.99	1.00
1	.73	.74	.72	.65
2	.63	.57	.50	.00

Note: In the display bias-partner design, all senders' eye movements were recorded ("eyetracked").

Table S3: Sender and Receiver's Payoffs, Display Bias-Partner

Bias	u_S (std)	u_R (std)	Pred. u_R (std)
0	109.14 (4.07) ^a	109.14 (4.07) ^a	110.00 (0.00)
1	93.35 (20.75)	94.01 (19.86)	91.40 (19.39)
2	41.52 (49.98)	85.52 (25.60)	80.80 (20.76)

Note: ^a Payoffs are exactly the same for senders and receivers due to the symmetry of the payoffs when $b=0$.

Table S4: Level-k Classification Results, Display Bias-Partner

Session	ID	log L	k	Exact	lambda
1	1	-36.33	L0	0.71	0.06
2	2	-51.47	L0	0.64	0.00
3	3	-33.01	L0	0.78	0.03
4	4	-19.81	L1	0.82	0.49
5	5	-38.93	SOPH	0.76	0.04
6	6	-45.05	EQ	0.69	0.05
7	7	-34.89	L0	0.80	0.00
8	8	-27.36	L2	0.84	0.04
9	9	-31.80	L1	0.80	0.04
10	10	-24.30	L1	0.84	0.48
11	11	-22.35	L2	0.87	0.45
12	12	-31.07	L2	0.73	1.00

Table S5: Average Sender Lookup Times (in sec.) across Game Parameters, Display Bias-Partner

Bias b	Response Time		State	Bias	Sender Payoffs	Receiver Payoffs	Sender-to- Receiver Ratio
	Periods	Periods					
	1-15	31-45					
0	5.42	2.39	0.65	0.41	0.73	0.27	2.70
1	7.92	5.44	1.47	0.99	2.29	1.05	2.18
2	9.73	8.12	1.72	1.52	3.03	1.50	2.02
all	8.07	5.25	1.34	1.02	2.14	1.00	2.14

Table S6: Average Lookup Time per Row Depending on the State, Display Bias-Partner

Bias b	True State Rows	Other State Rows	True-to-Other Ratio
0	0.54	0.11	4.91
1	2.06	0.32	6.44
2	2.24	0.57	4.28
overall	1.71	0.36	4.75

Table S7A: Average response time change for different biases, Display Bias-Partner

Bias	N	Average for first 15 rounds	N	Average for middle 15 rounds	N	Average for last 15 rounds
0	38	5.42	47	2.91	55	2.39
1	73	7.92	60	5.44	59	5.44
2	67	9.73	68	8.96	51	8.12
overall	178	8.07	175	6.13	165	5.25

* The numbers of observations are slightly different because we exclude 10 rounds where subjects had to use the keyboard to make their decision. Also, subject #4 had severe pain and the experimenter was forced to stop the experiment at the end of round 33.

Note: Since the bias was randomly determined each round, and subject #4 stopped at round 33 (due to excess pain wearing the eyetracker), numbers of observations are not equal. Dropping subject #4 does not change the results.

Table S7B: Average response time change for different biases, Hidden Bias-Stranger

Bias	N	Average for first 15 rounds	N	Average for middle 15 rounds	N	Average for last 15 rounds
0	30	9.78	24	5.54	29	7.24
1	56	11.77	58	10.78	59	8.76
2	61	16.84	65	10.23	49	8.99
overall	147	13.47	147	9.68	137	8.52

* The numbers of observations are slightly different because we exclude 12 rounds where subjects had to use the keyboard to make their decision. Also, subject #3 had calibration issues and the experimenter was forced to stop eyetracking at the end of round 40.

Note: Since the bias was randomly determined each round, and subject #4 stopped at round 40 wearing the eyetracker), numbers of observations are not equal.

Table S8: Pupil Size Regressions for 400 msec Intervals, Display Bias-Partner

Y	PUPIL _i	-1.2~	-0.8~	-0.4~	0.0~	0.4~
		-0.8sec	-0.4sec	0.0sec	0.4sec	0.8sec
constant	α	99.59 (2.45)	99.78 (2.41)	104.62 (2.19)	111.81 (1.84)	109.95 (2.07)
LIE_SIZE * BIAS _b interactions	β_{10}	1.20 (3.21)	6.41 (6.38)	3.92 (3.06)	-3.91 (2.76)	0.58 (7.36)
	β_{11}	2.79* (1.19)	3.40** (1.17)	3.28** (0.97)	4.55*** (0.86)	4.20*** (0.73)
	β_{12}	3.49*** (0.99)	3.71*** (0.98)	3.04*** (0.84)	2.90** (0.87)	3.28** (0.90)
	N	499	497	499	508	503
	χ^2	224.54	337.22	500.93	785.32	631.21
	R ²	0.271	0.346	0.455	0.539	0.557

Note: Robust standard error in parentheses; t-Test p-values lower than ^10 percent, *5 percent, ** 1 percent, and *** 0.1 percent. (Dummies for biases, states, individual subjects and individual learning trends are included in the regression, but results are omitted.)

Table S9: Predicting True States (Resampling 100 times, s.e. in parentheses), Display Bias-Partner

X	Display Bias-Partner	
MESSAGE * BIAS = 1	β_{11}	0.64* (0.22)
MESSAGE * BIAS = 2	β_{12}	0.91** (0.23)
ROW _{self} * BIAS=1	β_{21}	0.98** (0.21)
ROW _{self} * BIAS=2	β_{22}	1.00** (0.27)
ROW _{other} * BIAS=1	β_{31}	0.25 (0.16)
ROW _{other} * BIAS=2	β_{32}	0.39* (0.17)
total observations N ^a		208
N used in estimation		139.3
N used to predict		68.7
	Actual Data	Hold-out Sample
Percent of wrong prediction (b=1)	56.2	29.2
Percent of errors of size (1,2,3+) (b=1)	(80, 15, 5)	(74, 19, 7)
Average predicted payoff (b=1) ^b	93.4 (22.3)	100.7* (2.4)
Percent of wrong prediction (b=2)		70.9
Percent of errors of size (1,2,3+) (b=2)		(67, 26, 7)
Average predicted payoff (b=2) ^b		86.2 (23.8)
		91.8* (3.4)

Note: * and ** Denotes $p < 0.05$ and $p < 0.001$ (t-test)

^a Observation with less than 0.5 seconds lookup time and without the needed pupil size measures are excluded.

^b Two sample t-test conducted against the actual payoffs of receivers in the experiment who are paired with eyetracked senders.

Table S10: Average Sender Fixation Counts and Lookup Time across Game Parameters

Treatment	Bias b	Res- ponse time (sec.)	State		Bias		Sender Payoffs		Receiver Payoffs	
			Fixation (count)	Lookup (sec.)	Fixation (count)	Lookup (sec.)	Fixation (count)	Lookup (sec.)	Fixation (count)	Lookup (sec.)
Displayed - Partner	0	3.59	2.6	0.65	2.1	0.41	3.0	0.73	1.4	0.27
	Bias 1	6.86	5.0	1.47	3.9	0.99	8.1	2.29	3.9	1.05
	2	9.68	6.2	1.72	5.5	1.52	10.6	3.03	5.4	1.50
	overall	7.00	4.8	1.34	4.0	1.02	7.6	2.14	3.7	1.00
Hidden - Stranger	0	7.65	3.0	0.83	-	-	12.0	2.93	7.5	1.71
	Bias 1	10.95	3.1	0.81	-	-	14.2	3.80	10.7	2.66
	2	12.91	3.4	0.91	-	-	17.5	4.67	12.4	3.26
	overall	11.12	3.2	0.86	-	-	15.1	3.99	10.8	2.72

Table S11: Average Fixation Counts and Lookup Time per Row

Treatment	Bias b	True State Rows		Other Rows	
		Fixation Counts (counts per row)	Lookup Time (sec. per row)	Fixation Counts (counts per row)	Lookup Time (sec. per row)
Displayed - Partner	0	2.2	0.54	0.5	0.11
	Bias 1	6.8	2.06	1.3	0.32
	2	7.8	2.24	2.0	0.57
	overall 1	5.9	1.71	1.3	0.36
Hidden - Stranger	0	11.4	2.76	2.0	0.47
	Bias 1	14.4	3.88	2.6	0.64
	2	15.7	4.29	3.6	0.91
	overall	14.3	3.83	2.9	0.72

Table S12: Individual Types and Log Likelihood under Spike-logit and Logit Specification

Ses-	Sub-	Spike-logit (baseline)					Spike-logit (without bias=0)					Logit					Logit (without bias=0)				
		L0	L1	L2	L3	SOPH	L0	L1	L2	L3	SOPH	L0	L1	L2	L3	SOPH	L0	L1	L2	L3	SOPH
1	1	-60.20	-55.68	-46.36	-53.16	<u>-46.23</u>	-50.47	-43.08	-36.68	-41.48	<u>-35.28</u>	-66.92	-52.59	-50.83	-54.65	<u>-48.72</u>	-54.44	-42.12	-40.07	-43.31	<u>-38.50</u>
1	2	-67.54	<u>-25.99</u>	-55.16	-56.98	-55.82	-55.14	<u>-24.72</u>	-50.15	-51.96	-50.80	-66.85	<u>-36.95</u>	-49.41	-51.79	-48.10	-57.46	<u>-33.19</u>	-42.58	-44.18	-42.21
1	3	-72.16	-50.97	<u>-15.98</u>	-40.06	-22.60	-56.92	-42.76	<u>-8.82</u>	-33.29	-17.74	-72.21	-46.26	<u>-16.26</u>	-31.31	-19.94	-57.94	-39.55	<u>-10.24</u>	-24.95	-16.32
2	1	-55.43	<u>-37.32</u>	-43.27	-43.29	-41.45	-46.88	<u>-30.94</u>	-36.70	-36.82	-35.30	-56.20	-35.57	-36.33	-37.68	<u>-32.92</u>	-47.12	-29.33	-29.29	-30.01	<u>-26.56</u>
2	2	-49.08	-47.07	-45.17	<u>-37.34</u>	-43.01	-41.28	-39.95	-38.68	<u>-32.57</u>	-37.13	-54.41	-48.18	-44.00	-40.05	<u>-39.73</u>	-42.03	-37.90	-34.10	<u>-30.40</u>	-31.55
2	3	-63.73	-49.05	-33.23	-31.65	<u>-25.70</u>	-49.74	-40.08	-25.05	-23.07	<u>-17.26</u>	-63.97	-43.32	-28.04	-27.62	<u>-24.89</u>	-49.89	-35.66	-20.95	-20.18	<u>-19.66</u>
3	1	<u>-68.32</u>	-68.84	-71.93	-71.16	-71.29	-56.34	<u>-54.93</u>	-57.48	-56.92	-57.48	<u>-69.40</u>	-71.94	-72.43	-72.43	-72.42	<u>-56.72</u>	-57.93	-57.94	-57.94	-57.94
3	2	-71.84	-47.10	-22.95	-30.78	<u>-17.71</u>	-62.49	-43.98	-21.86	-28.76	<u>-16.95</u>	-71.79	-41.49	<u>-18.26</u>	-27.31	-21.02	-62.77	-38.52	<u>-16.86</u>	-24.02	-19.27
3	3	-72.35	-71.84	-59.83	<u>-54.73</u>	-55.24	-64.40	-65.98	-57.14	<u>-52.56</u>	-53.07	-72.43	-71.80	-63.97	<u>-61.77</u>	-62.83	-65.99	-65.85	-59.46	<u>-57.33</u>	-58.77
4	1	-54.83	<u>-50.86</u>	-57.41	-62.43	-58.71	-48.26	<u>-43.88</u>	-49.87	-54.04	-50.51	-54.81	<u>-49.71</u>	-57.41	-61.08	-56.59	-47.41	<u>-42.74</u>	-48.53	-51.20	-48.16
4	3	-69.49	-43.38	-29.43	<u>-25.22</u>	-27.41	-56.24	-36.20	-22.70	<u>-18.81</u>	-21.88	-69.77	-38.12	-23.20	-22.61	<u>-20.73</u>	-56.15	-31.29	-16.80	<u>-15.57</u>	-15.89
5	1	-68.90	<u>-22.26</u>	-44.60	-42.75	-40.74	-61.32	<u>-21.50</u>	-41.94	-40.52	-38.51	-67.29	<u>-23.01</u>	-33.07	-35.16	-29.89	-60.41	<u>-21.50</u>	-29.46	-30.98	-27.38
5	2	-69.84	-54.26	<u>-35.77</u>	-48.07	-40.75	-54.31	-42.78	<u>-21.10</u>	-37.72	-30.23	-69.44	-48.58	-40.71	-45.07	<u>-39.44</u>	-54.20	-37.41	<u>-29.60</u>	-33.32	-30.16
5	3	-70.23	-44.73	-30.63	<u>-25.17</u>	-29.33	-61.00	-40.19	-26.93	<u>-21.44</u>	-26.26	-71.66	-41.34	-21.23	-19.50	<u>-17.81</u>	-61.16	-36.49	-17.38	-15.43	<u>-15.35</u>
6	1	-70.88	-46.20	<u>-16.27</u>	-35.62	-22.96	-57.94	-39.17	<u>-9.11</u>	-29.26	-17.88	-70.51	-38.41	<u>-14.12</u>	-23.23	-15.98	-57.89	-32.36	<u>-8.53</u>	-17.17	-12.80
6	2	-65.57	-49.32	-43.38	-47.52	<u>-42.02</u>	-56.82	-44.38	-38.05	-43.33	<u>-37.08</u>	-70.22	-47.91	-48.39	-52.75	<u>-45.64</u>	-57.70	-41.36	-40.60	-43.83	<u>-38.75</u>
6	3	<u>-53.12</u>	-68.57	-70.88	-71.41	-70.87	<u>-46.26</u>	-59.73	-62.40	-62.66	-62.35	<u>-56.49</u>	-67.30	-71.21	-71.31	-70.36	<u>-48.23</u>	-58.70	-62.09	-62.15	-61.50

Note: Maximum likelihood for each specification underlined. Classification results that are consistent with the baseline specification (spike-logit) are in **bold**.

Subject 3-1 has compliance rates less than 20 percent for all types under both spike-logit specifications, and hence, is deemed as unclassified.