Promises and Lies: An Experiment on Detecting Deception

Jingnan (Cecilia) Chen and Daniel Houser

February 2013

Discussion Paper

Interdisciplinary Center for Economic Science
4400 University Drive, MSN 1B2, Fairfax, VA 22030
Tel: +1-703-993-4719 Fax: +1-703-993-4851
ICES Website: http://ices.gmu.edu
ICES RePEc Archive Online at: http://edirc.repec.org/data/icgmuus.html

Promises and Lies: An Experiment on Detecting Deception

By

Jingnan (Cecilia) Chen, Daniel Houser*

Interdisciplinary Center for Economic Science

George Mason University

13 February, 2013

Abstract: Although economic and social relationships can involve deception (Gneezy 2005), such relationships are often governed by informal contracts that require trust (Berg et al. 1995). While important advances have been made concerning deception in economics, the research has focused little on written forms of communication. Are there certain systematic cues that signal written communications as dishonest? Are those signals accurately detected and used by message receivers? We fill this gap by studying messages written in a novel three-person trust game (we call it the "Mistress Game"). We find that: (i) messages that use encompassing terms, or a greater number of words, are significantly more likely to be viewed as promises; and (ii) promises that mention money are significantly more likely to be trusted. Notwithstanding the latter finding, we find senders who mention money within their promises to be significantly less likely to keep their word than those who do not.

---

* Chen: Interdisciplinary Center for Economic Science, George Mason University, email: jchenw@gmu.edu; Houser: Interdisciplinary Center for Economic Science, George Mason University, email: dhouser@gmu.edu. We gratefully acknowledge the helpful comments from the participants at 2012 International Economic Science Association Conference in NYC, 2012 University of Mainz Workshop in Behavioral Economics in Mainz, Germany and 2012 North American Economic Science Association Conference in Tucson.

All truth is simple, is that not doubly a lie?
- Friedrich Nietzsche

I. Introduction

Many economic and social relationships involve deception (Gneezy 2005). Such relationships are often governed by informal contracts that require trust (Berg et al. 1995). Substantial research has focused on deception in economics (see, for example, Hao and Houser 2013; Erat and Gneezy 2011; Rosaz and Villeval 2011; Kartik 2009; Sutter 2009; Dreber and Johannesson 2008; Charness and Dufwenberg 2006; Ellingsen and Johannesson 2004). Recently, increasing attention has been devoted to the question of whether deception or trustworthiness can be detected (see *e.g.*, Belot et. al 2012; 2010; Darai and Grätz 2010). Although important advances have been made, the research has focused only on face-to-face communication, and has not yet addressed other, *e.g.*, written, forms of communication. In particular, research has not revealed whether certain systematic cues may help signal a written communication as dishonest, and, if so, whether those cues can be detected and accurately used by readers of messages. This paper fills that gap by studying this question within the context of a controlled laboratory experiment.

While trust is essential for an economy, the knowledge of who and when to trust, *i.e.* deception or trustworthiness detection, is equally critical (see, *e.g.*, Belot et al. 2012). Much deception detection literature has focused on verbal and non-verbal interactions (Zuckerman et al. 1981). However, to our knowledge, there have been no studies of deception detection in informal written communication [1] in economics. This is unfortunate, since informal written communication (*e.g.*, via emails, texting, tweeting, or facebooking) plays an increasingly important role in social and economic exchange decisions.

[1] Our interest is in understanding cues used in informal written communication of the sort that people might send in instant messages or other forms of casual (and often electronic) communication. Our focus is not, for example, formal legal documents which are typically constructed with the goal of reducing ambiguity (at least to people trained in reading the contracts).

The effect of informal communication has been widely studied in the context of " cheap talk"[2] (see, *e.g.*, Farrell and Rabin 1996; Crawford 1998). Cheap talk is effective at conveying private information and/or signaling intention; as such, it has proven effective in coordination and coordination-like games (Farrell and Rabin 1996), in impacting the outcome of bargaining games (Croson et. al. 2003), and in promoting trust and trustworthy behaviors (Charness and Dufwenberg 2006). As noted by Farrell and Rabin (1996), cheap talk matters because people respond to it. In this paper, we take a step towards better understanding the nature of people's responses to different types of cheap talk promises; that is, we inform when cheap talk is more likely to evoke responses, as well as the nature of those responses.

Informal written cheap talk occurs frequently during Internet dating[3]. In these cases, the interactions begin with initial informal written message exchanges. The purpose of these exchanges is to build a foundation of mutual trust upon which a real (as compared to virtual) relationship can develop[4] (Lawson and Leck 2006). Evidently, during this process of written exchanges, decisions must be made regarding the trustworthiness of one's partner. Consequently, the ability to write trustworthy-sounding messages, as well as to detect insincere messages, is an important and adaptive skill. In this paper, we further the understanding of cues and the interpretation of cues in informal written exchanges.

It is important to note online dating as an example of an environment where trust and deception are most important to understand. In particular, trust is not a consideration in cases where interests are either fully aligned (as in members of teams during competition or parental care of young children, where trust is not needed) or perfectly misaligned (as in cases of fraud or other zero-sum activity, where trust is not an option). Trust, and the

[2] Communication that has no direct effect on players' payoffs, is costless and unverifiable.
[3] Through for instance, match.com and many other websites.
[4] For anecdotal evidence see "A Million First Dates", http://www.theatlantic.com/magazine/archive/2013/01/a-million-first-dates/309195/?single_page=true

consequent possibility of deception, is critically important in cases where gains from exchange are possible, but there also exist incentives for one side to defect and appropriate the surplus. It is precisely in these situations that people may send informal "promises" of future behavior; these messages must be interpreted to gauge the extent to which they can be trusted.

Research on this topic has appeared in both economics and psychology. As discussed below, key findings from economics indicate that people notice and respond to some cues (for example, gender and presence of a handshake), but not others (*e.g.*, participants' past behavior) (*e.g.*, Belot et al, 2012; Darai et al, 2010). These results, however, are based only on face-to-face communication. The psychology literature studies the same question, but within the context of qualitative cues such as facial movements or expressions (*e.g.*, Ekman, 2009b). The main finding from this literature is that people do not know what to look for to identify cheating, and consequently perform poorly – not much better than chance – at detecting deception.

As noted, the previous literature focuses on face-to-face communication, with little attention paid to informal written communication[5]. In order to fill this gap, we introduce a novel variant of a trust game (building on the hidden action game of Charness and Dufwenberg, 2006). Our game captures an environment with misaligned incentives and opportunities to defect, but also includes potential gains from trade. In this context, we offer participants the opportunity to communicate promises to one another using hand-written messages. We use this design to answer three primary questions. First, are quantifiable features of natural language messages that make a message more likely to be viewed as a promise? Second, are there features of messages that leave some promises more likely to be trusted? Finally, are there objectively quantifiable cues that discriminate promises that will be kept from those that will be broken?

[5] The computer science and linguistics literatures include examples of computer-assisted text analyses. These studies provide evidence of differences between deceptive and non-deceptive texts (see, e.g., Zhou et. al 2002; 2003).

We find, first, that using either encompassing terms or a greater number of words leaves a message significantly more likely to be viewed as a promise. Second, we find that messages that mention money are much more likely to be believed than those that do not. On the other hand, and in answer to our third question, senders who discuss money as part of their promises are significantly more likely to break their promises.

These findings resonate with life experiences. Advertisements often use encompassing words and refer to money benefits. These promises are made in this way, presumably, because they are often believed. On the other hand, these promises are also often broken, in the sense that the advertised monetary savings are not equivalent to the actual benefits received in the transaction.

The remainder of this paper is organized as follows. In Section II, we discuss the related literature. Section III explains the context from which we obtain the message data and also the experimental setup. In Section IV, we report our analysis and results. Section V summarizes and concludes the paper.

II. Related Literature

2.1 Deception

Deception is a socially and economically relevant topic (Mazar and Ariely 2006). "Business people, politicians, diplomats, lawyers, and students in the experimental laboratory who make use of private information do not always do so honestly" (p.384, Gneezy (2005)). For instance, white-collar workers do not always pay for the bagels and donuts they purchase (Levitt 2006); neither do newspaper purchasers on the street (Pruckner and Sausgruber 2008). Even children sometimes report their die roll outcomes dishonestly (Bucciol and Piovesan 2011), let alone students in the experimental lab (see for example, Gneezy 2005; Fischbacher and Heusi 2008; Lundquist et al. 2009; Houser, Vetter, and Winter 2010; Hao and Houser 2013).

Economists are increasingly interested in deception-related topics such as dishonesty, cheating, lying, and promise breaking, and are making advances in understanding these

behaviors. The general finding is that people are averse to lying; nonetheless, there is a non-negligible percentage of people who do lie. Gneezy (2005) implemented a cheap talk sender-receiver game to investigate empirically the effect of consequences[6] on people's tendency to lie. In the game, senders chose to send truthful or false prefabricated messages. The author found that people are averse to lying (averse to choosing the false message) even when there are small benefits to self at the cost of others. Further, Erat and Gneezy (2011) used a modified sender-receiver game and discovered that not only are people averse to "selfish black lies"[7], they are also reluctant to tell "Pareto white lies."[8] In a modified trust-game environment where promises could be made, Charness and Dufwenberg (2006) found that about a quarter of the subjects broke their promises[9].

2.2 Deception detection/Promise Identification

Since deception is a common part of many social and economic interactions, a natural question arises: can deception be detected? More specifically: i) are there systematic features associated with deceptive actions? ii) Can people correctly identify and utilize those features to detect deception? The converse side of these questions is also quite interesting: are there certain attributes of actions that are considered signs of sincerity and trustworthiness?

Deception detection, especially lie detection, is widely studied in psychology. Generally speaking, psychologists find that there are behavioral clues to deceit, such as facial

[6] Changes in wealth resulted from the lie.
[7] Lies that benefit self but harm other.
[8] Lies that benefit both self and other.
[9] In an effort to account for the empirical observations, some scholars have suggested that people experience variable disutility from guilt, where the disutility depends on the degree subjects think they let others down—the greater the let-down, the more disutility they induce (*e.g.*, Charness and Dufwenberg (2006; 2010)). Some propose self-concept maintenance theory to account for dishonest behaviors from mostly honest people (Mazar et. al. 2008); that is, people suffer small fixed disutility from lying and make their deception decisions based on the net benefits of such action. Others, however, hypothesize that there may be two types of people: one type will never lie, and the other type are able to lie because they do not suffer guilt (see for example, Vanberg 2008; Ellingsen et al. 2010).

movements (Ekman and Friesen 1974), voice and speech pattern (Chichester 2008). People who can accurately detect liars seem to use some of the physical clues (Ekman and O'Sullivan 1991). A person uses all of the clues correctly can achieve over 80% accuracy in deception detection (Ekman, O'Sullivan, and Frank 1999; O'Sullivan and Ekman 2005; Ekman 2009b). While people are not good at catching lies based on demeanor (not much better than chance), groups of professionals (*e.g.*, United States Secret Service Professionals, interrogators) with special training can perform significantly better[10] (Ekman 2009b).

The common setups in the above-mentioned psychology studies generally include actors (usually students) who are instructed to tell the truth or a lie, and observers who evaluate the trueness of the actors' statements upon watching the videotaped recordings (see Ekman 2009a; or Ekman 2009b for a short review). For most of those studies, neither the actors nor the observers were incentivized to perform (Zuckerman, DePaulo, and Rosenthal 1981; Vrij et al. 2004). Later research (see *e.g.*, Ekman and Friesen 1974; Ekman, O'Sullivan, and Frank 1999) improved the experimental design by, for instance, making the lies more relevant and voluntary and increasing the stakes for success or failure (for both actors and observers, sometimes even punishment for actors if they are considered lying). The findings, however, did not differ.

Another focus of deception detection concerns people's ability to predict trustworthiness through face-to-face communication or through observing such interaction. Most of the early research on this topic centers either on communication's ability to facilitate cooperative outcomes (see, *e.g.*, Cooper et al. 1989; Cooper et al. 1992; Isaac and Walker 1988; Miettinen and Suetens 2008; Charness and Dufwenberg 2006; Charness and Dufwenberg 2010; Belot, Bhaskar, and van de Ven 2010), or on people's ability to predict outcomes accurately in light of that communication. This research does not delve

[10] Ekman (2009b) also suggests that regular people can use micro experssion training tools (METT) to be trained to recognize micro-experssion (www.paulekman.com) and use them to successfully detect deception. A popular show – Lie To Me - on Fox is inspired by Ekman's research in lie detection using micro-experssion.

into questions such as the nature of cues available to participants in these environments, or whether participants identify and use those cues.

Despite the success of communication in promoting cooperation, research indicates that people are not good at predicting people's decisions following communication[11]. Questions remain as to why people are not more successful at detecting deception given its important role in daily life. Could it be that people do not know what cues to look for, or that they able to detect cues, but unable to use them correctly, or both?

Ekman (2009b) suggests that humans (without training) lack the ability to identify traits correlated with deceit. Reasons include lack of exposure to lying during our ancestral history, lack of adaptive value for lying detection, propensity to trust rather than doubt, and wanting to be misled rather than know the truth. Wang et. al. (2010) used a sender-receiver game in which senders had the incentive to lie to demonstrate that certain systematic features are associated with lying subjects; these included payoff lookup patterns of the eyes and pupil dilation. The authors also calibrated that if the receivers were able to use the predicative features (the eye movement and the messages the senders sent), they could have earned up to 21% more. Additionally, Darai and Grätz (2010) used data from the British television game show "Golden Balls" to discover that certain features of the game/player, such as stake size and the occurrence of handshake, are highly predictive of the rate of cooperation in the face-to-face prisoners' dilemma game.

Furthermore, Belot et. al. (2012) reports that subjects in an economic experiment were able to use some objective cues (while ignoring some of the others) to help correctly

[11] See for example, Dawes et. al. (1977), Frank et. al. (1993) and Brosig (2002) implemented prisoner's dilemma with pre-game face-to-face communication, after which subjects were asked to predict their opponents' move. They found that people perform slightly better than chance in their predictions. Similarly, Ockenfels and Selten (2000) looked at the two-person bargaining game with incomplete information and had on-lookers observe the bargaining process. The on-lookers' detection accuracy exceeds chance as well. The authors attribute the results mostly to the success of on-lookers utilizing the objective feature of the bargaining process (*i.e.*, the length of bargaining time).

detect deception and predict trustworthiness. The authors made a novel use of data from a high-stakes prisoner's dilemma game show to investigate trustworthiness detection. Subjects watched the clips and rated the likelihood that players cooperated pre- and post-communication on a scale of zero to one (with the increment of 0.1). The authors discovered that subjects were able to use some[12] objective features of the game's players (such as gender and past behaviors) to make pre-communication predictions. Although subjects did not seem to improve their overall predictions after watching the communication between the players, they did respond positively to the "elicited promise" communication group[13]. The authors concluded that previous research might have underestimated people's ability to discern trustworthiness.

In sum, most research to date has emphasized people's ability to detect deception or trustworthiness in face-to-face[14] encounters. Face-to-face interaction is a very rich and relevant environment to access people's ability to detect deception; however, the environment may be too complex to enable one to draw inferences regarding the reasons for people's performance. The reason is that too many factors are at play, including facial expressions, body movements, hand gestures, and the language being used, with some of them quite hard to measure. It is difficult for researchers to pinpoint the information people acquire or use, or what information is actually perceived by subjects[15]. In addition, there may be several confounding factors. For example, in Belot et. al. (2012), the authors infer that subjects are able to correctly predict females as relatively more trustworthy than males. However, it may be the case that: 1) females are more sensitive to guilt, and thus less likely to lie (and more trustworthy in general) (*e.g.*, Dreber and

[12] The subjects weren't able to recognize or use all the objective features of the game show, *e.g.*, the relative contribution to the prize.
[13] Belot et. al. (2012) categorized communication into three different groups: no promise-where no promises are made, voluntary promise-where players voluntarily make promises, and elicited promise-where the subjects were prompted by the game show host to indicate their intention to either cooperate or defect.
[14] In some of the cases, face-to-face encounter includes subjects watching a video recording of the game players.
[15] As noted in Ekman et. al. (1999), successful subjects were able to use facial clues to detect liars, as opposed to others who were not able to do so when presented with the same video recordings.

Johannesson (2008), Erat and Gneezy(2011)); or 2) females are less capable of concealing their emotions (*e.g.*, Papini et al(1990)) in their facial expressions, and thus are more likely to be considered trustworthy by observers.

While we note above that previous experiments have established that people perform poorly at distinguishing truth from lies in face-to-face interactions, previous research has systematically investigated neither the causes of these relatively low success rates nor the ability to predict trustworthiness with other forms of communication (*e.g.*, online written communication such as that used in dating websites). In these cases, deception can have significant impact. This paper contributes to the literature by using a controlled laboratory experiment to investigate cues of deception (untrustworthiness) as well as potential explanations as to why people do or do not detect untrustworthiness. Finally, our analysis offers new insights into how to convey trustworthiness.

III. The Game, Messages and Evaluations

3.1 The Mistress Game

We devised a novel three-person game[16] to generate written messages (these messages were evaluated in a subsequent experiment, described below). The extensive form of the Mistress Game is shown in Figure 1. Payoffs are in dollars.

The Mistress Game builds on the hidden action trust game (Charness and Dufwenberg, 2006). In our game, chance (the die roll) is replaced with a strategic third player C. Our payoff structure offers incentives that suggest the following interpretation.

[16] This game is a modification of an extended three-person trust game with different multipliers for different trustees. Related games are Charness and Dufwenberg (2006) – two-person trust game with a hidden action, Sheremeta and Zhang (Sheremeta and Zhang 2010) and Rietz et al. (2011) – sequential three person trust game, Cassar and Rigdon (2011) – three person trust game with one trustee two trustor or one trustor two trustee, and Bigoni et al. (2012) – two person trust game with an add-on dominant solvable game between the trustee and a third player.

We denote Roles A, B and C as Wife, Husband and Mistress[17] (or Husband, Wife and Paramour): Wife and Husband consider whether to form a union. If no "marriage" occurs, then both parties receive the outside option payoff (to being single) of $5. In this case, the Mistress receives $10. If a marriage contract is formed, a trust relationship emerges, and the payoffs to this relationship depend on the Husband's decision. The Husband (Role B) is faced with a dilemma—either to stay with the current trust relationship (corresponding to B's *Out* option) or form an additional trust relationship with a third person and enjoy a potentially higher payoff (corresponds to B's *In* option). Note that Wife is NOT any better off (maybe even worse off) by Husband's choosing IN. If the Husband chooses to stay with the Wife (corresponding to the strategy profile (*In, Out, Left/Right*)), both Wife and Husband are better off (with the payoff of $10 for each), and Mistress (who has no move) again earns the outside option of $10. The strategy profile (*In, Out, Left/Right*) corresponds to the situation where the marriage contract is enforceable. However, the marriage contract may not be enforceable. Indeed, the Husband's choice may not be observable to the Wife, depending on the Mistress's decision. Our game captures this as discussed below.

If the Husband chooses to form a new trust relationship with the Mistress (corresponding to B's *In* option), the Mistress can either be cooperative and reciprocal by choosing *Left*, or defect by choosing *Right*. Note that if the Mistress chooses *Left*, the Husband's behavior is unknown to the Wife. However, if the Mistress chooses *Right*, not only does the Husband receive nothing from the newly-initiated trust (the Mistress takes all), the Wife is also impacted and receives nothing. In this case, the Wife knows the Husband's choice. Note that the Wife may foresee such outcomes and choose not to enter the trust relationship with the Husband. The players' choices *Out, In* and *Right* describe those possibilities. It is easy to verify that the sub-game perfect equilibrium of this game for selfish and risk-neutral players is (*In, Out, Right*), which is also inefficient.

[17] This game is not intended to capture the intricate and complicated features of marriage; the analogy used here is merely to facilitate understanding of the tradeoffs that each player faces in the game. More importantly, we are interested in the dynamics of multi-player trust relationship.

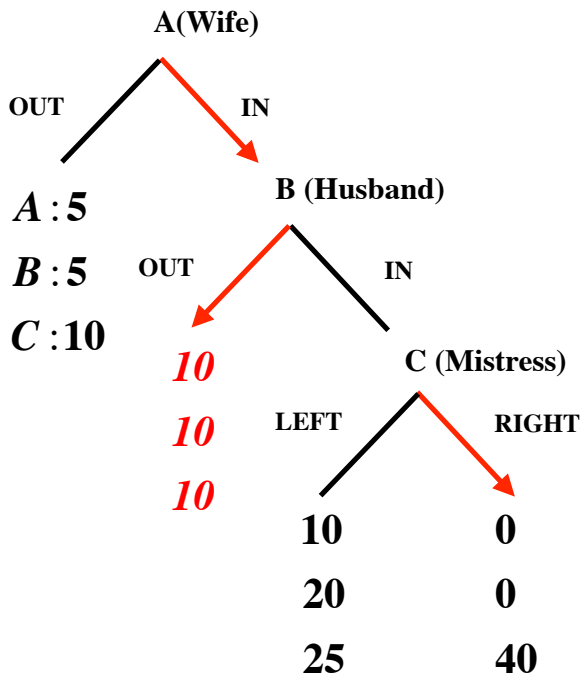

Figure 1. The Mistress Game

3.2 The Messages

In addition to the regular no-communication game play, we also introduce one-sided pre-game communication to the environment: the players have an opportunity to send a handwritten note to their counterparts. In particular, for the purpose of this paper, we focus on the messages from the Mistress (Role C) to the Husband (Role B) under two different environments: single message and double message.

3.2.1 Single message environment

Before the subjects play the Mistress Game, the Mistress (Role C) has the option to write a message to Husband (Role B). The experimenter then collects the messages and passes them as shown in Figure 2. That concludes the communication phase, and the subjects start to play the game[18].

[18] The authors also implemented other versions of the communication treatment (*e.g.*, only Role B sends messages to Role A). These data are reported in Chen and Houser 2012. Here we only focus on the C to B message treatments.

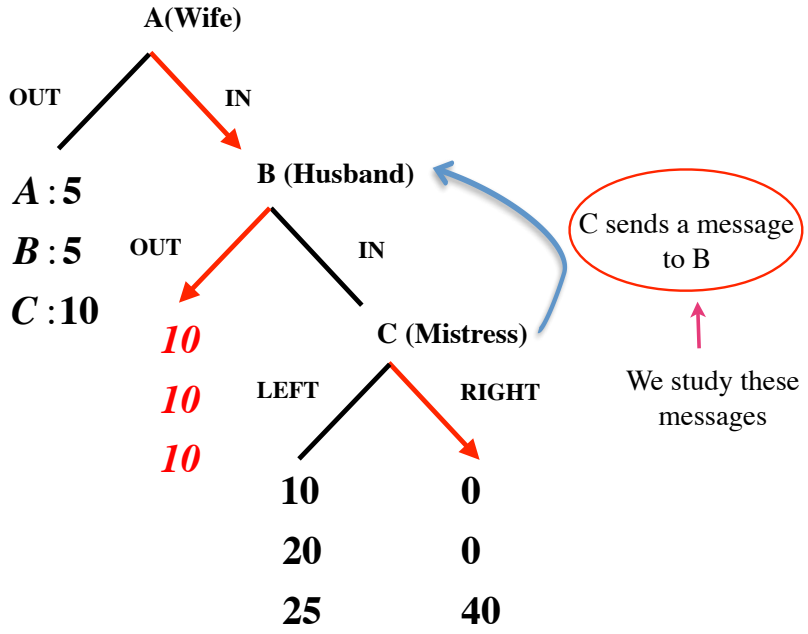

Figure 2. The Single Message Communication Phase

3.2.2 Double message environment

As shown in Figure 3, the double message environment is similar to the single message environment, except that the opportunity for the Mistress (Role C) to send a message to the Husband (Role B) comes as a surprise.

It is common knowledge from the beginning of the experiment that the Husband (Role B) has an opportunity to send a hand-written message[19] to the Wife (Role A). After the messages are transmitted, the experimenter announces a surprise message opportunity: the Mistress (Role C) can also send a message to the Husband (Role B). The experimenter waits for the messages to be written by Role C and then passes the messages on to their paired Bs. Upon completion of the message transmission, subjects start to play the game.

[19] It is well understood amongst subjects that they cannot write anything that is self-identifiable, and the experimenter monitors the messages to make sure this rule is followed.

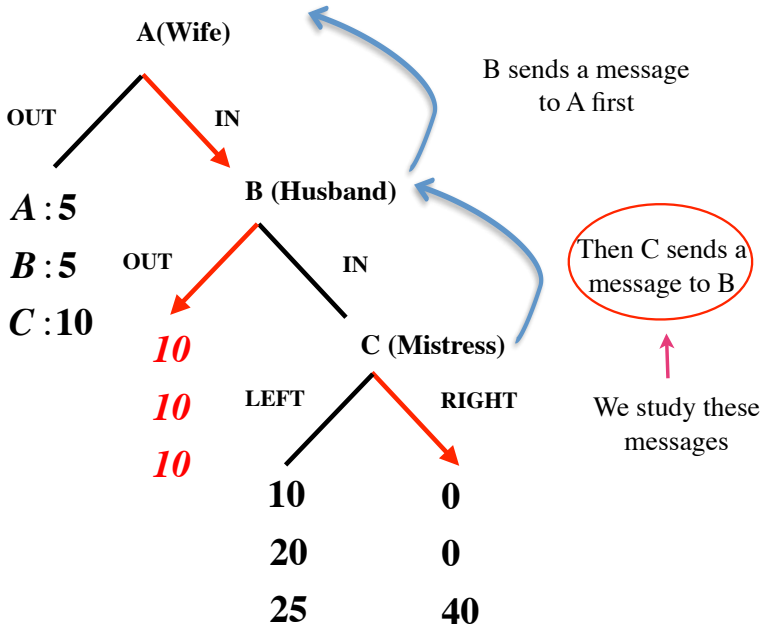

Figure 3. The Double Message Communication Phase

In both single and double message environments, the Mistress (Role C) is better off when the Husband (Role B) chooses IN; therefore, it is natural to assume that the Mistress (Role C) would use the messages as a means to persuade the Husband (Role B) to choose IN. However, the two environments also depart significantly from each other. Specifically, in the double message environment, where everybody knows that the Husband (Role B) has already sent a message to the Wife (Role A), it is reasonable to presume that the Husband (Role B) might have conveyed his intention to stay with Wife (Role A) and choose OUT. Therefore, it is very likely the case that Mistress (Role C) needs to do a better job in convincing Husband (Role B) to choose the Mistress instead by choosing IN[20].

3.3 Experimental Design, Procedure and Hypothesis

3.3.1 Experimental Design and Procedure

The evaluation sessions were conducted in the experimental laboratory of Interdisciplinary Center for Economic Science at George Mason University[21]. We recruited 45 evaluators from the general student population (22 evaluators to evaluate

[20] We did find some evidence suggesting that Mistresses (Role C) worked harder in crafting their messages, as shown in section 5.
[21] The game sessions were also conducted in George Mason University.

messages from single message environment and 23 to evaluate messages from double message environment). None of the evaluators had previously participated in the Mistress game experiment. Average earnings were $18 (including the $5 show-up bonus); sessions lasted about one hour.

Before reviewing any messages, evaluators were acquainted with the Mistress Game and were provided with a transcript of the Mistress Game instructions for either the single message environment or the double message environment. A quiz was administered to ensure that all the evaluators understood their tasks as well as the context when the messages were written.

There were in total 20 and 32 messages collected from the Mistress Game single and double message sessions respectively, all of which were scanned into PDF files and displayed on the computer screen in a random order for the evaluators to look through. Each evaluator worked on all messages independently inside their own visually separated cubicles. They were not given any information regarding the decisions of the message-senders or their partners. Nor were the evaluators given any information regarding the purpose of the study, or the hypotheses of interest. Evaluators were instructed to first classify each message as either "Promise or Intent" or "Empty Talk,"[22] and then make guesses with regard to what the message senders actually did. We followed Houser and Xiao's (2010) coordination classification procedure to incentivize the evaluation tasks. For the first task, two messages were randomly chosen for payment, and the evaluators were paid based on whether their classifications coincided with the medium choice of the population; similarly, another two messages were randomly chosen for the payment of

[22] We adapted Houser and Xiao (2010) weak promise treatment procedure to instruct evaluators on how to categorize promise/intent or empty talk. On the instructions, it is stated clearly that a message should be categorized as a statement of intent or promise if at least one of the following conditions is probably satisfied: 1) The writer, subject C, indicates in the message he/she would do something favorable to subject B or refrain from doing something that harms subject B; 2) The message gives subject B reasons to believe or expect that subject C would do something favorable to subject B or refrain from doing something that harms subject B. A message should be coded as empty talk if none of the above conditions are satisfied.

the second task, and the evaluators were paid based on whether their guesses match the actual behavior of the message senders. Upon completion of the evaluation tasks, the evaluators were given a survey with questions such as how they made their classification or guess decisions. The experimental instructions are available as an appendix to this paper.

3.3.2 Hypotheses

One advantage of written messages is that they have fewer cues, in relation to face-to-face communication, that one can make use of and quantify. In view of the literature, we developed several hypotheses in regard to cues of written messages that may impact both evaluators' and message senders' behaviors:

Hypothesis 1: The mention of money can impact: (i) the belief that a message is a promise; (ii) the belief a promise will be kept; and (iii) the trustworthiness of the message sender.

We hypothesize that the mention of money impacts how evaluators assess the trustworthiness of a message. The reason is that the mention of money contains information that is relevant to game play, and thus gives credibility to the message, perhaps makes the sender seem more trustworthy; consequently, the message is more likely to be evaluated as a promise (see, *e.g.*, Rubin and Liddy, 2006). Similarly, if the message is viewed as a promise, then the fact that the promise includes money may again add credibility to the message, impacting the evaluators' perceptions of whether the promise will be kept.

The mention of money could also have the effect of "monetizing" the exchange, and thus could impact message senders' behaviors. Such an effect is suggested by a sizable "crowding out" literature (see for example, Ariely and Bracha 2009; Lacetera and Macis 2010; Mellstrom and Johannesson 2008; Gneezy and Rustichini 2000; Gneezy and Rustichini 2000; Fehr and Falk 2002; Li et al. 2009; Houser et al. 2008). This literature emphasizes the idea that monetizing choices may crowd out extrinsic incentives, shift

decision-makers' perception of the environment into a "business" frame, and focus their attention on self-interested decision making.

Hypothesis 2: The use of encompassing words such as "we" or "us" can create an "in-group" effect that promotes the perception that a message is a promise and that the promise will be kept. Further, the use of these words may be associated with the sender's likelihood of keeping her promise.

The use of encompassing words can foster a common social identity among message senders and receivers (Hall 1995). This sort of "in-group" effect can impact the sense that a message is a promise, as well as the belief that a promise will be kept. Indeed, being part of an in-group can also impact reciprocity decisions. A rapidly growing literature supports these observations. For example, Kimbrough et al. (2006) found that it is more common to mention "we" or "us" during chat with in-group than out-group members, and that the mention of these encompassing words is positively correlated with cooperation and the willingness to make and keep promises to do personal favors. In-group effects are strong and systematic. While the literature is extensive, recent contributions include Bernhard et al. (2006), which demonstrates that people behave more altruistically when matched with in-group members than when matched with distinct social groups. Some recent contribution in economics such as Chen and Li (2009) and Pan and Houser (2013), finds similar results with students in the lab.

Hypothesis 3: Longer messages signal greater effort, and this can affect the perception that the message is a promise, that the promise will be kept, and also correlate with sender's decisions.

Longer messages signal that the writer has made greater effort. Some studies suggest that a person who invests greater effort into composing a message may seem more trustworthy, and yet be less trustworthy. For example, Wood et al. (1985) notes that the perception of trustworthiness is positively associated with longer messages, though they also report that those who send longer messages are in fact less likely to keep their

promise. That is, they find that there is an inverse correlation between word length and promise-keeping among senders of messages, but a positive effect of word length on trust for receivers of the messages. On the other hand, Ockenfels and Selten (2000) observe the positive correlation between detection accuracy rate and greater effort (in their case, longer negotiation time) exerted during the bargaining process.

IV. Data Analysis and Results

4.1 Data and Descriptive Statistics

We obtained 52 messages in total from the communication phase of the Mistress Game: 20 messages from Single, and 32 from Double[23], all of which were classified by our evaluators. Among the 20 messages from Single, 80% were categorized as promises or statement of intent[24], 75% of the 32 messages from Double were classified as including a promise or intent[25] (See Table 1).

Table 1. Message Evaluation Results

	Single Msg	Double Msg
Promises/Statement of Intent	16 (80%)	24 (75%)
Empty Talk	4 (20%)	8 (25%)
All Messages	20	32

[23] The messages are available from the authors upon request.

[24] A message is coded as a promise if a majority of the evaluators (more than 50%) coded the message as so.

[25] Our findings regarding promise frequency are consistent with previously reported data. For example, Charness and Dufwenberg (2006) classified 57% of their messages from B in the (5,5) treatment as promises; Vanberg (2008) classified 85% of the messages as promises in No Switch and 77% of the messages as promises in Switch. Using the same procedure as do we, Houser and Xiao (2009) found that 74% of the B messages from Charness and Dufwenberg (2006) (5,5) experiment were categorized as promises by the evaluators in their weak promise treatment.

Table 2. Comparison of the Messages From Single and Double

	Observations		Mean		Z Stat
Environment	Single	Double	Single	Double	
Mention of Money[26]	20	32	0.20 (0.09)	0.31 (0.08)	0.88
Mention of "We/us"[27]	20	32	0.20 (0.09)	0.34 (0.09)	1.10
Word Count[28]	20	32	7.85 (1.47)	14.78 (2.45)	1.93*

Standard errors are reported in the parentheses. The Z statistic derives from two-sided Mann-Whitney tests. * indicate $p < 0.10$ two tailed tests.

The messages from both environments are statistically identical in terms of mentions of money and we/us. However, they differ in terms of message length. As shown in Table 2, around a quarter of the messages include money mentions, and less than one third involve the use of "we", "us" or "let's". Messages from Double are significantly longer than those from Single. This may stem from the fact that in the double message environment Mistress understands that Husband communicated with Wife, and thus it may be more difficult to convince Husband to select *In*. Consequently, they exert more effort and write longer messages.

4.2 Perceived Cues For Trustworthiness From the Observer

4.2.1 Perceived trustworthiness of the message

In this section, we investigate objective features that receivers perceive as indicative of more trustworthy messages. In particular, we attempt to discover whether any of the objective features of the messages discussed above are significantly (positively or negatively) correlated with whether the message was classified as a promise, and, if so, the extent to which that promise is trusted. Our analysis is based on regressions with

[26] Mention of money is a binary variable; it is coded as 1 if there is any money/payoff related discussion in the message (payoff for the game, benefit from the game, and so on) and 0 otherwise.

[27] Mention of we/us is also a binary variable: =1 if in the message the send used "we", "us" and the abbreviated form, e.g., "let's", and 0 otherwise.

[28] Word Count is the number of words in the messages.

pooled or partially pooled data, according to the outcome of tests described below.

To begin, we investigated whether the evaluation data from the single and double message environments could be pooled. The reason to test is that evaluators may interpret messages emerging from different contexts in different ways. To assess whether pooling was appropriate, we performed Tobit regression analyses with the frequency with which evaluators thought a message was a promise as the dependent variable, and the three above-mentioned objectively quantifiable variables and the constant all interacted with a dummy for the single message treatment (for a total of eight variables on the right hand side).

$$Y = X'\beta + (DX)'\gamma + \varepsilon$$

where Y (all elements of which lie between zero and one) is the frequency with which a message is categorized by the evaluators as a promise, X_i is a vector of observable characteristics of message i, D' is a dummy variable for single message environment, $D = \begin{bmatrix} 0\ if\ belongs\ to\ double\ msg\ environment \\ 1\ if\ belongs\ to\ single\ msg\ environment \end{bmatrix}$, and ε is the idiosyncratic iid error term.

The results indicate that it is appropriate to pool the slope variables (F-test, $P = 0.45$), while the constants are statistically significantly different (F test, $P = 0.08$). This implies that the context (Single or Double) affects the chance that message receivers believe a message is a promise. We report pooled regression results (including only the dummy for the constant) in Table 3. Subjects are 15% more likely to consider a message as a promise if encompassing words such as "we" or "us" are mentioned; longer messages are significantly more likely to be regarded as promises; on average, a message from the Mistress under the single message environment is 19% more likely to be considered a promise than an otherwise identical message from the Mistress but under the environment where Husband has previously sent a message to Wife[29].

[29] We also performed a panel data analysis with random effects, and the results are qualitatively identical. Details available from the authors on request.

Table 3. Message Classification and Perceived Cues

Dependent Variable: Frequency Considered As Promise	(1)
Mention of Money	-.03
	(.12)
Mention of We/Us	.15***
	(.05)
Word Count	.01***
	(.00)
Single Msg Treatment	.19***
	(.01)
No. of Observation	52

Standard errors (clustered by treatment) are reported in parentheses, *** indicates significance at the 1% level.

4.2.2 Promises

Next, we turn to those messages that were coded as promises by the majority of the evaluators. Our goal is to understand the cues that are used by the evaluators in guessing whether a promise (as agreed by the majority) is likely to be trusted. As with the previous analysis, we again had to consider whether the guessing data were appropriate to pool. To assess this, we performed a regression analysis exactly analogous to that described in section 4.2.1 using frequency of trust[30] as the dependent variable. Consistent with the coding data, we found that it is appropriate to pool the slope variables (F test, $P = 0.88$), however, the constants are statistically different (F test, $P = 0.01$). This implies that the context also affects the chance that message evaluators believe a promise will be kept. In the following, we focus on the pooled analysis.

Column 2 in Table 4 shows how characteristics of messages determine the evaluators' guesses. We find that evaluators are significantly more likely to trust the promise when it mentions money, uses encompassing words, and is longer. For example, a promise with 10 additional words is 4 percentage points more likely to be trusted, all else equal.

[30] The average trust rate is defined as the percentage of evaluators who believed the message sender chose *Left* (the cooperative option).

Table 4. Perceived Cues and Trust for Promises

Dependent Variable: Frequency of Trust For Promises	(1)
Mention of Money	.02*
	(.01)
Mention of We/Us	.03***
	(.01)
Word Count	.004***
	(.00)
Single Message	.13***
	(.01)
Number of Observations	40

Standard errors (clustered by treatment) are reported in parentheses.
* and *** correspond to 10% and 1% significance levels, respectively.

Actual Cues For Promise Trustworthiness For The Senders

We now investigate which cues predict senders' actual decisions. As with the previous analyses, we again investigated whether we could pool data from the Double and Single treatments. We performed a Probit regression analysis analogous to that described in section 4.2.1 with the Mistress's actual behavior as dependent variable[31]. We found that the actual behavior of the sender can be pooled (F test, $P = 0.42$). This implies that the message senders' behavior is invariant to context.

As shown in Table 5, broken promises mention more money, use more encompassing words, and also include more words. Next, we control for the possible partial correlations between the cues and report our results in Table 6. Mention of money is the single best predictor of senders' defections: Mistresses are 35% more likely to defect when they mention money in their messages.

Our evaluators identified mention of money as a cue of senders' actions; however, they

[31] We estimate a Probit model where y indicates the decision to choose *Left* (y=1) or *Right* (y=0) and z the corresponding latent variable: $Y = 1\ if\ z = \beta' X + \gamma'(DX) + \varepsilon > 0$ and $y = 0$ otherwise.

used the cue in the wrong way. In addition, the receivers picked up on both the mention of "we/us" and word count as positive indicator of senders' trustworthiness. In contrast, both cues were more likely to present untrustworthiness. In particular, as seen in Table 6, evaluators used cues in a statistically significantly incorrect way in all three cases.

Table 5. Actual Cues For Promises

	Promise		
	Kept	Broken	Z Stat
Mention Money	.16	.60	2.84***
	(.03)	(.02)	
Mention "We/Us"	.24	.60	2.25**
	(.02)	(.03)	
Word Count	12	19.07	1.60
	(1.97)	(4.41)	
Observations	25	15	

The Z statistic derives from two-sided Mann-Whitney tests of the null hypothesis that means in Kept and Broken are identical. *,**,*** indicate $p < 0.10$, 0.05 and 0.01, respectively, two-tailed tests.

Table 6. Actual Cues and Perceived Cues

Dependent Variable:	Actual Realization	Evaluators' Prediction
Cooperative Decision	(1)	(2)
Mention of Money	-.35***	.02*
	(.07)	(.01)
Mention of We/Us	-.16	.03***
	(.23)	(.01)
Word Count	-.003	.004***
	(.005)	(.00)
Single Message		.13***
		(.01)
No. of Observation	40	40

Standard errors are in parentheses. * and *** correspond to 10% and 1% significance levels, respectively. Column 1: bivariate probit estimates, marginal effects, standard errors clustered by treatment. Column 2: Tobit estimates, with standard errors clustered by treatment.

4.3 A Closer Look

Table 7 reports the results of evaluators' guesses regarding whether the message would be believed to lead to a cooperative action, and also whether the subsequent action was

actually cooperative. We find that among the messages that were identified as promises, 70% of evaluators believed that message senders kept their promise (choose *Left*). This belief is statistically identical to the overall actual 63% of promises were kept. A different picture emerges, however, when one considers promises that included mentions of money, encompassing terms, or were greater than median length. In these cases, evaluators were substantially over-optimistic regarding the likelihood that the promise would be kept. In particular, while evaluators believed roughly 75% of these promises would be kept, in fact at most 52% of such promises were actually kept. The differences between evaluators' beliefs and actual behavior are statistically significantly different in these cases. In contrast, for the messages that are identified as empty talk, only 25% of the evaluators believe that the message sender chose *Left*. This is statistically indistinguishable from the one third of senders who did actually choose left. In contrast with promises, beliefs are statistically correct in all of three sub-categories of messages.[32]

As for the accuracy rate, overall, 56% of evaluators were able to make correct predictions based on the messages; however, when considering messages categorized as promises, about the same rate of the evaluators were able to make the correct predictions, while 61% of the evaluators predicted the sender's decisions correctly for the empty talk messages. When we further break down the data, it is clear where mistakes were made: evaluators placed higher trust on promises that mentioned money, used "we/us" and were longer, while at the same time those messages were most likely to be defected upon. In contrast, the empty talk messages that did not mention money or use encompassing words, or were shorter were also less trusted by evaluators, consequently, the evaluators achieved higher rates of accuracy.

[32] These results are consistent with earlier findings by Belot et al (2012).

Table 7. Prediction By Receivers: Summary Statistics

Message Type	Obs	Average Prediction[33]	Actual Rate of Cooperation[34]	T- Stat[35]	Rate of Accuracy[36]
Promises/Statement of Intent	40	.70(.02)	.63(.08)	0.96	.56(.02)***
Money Mention=1	13	.75(.02)	.31(.14)	3.30***	.46(.03)
Us Mention=1	15	.76(.02)	.40(.14)	2.71**	.52(.03)
Word Count = Long	25	.74 (.02)	.52(.10)	2.17**	.58(.03)***
Empty Talk	12	.25(.04)	.33(.15)	0.61	.56(.03)**
Money=0	11	.25(.05)	.27(.15)	0.18	.60(.03)***
Us=0	12	.25(.04)	.33(.15)	0.61	.56(.03)***
Word Count = Short	11	.24(.05)	.27(.15)	0.25	.59(.03)**
All Messages	52	.60(.03)	.56(.07)	0.56	.56(.04)

Standard errors are in the parenthesis. *,**,*** indicate $p < 0.10, 0.05$ and 0.01,

V. Discussion

This paper has drawn attention to the importance of understanding cues for deception (or honesty) in natural language written messages. It is well established that people respond to cheap talk communication; however, some communication works well, while some does not. We further this literature by investigating the type of cheap talk communication that has the most positive effect. We conducted a laboratory experiment in which people could offer written promises of cooperative actions. The messages were evaluated by independent observers, and we used these evaluations, as well as the behaviors in the

[33] The average prediction is defined as the percentage of population that believes the promise is kept.

[34] Actual rate of cooperation is defined as the percentage of messages that are followed by a cooperative move from the message sender.

[35] The statistics indicate the t-test for the null hypothesis that the Average Prediction and Actual Rate of Cooperation are equal.

[36] We define the rate of accuracy as the average percentage of right guesses for all the evaluators (the guess matches the actual behaviors of the message senders). The * indicates significance of two-tailed tests under the null hypothesis that the rate of success is chance (0.5).

game, to answer three questions: i) are there objective cues that correlate with a person's willingness to break a promise? ii) do people recognize these cues? iii) do people use the cues correctly?.

We found that systematic evidence that: (i) there are cues that correlate with promise-breaking; (ii) people do recognize these cues; (iii) people do not always use those cues correctly. In particular, we found that; (i) a message was more likely to be trusted as a promise if it included encompassing words and included more words; and (ii) promises that mentioned money were more likely to be believed; but (iii) that promises that mentioned money were more likely to be broken.

Moreover, we were surprised to find that messages in Double were less likely to be trusted, all else equal, than messages from Single. There are a least two explanations for this. First, it may be that the independent evaluators hold the double-message promises to a higher standard of credibility than those in the single-message environment. One reason is that in double message Mistress must convince Husband to break his previous promise, and does so by offering Husband a new promise. In contrast, in the single message treatment, Husband need not break a previous promise, so evaluators may view the promise from Mistress as needing to be less strong, all else equal, in order to be equally credible. In fact, we find that the messages written in the two treatments are largely identical, except that the messages are a bit longer in the double-message treatment.

Note that in that explanation, Mistresses are viewed as equally trustworthy in both environments, but the messages are held to different standards. Alternatively, the Mistresses in the double message environment may be viewed as less trustworthy. The reason is that the Mistresses are choosing to encourage Husbands to violate a previous promise, and this might lead Mistresses to be viewed by the external evaluators as unethical and untrustworthy. Therefore, all else equal, messages from the double-message environment would be less trusted than those from the single-message environment. Our design cannot distinguish these explanations, but it would be profitable to do further research to disentangle the impact of context on perceptions of

trustworthiness.

Our results might explain some patterns in previously published data. For example, Charness and Dufwenberg (2010) offered new data on their hidden action trust game (Charness and Dufwenberg, 2006) and found that, in contrast with their original data, the bare statements "I will not roll" or "I will roll" do not promote trust or cooperation. Charness and Dufwenberg indicate that this might be due to the impersonal nature of the message. Another factor might be that these statements do not include encompassing terms (*e.g.*, we or us), and are quite short. The results of our paper suggest that both of these features would make any message, personal or otherwise, less likely to be viewed as a promise.

Our results have clear implications for a wide variety of areas. One is political campaign, where written slogans and rapid communication are typically making various types of promises, and are required to win an election or pass a certain policy platforms. Our findings provide an explanation for the popularity of catch-phrases that use words such as "we" or "together," which presumably are used by candidates in an effort to inspire confidence in the candidate's platforms. Another important example relates to the receivers of promises that include mentions of money. For example, billboards advertising large monetary benefits (discounts or savings) to people who choose to shop at a particular retail location should be aware that such promises may be likely to be broken, and that the reality of the savings may be less than the advertised amount[37]. Our results indicate that consumers of advertisements should be especially cautious of promises that include specific monetary commitments.

Our study is only a first step on this important topic, and is limited in a number of ways. One is that the promises in our environment all related to money, while in many natural contexts promises either are not explicitly about monetary payoffs or, even if so, it would be unnatural to refer to money as part of the promise process. Similarly, we studied a

[37] For example, one highway billboard near us reads: "$700 Cash today, the Ca$h Store". Preceding the "$700" there is an almost entirely unnoticeable "Up to."

particular game within which these promises were made, and different games may lead people to use or to recognize different cues that we discovered, or to use or recognize the same cues differently. Finally, our results were derived from a particular cultural environment. The same games played with different cultural groups may generate different types of cues (*e.g.*, some cultures may be reluctant to use "we" or "us" with strangers.) There is no question that cross-cultural research on this topic will be profitable.

Reference:

Ariely, Dan, and Anat Bracha. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *The American Economic Review* 99 (1): 544–555. http://www.jstor.org/stable/10.2307/29730196.

Belot, Michèle, V Bhaskar, and Jeroen van de Ven. 2012. "Can Observers Predict Trustworthiness?" *The Review of Economics and Statistics* 94 (1): 246–259. http://www.mitpressjournals.org/doi/abs/10.1162/REST_a_00146.

Belot, Michèle, V. Bhaskar, and Jeroen van de Ven. 2010. "Promises and Cooperation: Evidence from a TV Game Show." *Journal of Economic Behavior & Organization* 73 (3) (March): 396–405. doi:10.1016/j.jebo.2010.01.001. http://linkinghub.elsevier.com/retrieve/pii/S0167268110000041.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–142. doi:10.1006/game.1995.1027. http://linkinghub.elsevier.com/retrieve/pii/S0899825685710275.

Bernhard, H, Ernst Fehr, and Urs Fischbacher. 2006. "Group Affiliation and Altruistic Norm Enforcement." *The American Economic Review* 96 (2): 217–221. http://www.jstor.org/stable/10.2307/30034645.

Bigoni, M, S Bortolotti, and Marco Casari. 2012. "Trustworthy by Convention." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2057607.

Brosig, Jeannette. 2002. "Identifying Cooperative Behavior: Some Experimental Results in a Prisoner's Dilemma Game." *Journal of Economic Behavior & Organization* 47 (3) (March): 275–290. doi:10.1016/S0167-2681(01)00211-6. http://linkinghub.elsevier.com/retrieve/pii/S0167268101002116.

Bucciol, Alessandro, and Marco Piovesan. 2011. "Luck or Cheating? A Field Experiment on Honesty with Children." *Journal of Economic Psychology* 32 (1): 73–78. doi:10.1016/j.joep.2010.12.001. http://linkinghub.elsevier.com/retrieve/pii/S0167487010001352.

Cassar, Alessandra, and Mary Rigdon. 2011. "Trust and Trustworthiness in Networked Exchange." *Games and Economic Behavior* 71 (2) (March): 282–303. doi:10.1016/j.geb.2010.04.003. http://linkinghub.elsevier.com/retrieve/pii/S0899825610000643.

Charness, G., and M. Dufwenberg. 2006. "Promises and Partnership." *Econometrica* 74 (6): 1579–1601. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2006.00719.x/abstract.

Charness, Gary, and Martin Dufwenberg. 2010. “Bare Promises: An Experiment.” *Economics Letters* 107 (2) (May): 281–283. doi:10.1016/j.econlet.2010.02.009. http://linkinghub.elsevier.com/retrieve/pii/S0165176510000728.

Chen, Yan, and Sherry Xin Li. 2009. “Group Identity and Social Preferences.” *American Economic Review* 99 (1): 431–457. doi:10.1257/aer.99.1.431. http://pubs.aeaweb.org/doi/abs/10.1257/aer.99.1.431.

Chichester, Aldert Vrij. 2008. “Detecting Lies and Deceit: Pitfalls and Opportunities (second Edition).” *Analysis*: 344–347.

Cooper, Russell, Douglas V DeJong, Robert Forsythe, and Thomas W Ross. 1989. “Communication in the Battle of the Sexes Game: Some Experimental Results.” *The RAND Journal of Economics* 20 (4): 568–587. doi:10.2307/2555734. http://www.jstor.org/stable/2555734.

———. 1992. “Communication in Coordination Games.” *Quarterly Journal of Economics* 107 (2): 739–771. doi:10.2307/2118488. http://qje.oxfordjournals.org/content/107/2/739.short.

Crawford, Vincent. 1998. “A Survey of Experiments on Communication via Cheap Talk.” *Journal of Economic Theory* 298: 286–298. http://www.sciencedirect.com/science/article/pii/S0022053197923590.

Croson, Rachel, Terry Boles, and J.Keith Murnighan. 2003. “Cheap Talk in Bargaining Experiments: Lying and Threats in Ultimatum Games.” *Journal of Economic Behavior & Organization* 51 (2) (June): 143–159. doi:10.1016/S0167-2681(02)00092-6. http://linkinghub.elsevier.com/retrieve/pii/S0167268102000926.

Darai, Donja, and Silvia Grätz. 2010. “Determinants of Successful Cooperation in a Face-to-Face Social Dilemma.”

Dawes, RM, J McTavish, and Harriet Shaklee. 1977. “Behavior, Communication, and Assuptions About Other People’s Behavior in A Commons Dilemma Situation.” *Journal of Personality and Social Psychology* 35 (1): 1–11. http://psycnet.apa.org/journals/psp/35/1/1/.

Dreber, Anna, and Magnus Johannesson. 2008. “Gender Differences in Deception.” *Economics Letters* 99 (1) (April): 197–199. doi:10.1016/j.econlet.2007.06.027. http://linkinghub.elsevier.com/retrieve/pii/S0165176507002534.

Ekman, P, M Frank, and M O’Sullivan. 1999. “A Few Can Catch A Liar.” *Psychological Science* 10 (3): 263–266. http://pss.sagepub.com/content/10/3/263.short.

Ekman, P, and M O'Sullivan. 1991. "Who Can Catch a Liar?" *The American Psychologist* 46 (9) (September): 913–20. http://www.ncbi.nlm.nih.gov/pubmed/1958011.

Ekman, Paul. 2009a. "Lie Catching and Microexperssions." In *The Philosophy of Deception*, ed. Clancy W Martin, 118–138. New York: Oxford University Press.

———. 2009b. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, Third Edition. Clues*. W. W. Norton & Company. http://www.amazon.com/dp/0393337456.

Ekman, Paul, and Wallace V. Friesen. 1974. "Detecting Deception From The Body Or Face." *Journal of Personality and Social …* 29 (3): 288–298. http://psycnet.apa.org/journals/psp/29/3/288/.

Ellingsen, Tore, and Magnus Johannesson. 2004. "PROMISES , THREATS AND FAIRNESS *." *The Economic Journal* 114: 397–420.

Ellingsen, Tore, Magnus Johannesson, Sigve Tjøtta, and Gaute Torsvik. 2010. "Testing Guilt Aversion." *Games and Economic Behavior* 68 (1) (January): 95–107. doi:10.1016/j.geb.2009.04.021. http://linkinghub.elsevier.com/retrieve/pii/S089982560900092X.

Erat, S., and U. Gneezy. 2011. "White Lies." *Management Science* (November 4): 1–11. doi:10.1287/mnsc.1110.1449. http://mansci.journal.informs.org/cgi/doi/10.1287/mnsc.1110.1449.

Farrell, Joseph, and Matthew Rabin. 1996. "Cheap Talk." *The Journal of Economic Perspectives* 10 (3): 103–118. http://www.jstor.org/stable/10.2307/2138522.

Fehr, Ernst, and Armin Falk. 2002. "Psychological Foundations of Incentives." *European Economic Review* 46 (4-5) (May): 687–724. doi:10.1016/S0014-2921(01)00208-2. http://linkinghub.elsevier.com/retrieve/pii/S0014292101002082.

Fischbacher, Urs, and F Heusi. 2008. "Lies in Disguise: An Experimental Study on Cheating." http://kops.ub.uni-konstanz.de/bitstream/handle/urn:nbn:de:bsz:352-opus-74612/62_No40_08_11_TWI_RPS_Fischbacher_Heusi.pdf?sequence=1.

Frank, RH, T Gilovich, and Dennis T. Regan. 1993. "The Evolution of One-Shot Cooperation: An Experiment." *Ethology and Sociobiology* 14 (4): 247–256. http://www.sciencedirect.com/science/article/pii/016230959390020I.

Gneezy, U, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics* 115 (3): 791–810. http://qje.oxfordjournals.org/content/115/3/791.short.

Gneezy, U. 2005. “Deception: The Role of Consequences.” *The American Economic Review* 95 (1): 384–394. http://www.jstor.org/stable/10.2307/4132685.

Gneezy, Uri, and Aldo Rustichini. 2000. “A Fine Is a Price.” *The Journal of Legal Studies* 29 (1): 1–17.

Hall, JK. 1995. “(Re)creating Our Worlds with Words: A Sociohistorical Perspective of Face-to-face Interaction.” *Applied Linguistics* 16 (2): 206–232. http://applij.oxfordjournals.org/content/16/2/206.short.

Hao, Li, and Daniel Houser. Working Paper. 2013. “Perception and Cheating: An Experimental Analysis.”

Houser, Daniel, Stefan Vetter, and Joachim Winter. 2012. “Fairness and Cheating.” *European Economic Review* 56 (8): 1645–1655. http://ideas.repec.org/p/trf/wpaper/335.html.

Houser, Daniel, and Erte Xiao. 2010. “Classification of Natural Language Messages Using a Coordination Game.” *Experimental Economics* 14 (1) (September 15): 1–14. doi:10.1007/s10683-010-9254-4. http://www.springerlink.com/index/10.1007/s10683-010-9254-4.

Houser, Daniel, Erte Xiao, Kevin McCabe, and Vernon Smith. 2008. “When Punishment Fails: Research on Sanctions, Intentions and Non-cooperation.” *Games and Economic Behavior* 62 (2) (March): 509–532. doi:10.1016/j.geb.2007.05.001. http://linkinghub.elsevier.com/retrieve/pii/S0899825607000802.

Isaac, RM, and James M. Walker. 1988. “Communication and Free-riding Behavior: The Voluntary Contribution Mechanism.” *Economic Inquiry* 26 (4): 585–608. http://onlinelibrary.wiley.com/doi/10.1111/j.1465-7295.1988.tb01519.x/abstract.

Kartik, Navin. 2009. “Strategic Communication with Lying Costs.” *Review of Economic Studies* 76 (4) (October): 1359–1395. doi:10.1111/j.1467-937X.2009.00559.x. http://restud.oxfordjournals.org/lookup/doi/10.1111/j.1467-937X.2009.00559.x.

Kimbrough, E, V Smith, and Bart J Wilson. 2006. “Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-distance Trade.” *American Economic Review* 98 (3) (May): 1009–1039. doi:10.1257/aer.98.3.1009. http://pubs.aeaweb.org/doi/abs/10.1257/aer.98.3.1009.

Lacetera, Nicola, and Mario Macis. 2010. “Do All Material Incentives for Pro-social Activities Backfire? The Response to Cash and Non-cash Incentives for Blood Donations.” *Journal of Economic Psychology* 31 (4) (August): 738–748. doi:10.1016/j.joep.2010.05.007. http://linkinghub.elsevier.com/retrieve/pii/S0167487010000735.

Lawson, H. M., and Kira Leck. 2006. “Dynamics of Internet Dating.” *Social Science Computer Review* 24 (2) (May 1): 189–208. doi:10.1177/0894439305283402. http://ssc.sagepub.com/cgi/doi/10.1177/0894439305283402.

Levitt, Steven D. 2006. “White-Collar Crime Writ Small: A Case Study of Bagels, Donuts, and the Honor System.” *American Economic Review* 96 (2): 290–294. doi:10.1257/000282806777212161. http://apps.isiknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=U2ppboH6a8gJHJimeH@&page=1&doc=1.

Li, Jian, Erte Xiao, Daniel Houser, and P Read Montague. 2009. “Neural Responses to Sanction Threats in Two-party Economic Exchange.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (39) (September 29): 16835–40. doi:10.1073/pnas.0908855106. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2744630&tool=pmcentrez&rendertype=abstract.

Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. 2009. “The Aversion to Lying.” *Journal of Economic Behavior & Organization* 70 (1-2) (May): 81–92. doi:10.1016/j.jebo.2009.02.010. http://linkinghub.elsevier.com/retrieve/pii/S0167268109000523.

Mazar, Nina, O. Amir, and D. Ariely. 2008. “The Dishonesty of Honest People: A Theory of Self-concept Maintenance.” *Journal of Marketing Research* 45 (6): 633–644. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=979648.

Mazar, Nina, and Dan Ariely. 2006. “Dishonesty in Everyday Life and Its Policy Implications.” *Journal of Public Policy & Marketing* 25 (1): 117–126.

Mellstrom, Carl, and Magnus Johannesson. 2008. “Crowding Out in Blood Donation: Was Titmuss Right?” *Journal of the European Economic Association* 6 (4): 845–863. http://onlinelibrary.wiley.com/doi/10.1162/JEEA.2008.6.4.845/abstract.

Miettinen, T., and S. Suetens. 2008. “Communication and Guilt in a Prisoner’s Dilemma.” *Journal of Conflict Resolution* 52 (6) (July 28): 945–960. doi:10.1177/0022002708322325. http://jcr.sagepub.com/cgi/doi/10.1177/0022002708322325.

Ockenfels, Axel, and Reinhard Selten. 2000. “An Experiment on the Hypothesis of Involuntary Truth-Signalling in Bargaining.” *Games and Economic Behavior* 33 (1) (October): 90–116. doi:10.1006/game.1999.0762. http://linkinghub.elsevier.com/retrieve/pii/S0899825699907625.

O’Sullivan, M, and Paul Ekman. 2005. “The Wizards of Deception Detection.” In *The Detection of Deception in Forensic Contexts*, ed. P Granhag and L Strömwall, 269–286. Cambridge University Press.

Pan, Xiaofei, and Daniel Houser. Working Paper. 2013. “Cooperation During Cultural Group Formation Promotes Trust Towards Member of Out-groups.”

Papini, D R, F F Farmer, S M Clark, J C Micka, and J K Barnett. 1990. “Early Adolescent Age and Gender Differences in Patterns of Emotional Self-disclosure to Parents and Friends.” *Adolescence* 25 (100): 959–976. http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com/login.aspx?direct=true&db=sph&AN=9603010054&site=ehost-live.

Pruckner, Gerald J, and Rupert Sausgruber. 2008. “Honesty on the Streets - A Natural Field Experiment on Newspaper Purchasing.” *October*. SSRN. doi:10.2139/ssrn.1277208. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1277208.

Rietz, T, R Sheremeta, T Shields, and Vernon Smith. 2011. “Transparency, Efficiency and the Distribution of Economic Welfare in Pass-through Investment Trust Games.” *Exchange Organizational Behavior Teaching Journal*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1830622.

Rosaz, Julie, and Marie Claire Villeval. 2011. “Lies and Biased Evaluation: A Real-Effort Experiment.” *IZA Discussion Paper Series*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1906187.

Rubin, Victoria L, and Elizabeth D Liddy. 2006. “Assessing Credibility of Weblogs.” In *AAAI Symposium on Computational Approaches to Analysing Weblogs AAAICAAW*, 187–190. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Assessing+Credibility+of+Weblogs#0.

Sheremeta, Roman M, and Jingjing Zhang. Working Paper. 2010. “Multi-Level Trust Game with ‘Insider’ Communication.”

Sutter, Matthias. 2009. “Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams*.” *The Economic Journal* 119 (534): 47–60. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2008.02205.x/full.

Vanberg, Christoph. 2008. “Why Do People Keep Their Promises? An Experimental Test of Two Explanations.” *Econometrica* 76 (6): 1467–1480. doi:10.3982/ECTA7673. http://doi.wiley.com/10.3982/ECTA7673.

Vrij, Aldert, Lucy Akehurst, Stavroula Soukara, and Ray Bull. 2004. “Detecting Deceit Via Analyses of Verbal and Nonverbal Behavior in Children and Adults.” *Human Communication Research* 30 (1) (January): 8–41. doi:10.1111/j.1468-2958.2004.tb00723.x. http://www.blackwell-synergy.com/doi/abs/10.1111/j.1468-2958.2004.tb00723.x.

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. doi:10.1257/aer.100.3.984. http://pubs.aeaweb.org/doi/abs/10.1257/aer.100.3.984.

Wood, W, CA Kallgren, and Rebecca Mueller Preisler. 1985. "Access to Attitude-Relevant Information in Memory as a Determinant of Persuasion : The Role of Message Attributes." *Journal of Experimental Social Psychology* 21: 73–85. http://www.sciencedirect.com/science/article/pii/0022103185900071.

Zhou, Lina, J Burgoon, and D Twitchell. 2003. "A Longitudinal Analysis of Language Behavior of Deception in E-mail." In *Intelligence and Security Informatics*, 2665:102–110. http://www.springerlink.com/index/G27D8LLNDMJH6C97.pdf.

Zhou, Lina, DP Twitchell, and T Qin. 2002. "An Exploratory Study into Deception Detection in Text-based Computer-Mediated Communication." *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1173793.

Zuckerman, Miron, Bella M DePaulo, and Robert Rosenthal. 1981. "Verbal and Nonverbal Communication of Deception." Ed. L Berkowitz. *Advances in Experimental Social Psychology* 14: 1–59. doi:10.1016/S0065-2601(08)60369-X. http://www.sciencedirect.com/science/article/pii/S006526010860369X.