UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Essays on Deception and Lying Aversion

Permalink
https://escholarship.org/uc/item/0n1899r5

Author
Gawn, Glynis

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Essays on Deception and Lying Aversion

By

Glynis Margaret E. Gawn

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor Jeffrey M. Perloff, Chair
Professor Alain de Janvry
Professor Stefano DellaVigna

Spring 2015

Essays on Deception and Lying Aversion

Copyright 2015

by

Glynis Margaret E. Gawn

Abstract

Essays on Deception and Lying Aversion

By

Glynis Margaret E. Gawn

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Jeffrey M. Perloff, Chair

This dissertation consists of three experimental essays on deception and lying aversion. Chapter 2, "Do Lies Erode Trust?" studies the interaction between honesty and trust and trustworthiness. Specifically, the chapter investigates the effect of being lied to or told the truth in a Gneezy (2005) deception game on behavior in a subsequent trust game with different players. Treatment effects are decomposed between the impacts of being "burned" by a low payoff in the deception game, mood change, and the specific experience of a lie. The specific experience of being lied to significantly erodes trust, trustworthiness, and the use of communication to promote trust. However, the experience effect on trustworthiness occurs only for subjects who are burned.

Chapter 3, "Pure Lying Aversion", studies several factors affecting the propensity to tell the truth when no one would be directly negatively impacted by the lie. Utilizing a simple experiment, the effect of the strength of the message one is using to convey information is examined, while the economic incentive to lie is also varied. The effect of being lied to in a prior interaction on one's subsequent truthfulness is also studied in a separate set of experiments. The strength of the message has a strong effect on truthfulness regardless of the incentive to lie, while the effect of the size of the economic gain from lying has a non-monotonic effect on truthfulness. Additionally, the effect of knowledge about whether one has been lied to before interacts with the payoff outcome received in the prior interaction to reduce truthfulness in some cases and increase it in others.

Chapter 4, "Lying Through Others", considers the question of how agency relationships, ubiquitous in economic interactions, affect an individual's willingness to lie for monetary advantage? Does individual lying aversion tend to decline if the lie (or truth) is sent through an agent, rather than sent directly by the individual? In three experiments that control for the effects of delegation on preferences over payoffs and probabilities of actions, it is found that delegation reduces – but does not eliminate – lying aversion.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am very grateful to the Agricultural and Resource Economics department for allowing me to return and complete my PhD after a long break. I am especially indebted to Jeff Perloff for helping me with the process of re-entering the program and gaining approvals for various aspects of my research, as well as chairing my thesis committee. I am also extremely grateful to Alain de Janvry for chairing my orals committee, and to the other members of my committees, Stefano DellaVigna, Elisabeth Sadoulet and Dave Sunding, for all their help. I am also grateful to all the other professors with whom I had the opportunity to study when I began the PhD program, especially David Zilberman.

I am extremely grateful also to Robert Innes, my husband and co-author, without whom I would not have been able to undertake the research on which this dissertation is based, and for his unwavering faith in my ability to complete the PhD.

Chapter 1
Introduction

While honesty has long been upheld by both religious and civic institutions as important both for its own sake and for the smooth functioning of society (Bok 1978), neoclassical economic theory has emphasized the maximization of material well-being as the primary focus of individuals and firms. Models, such as those examining principal-agent relationships, investigate how to structure incentives under the assumption that agents will lie if doing so increases their material well-being. However, much evidence has now accumulated that many people are not in fact willing to lie to increase their material well-being, under at least some circumstances. Economists have begun to explore the implications of a preference for honesty in some theoretical models by including a cost of lying in the utility function. In a seminal paper on the role of communication, Crawford and Sobel (1982) study communication in the context of a situation where a Sender, who has private information about his type, has to communicate with a Receiver who then has to take an action which determines payoffs for both. Under the assumption of costless lying, they show that communication is "cheap-talk", that is, it has no consequence for the outcome. Kartik (2009)[1] builds on the model of Crawford and Sobel (1982) but introduces a cost of lying. He shows that when the cost of lying is small, Senders would continue to lie, and Receivers expect Senders to lie. He gives as an example of this, analysts at brokerage firms where there is an underwriting relationship with the issuers of a stock, providing inflated recommendations on that stock. However, as lying costs increase, Kartik shows that there is more truthfulness. In a different context, that of coordination on a superior equilibrium outcome for both players, Demichelis and Weibull (2008) show how even a very small preference for honesty or cost of lying makes the superior outcome the only feasible equilibrium in a repeated game, rather than unattainable as it would be with costless lying as found by Aumann (1990).

The anonymity of the experimental setting is particularly valuable for studying lying behavior, since many people might prefer to lie in some circumstances, but would not wish to admit to violating what is widely perceived as a social norm of honesty. The experimental literature on lying has grown rapidly since the publication of Gneezy's (2005) experimental paper based on a simplified version of the Sender-Receiver model of Crawford and Sobel (1982). Gneezy's experiments studied how truthfulness varies with the gain to the Sender and the harm to the Receiver, and he found that while people tend to lie more if they can benefit from it, they lie less if the Receiver is harmed more for a given gain to themselves. Subsequent studies based on Gneezy's Sender-Receiver game have found that lying is affected by a variety of factors. In addition to being sensitive to its direct monetary consequences for those on both sides of the interaction (Gneezy, 2005; Gibson et al., 2013; Freeman and Gelber, 2010), honesty has been found to be affected by strategic considerations (Sutter, 2009), a norm of honesty (Pruckner and Sausgruber, 2013), social

[1] Kartik(2009) reviews many other contributions to the theoretical literature.

cues on how often others lie (Innes and Mitra, 2013), gender (Dreber and Johanneson, 2008), the extent of the lie (Lundquist et al., 2009; Fischbacher and Heusi, 2013), and cooperation in prior play (Ellingsen et al., 2009) but not cooperative (vs. competitive) priming (Rode, 2010). Chapter 3 extends the study of determinants of lying or truthful behavior by examining the role of communication language in lying, in a different context than previous studies focusing on language (Lundquist et al., 2009; Cappelen et al. 2013) by varying the strength of the statement that is being made when a lie is being told, while simultaneously varying the monetary reward to lying. The strength of the statement is found to have a very significant effect on the rate of truth-telling that is invariant to the economic cost of the lie, while the rate of truth-telling is at the same time also directly affected by the economic cost of the lie.

Chapter 4 extends the examination of the circumstances under which people are averse to lies by examining the question of whether lying aversion is lower when one is not telling a lie directly, but rather through delegation to an agent. Agency relationships are prevalent in economic interactions. Employees act for firms. Subcontractors and outsourced suppliers produce for contracting companies. Intermediaries act on behalf of clients in employment and professional services. In many of these settings, agents make decisions with ethical overtones that affect economic costs and benefits to their principals. In a series of experiments that disentangle effects on willingness to make a decision, and preferences over allocations, the effect of agency on lying aversion is examined, and it is found that lying aversion is indeed lower when one is acting through an agent, but not entirely eliminated. These results hold even when the lie being told benefits both the liar and the recipient of the lie.

While much recent work has considered various factors affecting the propensity to lie, less attention has been focused in the deception and lying aversion literature on the consequence of being lied to for one's own subsequent behavior. There are potentially two kinds of behavior of interest – behavior toward the same person who has perpetrated the lie against us, and behavior in subsequent interactions with different people. In models of social preferences, direct reciprocity is considered an important driver of behavior; that is, people often exhibit a tendency to behave well toward those who have treated them well, and negatively or punitively toward those who have wronged them. This literature has mostly focused on allocations, such as behavior of proposers in ultimatum games or dictators in dictator games, rather than lying behavior, but reaction to a lie could similarly result in reciprocal behavior toward the person who lied to us. Behavior of one person toward another that is outside the immediate dyad but in response to the initial action, is termed indirect reciprocity. There are two main categories of this form of response to an initial action. When a third party, C, responds in kind toward A as a result of the action (positive or negative) that A directed toward B, this is termed indirect or social indirect reciprocity, while the behavior of B in response to the effect of A's action on him/her in subsequent interaction with C is termed generalized or generalized indirect reciprocity (Herne et al. 2013, Stanca, 2010).

This latter behavior is the focus of Chapter 2, which uses a Gneezy (2005) type Sender-Receiver game to study the effect of knowledge about whether a participant has been lied to or told the truth on subsequent behavior in a trust game with a different person than the one who either lied to or told the truth to the participant in the initial interaction. The importance of trust as the glue of a well-functioning economy makes it a natural arena in which to study the impact of knowledge of dishonesty on behavior. Experimental evidence is found for a

detrimental effect of being lied to on subsequent trusting behavior, and also on trustworthy behavior, but in the latter case only when the lie resulted in a poor payoff from the initial interaction.

The second part of Chapter 3 examines whether the effect of being lied to also encourages further lying behavior by the recipient of the original lie. This is examined in a context where no-ones' material outcome is directly affected by this new lie. Thus, unwillingness to lie in this situation is termed "pure lying aversion". In this setting, evidence is found that, when being on the receiving end of a lie has resulted in a poor payoff for the participant, he/she is more likely to lie. In contrast, and somewhat surprisingly, when the outcome of the lie was a good payoff, due to the participant correctly disbelieving the liar, being lied to actually increased truthfulness, relative to those who received the same payoff as a result of being told the truth, when controlling for underlying differences in behavior of those who believed and disbelieved the received messages.

The experimental research into lying has expanded rapidly in the last ten years, but in comparison to research into norms such as altruism, cooperation, fairness and trust, there is still much more heterogeneity in experimental design and measures of honesty or dishonesty (Rosenbaum et al. 2014). This dissertation utilizes several experiments to examine some of the determinants of lying and lying aversion, as well as the impact of being lied to on the subsequent behavior of the recipient of the lie.

Chapter 2
Do Lies Erode Trust?[2]

"Those who have been lied to … are resentful, disappointed, and suspicious. They feel wronged; they are wary of new overtures… Those who share (the perspective of the deceived) … are all too aware of the impact of discovered and suspected lies on trust and social cooperation." Sissela Bok (*Lying*, 1978).

2.1 Introduction

Lies are a common and frequent phenomenon in everyday life. DePaulo et al. (1996) found that college students and members of a U.S. community lied in 20 to 31 percent of social interactions recorded in daily diaries, with college students telling an average of two lies per day, and community members one lie per day. In a recent survey of over 23,000 high school students, 76 percent self-reported that they had lied about something significant in the past year, while 38 percent indicated that they sometimes lie to save money (Josephson Institute of Ethics, 2012).

Such statistics would seem to be consistent with the standard economic model of self-interested behavior that predicts lying whenever an individual can materially benefit from this behavior. However, a recent economics literature provides compelling evidence that many individuals are averse to lies; they are honest despite monetary incentives to be dishonest (see, for example, Gneezy, 2005; Gibson et al., 2013; Fischbacher and Heusi, 2013). While these results are a surprise from the standpoint of the baseline economic model, what if lies have serious economic consequences, beyond direct and immediate costs to the recipient of the lie? Perhaps lying aversion is a symptom of these consequences.

Two perspectives suggest that lies are likely to cause economic harm. On one hand, philosophers have elucidated the deleterious effects of lies, particularly on those who have been lied to, in works dating back as far as Aristotle (Bok, 1978). Lies, it is argued, erode trust and thereby deter social cooperation. On the other hand, recent economics research identifies significant benefits of generalized trust in promoting economic growth and progress.[3] If both perspectives are right, then values, norms and institutions that deter lies –

[2] This essay is joint work with Robert Innes, UC Merced.

[3] A compelling literature identifies close links between survey indicators of trust and, respectively, economic growth (e.g., Knack and Keefer, 1997; Zak and Knack, 2001; Guiso et al., 2004), international trade (Guiso et al., 2009), development of financial institutions (Guiso et al., 2008), and other indicators of economic success (LaPorta, et al., 1997; Bloom et al., 2009). A large experimental literature on trust games is arguably motivated by these links (see Johnson and Mislin, 2011, for a recent survey). Indeed, a growing body of work identifies the close relationship between behavior in experimental trust games and survey evidence on trust. See, for example, Glaeser et al. (2000), Lazzerini et al. (2004), Fehr et al. (2003), and Bellemare and Kroger (2007). This literature studies, among other things, the correlation between responses to survey questions on trust (such as answers to World Values Survey) and experimental indicators of trust and trustworthiness. Some of this

many of which we see in practice (Fischbacher and Heusi, 2013) – may deliver economic benefits by promoting trust.

In this paper, we study the effects of lies on those at the receiving end in order to test the general proposition that lies erode trust. We measure how being lied to, in a first-round Gneezy (2005) deception game, alters behavior on both sides of a second-round trust relationship with a different person. Second round outcomes include whether to trust a partner and whether to reciprocate trust with trustworthiness in a simple interaction patterned after the original game of Berg et al. (1995).[4] Overall, we find that being lied to in a prior interaction erodes both trust and trustworthiness.

This conclusion is both general – in the sense that lies erode trust overall – and specific in the sense that lies erode trust even when controlling for a variety of correlated effects. What is it about the experience of a lie that might deter trust? Is it because lies disappoint their recipients? Is it because lies "burn" their recipients by reducing payoffs they enjoy from the interaction? Is it because lies signal something about norms of honesty, so that a recipient of a lie reasonably infers that, generally speaking, others are less honest? Or is there something more fundamental about a lie that affects trust, separate from immediate disappointment, harm, and inferences about social behavior? For example, Bok (1978) stresses the fundamental nature of truthful communication as a cornerstone of human interaction; shaking this foundation with the experience of a lie, she suggests, can limit free will, jeopardize accumulation of knowledge, and, in the extreme, lead to the collapse of social institutions. These arguments suggest that lies may have intrinsic consequences.

Our experimental design allows us to isolate the intrinsic effect of lies by controlling for mood, "burns," and overall propensities for honesty. Stripped of these correlated effects, the proposition that lies erode trust becomes less clear as an implication of existing literature. Prior work documents that mood can affect behavior in a trust interaction (e.g., see Capra, 2004; Kirchsteiger et al., 2006). "Burns" can alter behavior due to effects on mood, preferences, and perceptions of procedural justice (Brandts and Charness, 2003; Sanchez-Pages and Vorsatz, 2007, 2009; Peeters et al. 2013). These general (mood/burned) effects are symptoms of the social interaction that we examine, but they are not unique to it; many other experiences may also generate these symptoms, including, for example, being at the receiving end of a low payoff from a dictator (e.g., Ben-Ner et al., 2004; Herne et al., 2013) or a first-round defector (Ellingsen et al., 2009). Similarly, prior work demonstrates that changes in perceived social norms can spill over from one context to another (Keizer et al., 2008; Houser et al., 2012). In our setting, changes in perceived propensities for honesty could reduce trust in a subsequent interaction. We control for this channel of effect by informing all players in our experiment about the overall proportion of lies, so that the experience of being lied to, or told the truth, is an individual experience and not a reflection of norms or general propensities for honesty. In sum, the existing literature might be interpreted as predicting effects of our treatments on trust due to potential impacts on mood, "burns," and norms; however, to our knowledge, it does not speak to the intrinsic effects of lies that we identify.

---

work indicates correlation between survey measures and trustworthiness, but not trust (Glaeser et al., 2000, for example); others document correlation between survey responses and trust (Fehr et al., 2003, for example). Sapienza et al. (2013) reconcile conflicting evidence in their recent study.

[4] See also the lost wallet game of Dufwenberg and Gneezy (2000).

These intrinsic effects are potentially relevant to both research and policy. If effects of lies are driven by resulting "burns" or bad moods, then any traits that reduce such effects, including lying aversion, but also many others, for example, resiliency, weaker preferences over relatively poorer payoffs, and stronger dictator preferences for generosity, will promote advantageous trust. However, the intrinsic experience effects of lies that we find here can only be mitigated by traits, policies and norms that deter lies – including, for example, ingrained lying aversion, societal values that promote veracity, and a culture of honesty in organizations.

2.2. Relationship to Prior Literature

Consistent with the normative perspective that lies are damaging to society, recent literature documents a broad departure from the benchmark model of self-interested behavior, namely, an unwillingness of many individuals to lie despite monetary benefits from doing so. Mostly drawing on Gneezy's (2005) initial deception game, recent studies identify propensities for honesty versus dishonesty, including how lying aversion varies across individuals (Gibson, Tanner and Wagner, 2013; Hurkens and Kartik, 2009) and how a variety of alternate circumstances affect lying behavior. For example, scholars have found that honesty is sensitive to its direct monetary consequences for those on both sides of the interaction (Gneezy, 2005; Gibson et al., 2013; Freeman and Gelber, 2010), strategic considerations (Sutter, 2009), a norm of honesty (Pruckner and Sausgruber, 2013), social cues on how often others lie (Innes and Mitra, 2013), gender (Dreber and Johanneson, 2008), the extent of the lie (Lundquist et al., 2009; Fischbacher and Heusi, 2013), and cooperation in prior play (Ellingsen et al., 2009) but not cooperative (vs. competitive) priming (Rode, 2010).[5]

In the present paper, we focus instead on the *consequences* of a lie, beyond its direct and immediate effect on payoffs to the liar and recipient of the lie, for trust. While there is a rich literature studying individual drivers of trust – including beliefs about behavior (e.g., Sapienza et al., 2013; Costa-Gomez et al., 2010), mood (e.g., Capra, 2004; Kirchsteiger et al., 2006), and a variety of preference attributes[6] – to our knowledge this is the first study investigating the impact of receiving a lie on trust outcomes.

A number of papers are closely related to this inquiry. A study in the psychology literature, Tyler et al. (2006), uses videotaped conversations to reveal lying behavior to the participants. The authors find that when participants witness more lying behavior, they like and believe their partner less and also increase their own use of deception in follow-up interactions with the same partner. Unlike our paper, however, Tyler et al. (2006) do not use economic incentives in their experiment; the subjects interacting in the follow-up round are

[5] See also Battigalli, Charness and Dufwenberg (2013) on the role of guilt, Shalvi et al. (2012) and Mazar et al. (2008) on the role of self-justifications and self-concept, and the recent survey by Rosenbaum et al. (2014).

[6] Relevant attributes include altruism (Cox, et al., 2008; Ashraf, et al., 2006), reciprocity (Charness and Rabin, 2002), inequity aversion (Fehr and Schmidt, 1999), risk aversion (Houser et al., 2010), values of social welfare (Charness and Rabin, 2002), benefits of a "warm glow" (Andreoni, 1990), and guilt aversion (Charness and Dufwenberg, 2006). (This is a small subset of the literature, and we apologize to authors of many key papers omitted here.)

the same as those who lied to them (or not) in the earlier round; and specific effects on trust are not the focus.

Arguably most relevant are key papers by Gneezy et al. (2013) and Brandts and Charness (2003), Sánchez-Pagés and Vorsatz (2007, 2009) and Peeters et al. (2013) studying effects on the deceived in the context of economic experiments. Gneezy et al. (2013) find that Receivers in a multi-round deception game are less likely to follow the recommendation of their Senders if they have been negatively affected by a lie in the previous round. While this result can be interpreted as a negative effect on trust, it may reflect learning from prior experience in the same (deception) game; and distinguishing between experience and "burned" effects of being lied to is not possible. Brandts and Charness (2003), in a simultaneous move game with prior communication by the Sender, find that Receivers are most likely to punish if a poor outcome results from a deceitful message on the part of the Sender, that is, when the action did not match that indicated in the message. Similarly, Sánchez-Pagés and Vorsatz (2007, 2009) and Peeters et al. (2013) examine the extent to which Receivers in a deception game punish their respective Senders. They find that the Receivers punish primarily when they have been lied to and been burned as a result (because they followed the lie). The punishment in these papers appears to reflect concerns for procedural justice, since much less punishment is meted out when the same low payout is received through failing to believe a truthful statement.[7] These interesting results come closest to distinguishing between intrinsic experience and "burned" effects in the punishment behavior of Receivers. In contrast to the trust interaction that we study, however, where participants are not playing with those who lied to them, punishment in these papers is a manifestation of direct reciprocity.[8]

Keizer et al. (2008) conduct public field experiments in which violation of one social norm (the target) is more likely when another norm (the contextual norm) is violated. For example, subjects are more likely to litter when they see graffiti (despite a "no graffiti" sign). Our study is similar in the sense that we study cross-context behavioral spillovers, and find that experience in one domain (receiving a lie) affects behavior in another (trust). Our focus is very different in other respects. Whereas Keizer et al. (2008) study public behavior with the potential for social or governmental sanctions (and observation of contextual norm violations may prompt inferences that sanctions are less likely), the behavior that we study is private with no possibility for sanctions or building of social esteem. In Keizer et al. (2008), observation of a contextual norm violation (e.g., graffiti) can also affect subjects' inferences about what social norms prevail in the target context (e.g., litter). We control for inferences about norms by distinguishing specific experience (of a lie) from the overall propensity for honesty in our experiment, the latter of which is conveyed to all subjects.

[7] Sanchez-Pages and Vorsatz (2009) introduce a costly "silence" option for Senders in a Gneezy (2005)-type game, showing how the presence of a punishment option promotes silence, while Peeters et al. (2013) allow participants to select into either a sanctioning or non-sanctioning institution.

[8] It is also possible that Receivers who are burned by lies are different types of people (those who follow Sender recommendations) than Receivers who are not burned by lies (those who reject their Sender recommendations). For example, first round "accepters" may be more willing to punish than first round "rejecters." If so, punishment by burned recipients of lies (the accepters) may reflect primarily a "burned" effect versus an intrinsic experience effect of lies (for the accepters). In our data, we find evidence that the two different types of Receivers tend to behave differently in the trust game. Identifying the pure experience effect of a lie (our objective) requires a decomposition that controls for this source of heterogeneity.

Al-Ubaydli et al. (2013) study how market priming promotes trust, indicating a causal connection between markets and the trust that is also associated with economic progress. Several papers study how people react to their treatment as Receivers in a dictator game. Much of this literature identifies generalized/indirect reciprocity by studying behavior in second-round dictator games with different players (e.g., Ben-Ner et al., 2004; Herne et al., 2013). An exception is Houser et al. (2012), who find that individuals who feel they are unfairly treated in a first-round dictator game are more likely to cheat in a subsequent (unrelated) game, which the authors interpret as evidence of cross-context spillovers in social norms. Results from this work are suggestive of those we seek to identify in the sense that a prior experience at the receiving end of perceived moral or immoral behavior is found to affect subsequent behavior in a game also with moral overtones. However, the nature of the treatments ("unfair" in dictator games vs. "lied to" in our context) is very different,[9] as are outcomes (fairness or cheating vs. trust in our context); indeed, cheating outcomes are more akin to the deception treatments, the effects of which we study. In view of these contrasts, results from the prior literature cannot be readily translated to the issue we raise here.[10]

The indirect reciprocity literature also includes studies on trust. For example, Dufwenberg et al. (2001) compare trust games in which a Returner, when trusted by a Sender, alternately makes a decision on reciprocating by returning money to the same Sender (direct reciprocity) or to a different Sender (indirect reciprocity).[11] Being trusted in these contexts arguably also reflects a not-burned situation. The intrinsic experience effects that we identify in this paper potentially reflect a type of indirect reciprocity, similar to that identified by Dufwenberg et al. (2001), but not driven by a burned or not-burned effect.

2.3. The Experiment

Our design involves two subject interactions, in two games, between different players. First is the deception game, followed by the trust experiment.

2.3.1 The Deception Game

The deception game follows the Gneezy (2005) design. In this game, Senders from one classroom are randomly paired with Receivers from another classroom, one Receiver for each Sender. The Sender observes two possible payoff allocations between the two players. In our game, the payoff options are as follows:

Option C: \$6 to the Sender and \$3 to the Receiver.

Option D: \$4 to the Sender and \$6 to the Receiver.

[9] The "selfishness" exhibited in dictator games is sometimes heralded (by economists in particular) for promoting effort and innovation that are central to successful market economies; in other contexts, it is derided as an impediment to cooperative relationships. In contrast, dishonesty and corruption are consistently scorned by churches, community leaders, and even economists for impairing economic progress.

[10] List (2007) and Bardsley (2008) find that arguably modest framing differences in dictator games can have significant effects on behavior, let alone more profound variations in the structure of an experiment.

[11] See also Greiner and Levati (2005). Other work studies indirect reciprocity in a gift exchange game (Stanca, 2009). A related literature – but less relevant to our experiment – studies whether a subject who observes his or her matched player helping someone else, is more likely to be generous toward that player.

The Sender chooses one of two Messages to deliver to the Receiver, one *truthful* (Message D) and the other *untruthful* (Message C). The two possible Messages are:

Message C: "Option C will earn you (the Receiver) more money than Option D."

Message D: "Option D will earn you (the Receiver) more money than Option C."

Based only on the Message chosen by the Sender, the Receiver chooses one of the two Options, which in turn determines payoffs to the two players, Sender and Receiver. In the experiment, Option labels are varied between subjects (sometimes Option C is better for the Receiver and sometimes Option D). Receivers are never told the dollar amounts in the two options, but are told that one of the two is better for the Receiver and the other is better for the Sender.

Our focus is on the Receivers. After all Receiver decisions are made in the Deception game, and the decisions collected by the experimenter, Receivers are exposed to our Treatments.

2.3.2 The Treatments

Receivers are randomly assigned to three Treatment groups. The first group is the set of Control subjects who are exposed only to common information about the Deception game – that is, information that is given to all Receivers. The purpose of this information is to control for subject beliefs about behavior in the Deception experiment. The specific information given to all subjects, after their decisions in the Deception game (Experiment 1) are made, is as follows:

"In Experiment 1, roughly 5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED."

The statement reports the approximate percentage of truthful subjects from the Sender side of our experiment. The precise percentage of truthful Senders was 47.8 percent.

Each subject in the second and third Treatment groups is told whether his or her own matched Sender *lied* (Treatment LT) or *told the truth* (Treatment TT) in the Message that was sent. We are interested in the effects of this specific experience on subsequent decisions in a trust game. How does being lied to (or being told the truth) affect (i) one's willingness to trust, (ii) one's trustworthiness, and (iii) the effect of communication in promoting trust?

In studying these questions, we note that treated subjects are distinguished not only by the Treatment information (whether they were lied to, for example) but also by their decisions in the Deception game. Subjects who *accepted* their Sender recommendations in Experiment 1 are hurt by a lie (and helped by a truthful message); conversely, subjects who *rejected* their Sender recommendations are helped by a lie – in the sense that they earn a higher payoff in Experiment 1. These distinctions will be crucial in the analysis of our experimental outcomes as we seek to disentangle effects of (i) specific experience of being lied to (or told the truth), (ii) being burned in the Deception game, and (iii) inherent differences between "accepters" and "rejecters."

Random assignment to Treatments is ensured by random matching at the start of the experiment. Each questionnaire identifies a participant by the registration number, which in turn determines the Treatment group (with the correspondence known only by the experiment manager). While the assignment of registration numbers to Treatments is determined a priori, assignment of registration numbers to subjects is purely random.

Registration numbers contained a numerical identifier specific to each individual subject, followed by an alphabetical identifier associated with the treatment group (Z, V, and W, for example). Alphabetical identifiers were different in different classrooms. After turning in their deception game decisions, Receivers were given an information sheet. Control subjects (with Z identifiers, for example) collected their sheet at one "station" to which they were directed (so that their information only reflected overall propensities for honesty of Senders). LT and TT treatment subjects (with V and W identifiers, for example) were each directed to one of two other "stations," where they were given an information sheet containing both information on overall propensities for honesty AND information on whether their own Sender lied or told the truth. This is the only point at which any reference was made to the alphabetical identifier. On the information sheet, the LT and TT treatment subjects were told:

▪If your Registration number ends with an V, your Sender TOLD YOU THE TRUTH in Experiment 1 about the Option that earns you more money.

▪If your Registration number ends with a W, your Sender LIED TO YOU in Experiment 1 about the Option that earns you more money.

To verify understanding, we also asked each of the LT and TT treatment subjects to circle whether they were Told the Truth or Lied To.

2.3.3 The Trust Game

After receiving the Treatments, subjects participate in a second experiment. Here, each subject is again matched with another player in a different classroom. None of the participants in this game are Senders from the Deception experiment, and subjects are told that their matched player is a different person than their Sender from the first (Deception) experiment.

Subjects are either in the role of Sender or Returner and each player starts with $4. The Sender chooses between two alternatives:

KEEP. Keep the initial $4, implying that both players earn the $4 allocated to them.

SEND. Send his/her $4 to the Returner.

If the Sender chooses SEND, the $4 sent becomes $8, which combined with the Returner's initial $4, makes $12 available. In this case, the Returner chooses between:

OPTION A. Return $7 to the Sender, so that the Returner receives $5 and the Sender receives $7.

OPTION B. Return $2 to the Sender, pay a fee of $2 and keep the remainder, so the Returner receives $8 and the Sender receives $2.

In this game, a "SEND" decision by the Sender is an indication of *trust*, and a Returner choice of Option A indicates *trustworthiness*. Table 2.1 summarizes the payments.

Table 2.1: Trust Game Payments (in $)

		If Sender Chooses SEND		If Sender Chooses KEEP	
		Payment to Returner	Payment to Sender	Payment to Returner	Payment to Sender
Returner's Option Choice	A	$5	$7		
	B	$8	$2	$4	$4

Before the Returner decides which Option to choose, he or she can deliver a message to the Sender. The message is:

MESSAGE A: I am going to choose Option A.

Alternately the Returner can choose to send **NO MESSAGE**. Returners are told that a decision to send Message A does not preclude them from choosing Option B.

The Sender decisions are elicited using the strategy method: Subjects are asked to make a choice (KEEP or SEND) for each of the two possibilities, if he or she receives Message A or No Message. Payments are then determined (by Table 1) according to the decision made by the Sender for the actual Message that was sent (Message A or No Message) and, if the Sender chooses SEND, the Returner's choice of Option (Option A or Option B). Option labels are again varied between subjects (sometimes Option A is generous, as above, and sometimes stingy).

Figures 2.1A and 2.1B below summarize the sequence of decisions and the payoff options in the two games.

Figure 2.1A Sender-Receiver Game and Receiver Treatment

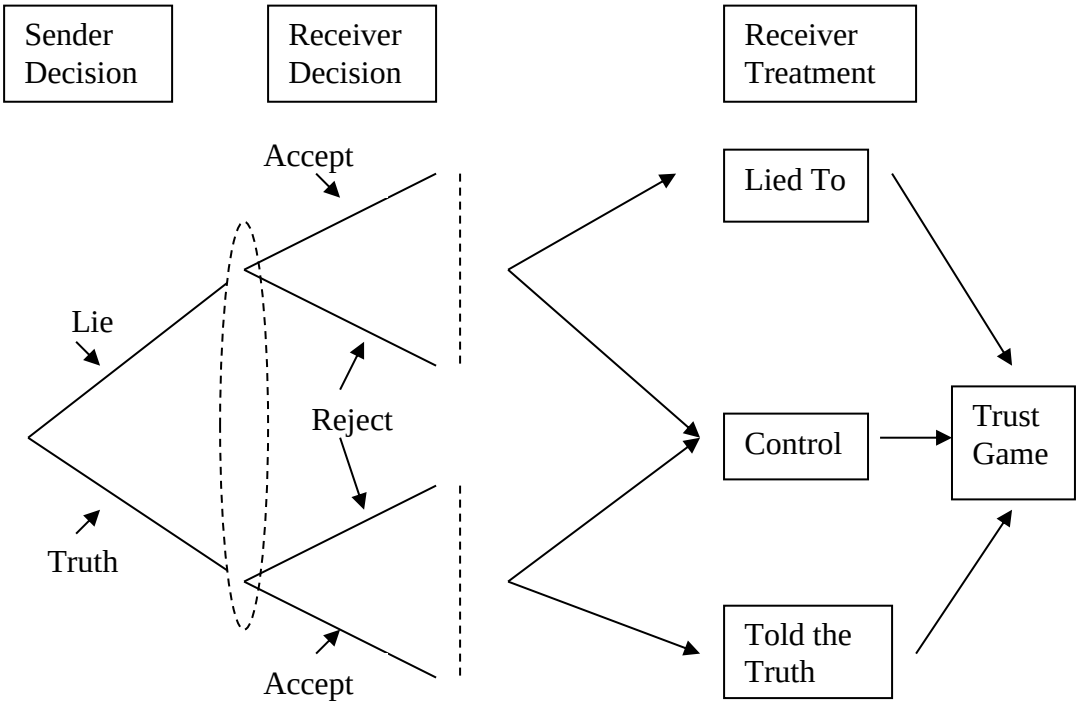
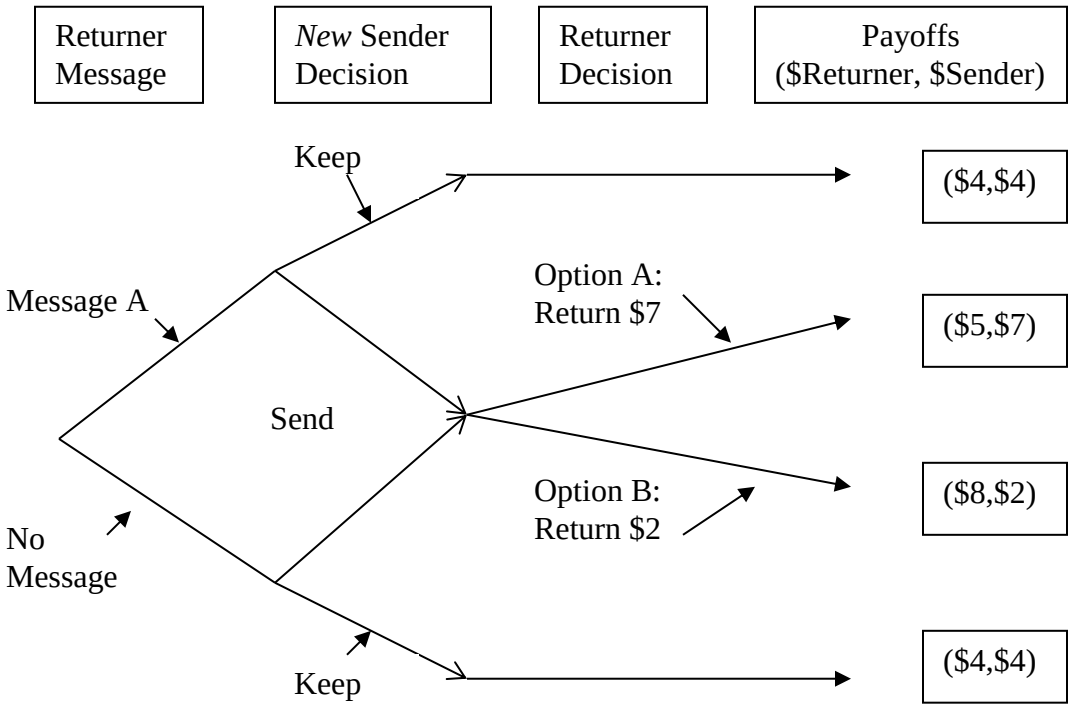


Figure 2.1B Trust Game



In the experiment, participants make decisions in both roles. Each matched pair is paid according to one player's decision as Sender and the other player's decision as Returner, with the allocation of roles determined by a coin flip after the experiment is completed. Subjects are told this procedure at the start of the trust experiment, with the corresponding instruction: "You should therefore make your decision in each situation (role) as if it is the one for which you will be paid." Because participants simultaneously and anonymously make choices in both roles (Sender and Returner), with payments determined according to one of the two roles, reputational motivations are avoided.

Some aspects of our design might limit comparison to some other experiments. The use of a two-role protocol could potentially lead to different behavior than in experiments where subjects play only one role.[12] We also have participants make each type of decision only once, whereas many experiments have participants make the same decision repeatedly. We use the strategy method for the Sender and simultaneously for the Returner (who answers contingent on a SEND decision), rather than a direct response approach.[13] We do not believe that these design choices are important factors in our results. What is important for our experiment is that the subjects' choices reflect "trusting" and "trustworthy" behaviors, an interpretation that is intrinsic to the standard trust game framework to which we adhere.

2.3.4 Measuring Mood

One possible mechanism by which specific experience (of being lied to, in our case) may affect behavior is due to its effect on a subject's mood. While mood effects may be driven to some extent by whether a subject is burned or not in the Deception experiment, we also seek a direct measure of mood in the experiment. To this end, we ask subjects to gauge their mood both at the start of the experiment (before instructions for the Deception game) and later in the experiment, after completion of the treatments but before the trust game, on the following scale:

bad down so-so good very good great

2.3.5 Logistics

The experiment was conducted in Economics and Sociology classes at the University of California, Merced, and Cal State East Bay. All three Treatments were conducted in all classes, resulting in a sample of 204 subjects. Sixty subjects were exposed to the Control Treatment; 72 subjects were exposed to the "Lied To" Treatment; and 72 subjects were exposed to the "Told the Truth" Treatment. The experimental instructions are attached as Appendix A.

[12] The literature gives a somewhat mixed picture on "role reversal" versus single role designs (Brandts and Charness, 2011). A number of authors have subjects play both roles in the trust game (for example, Chaudhuri and Gangadharan, 2007; Altmann et al., 2008). Charness and Rabin (2002), building on other literature, also have participants play both roles in a trust-type game that is played sequentially. In a subsequent paper, Charness and Rabin (2005) find that playing two roles (versus one) has no significant impact on their earlier results. Burks et al. (2003) study effects of two-role versus one (direct) role play in a trust game, when players are paid in both roles; they find that when participants are informed a priori that they will play both roles, there is a tendency to be less trusting and less trustworthy. These results suggest that the two-role design may potentially improve subjects' understanding of the game.

[13] Brandts and Charness (2011) provide evidence that the strategy method generally does not elicit significantly different responses in a variety of games, including trust.

2.4. Results

2.4.1. Baseline Results

Tables 2.2 and 2.3 describe broad results from our experiment. Table 2.2 presents proportions of subjects in total, and by treatment group (Control, Lied To, and Told the Truth), who made various decisions in the trust game. The decisions include: (1) Send1, whether to Send / trust when receiving a Message from the Returner indicating that s/ he (the Returner) intends to choose the generous option (Option A in Table 1); (2) Send2, whether to Send / trust when the Returner sends no Message; (3) OptGen, whether to choose the generous option (the Returner's decision), which we refer to as the trustworthy decision (following standard nomenclature); (4) MessGen, whether to Send a Message indicating selection of the generous option (again the Returner's decision); and (5) various combined choices of the Returner, including a Deceitful strategy (sending the Message, but choosing the ungenerous option, Deceit), an untrustworthy but not deceitful combination (not sending the Message and choosing the ungenerous option, UTBND), a trustworthy and truthful strategy (sending the Message and choosing the generous option, TWTruth), and a trustworthy strategy without communication (choosing the generous option and no Messaage, TWBNM).

Table 2.3 presents z-statistics for the respective differences between choices of the three treatment groups: (1) Lied To versus Control (column (1)), (2) Told the Truth versus Control (column (2)), and (3) Lied To versus Told the Truth (column (3)). The Table reveals that Lied To subjects were significantly less trusting (in terms of Send1); less inclined to send a Message; less likely to be trustworthy and truthful; and more likely to be trustworthy without sending a Message, all relative to both Control subjects and subjects who were Told the Truth. Lied To subjects were also significantly less likely to be trustworthy overall (by choosing the generous option), and more likely to be untrustworthy but not deceitful, relative to subjects who were Told the Truth. The absolute magnitudes of these differences are noteworthy. For example, 43 percent of Lied To (LT) subjects were trusting (Send1) compared with 61 percent of subjects who were Told the Truth (TT); 46 percent of LT subjects chose the generous option, compared with 62.5 percent of TT subjects; and 53 percent of LT subjects sent a Message, compared with over 76 percent for the TT group.

The bottom panels of Tables 2.2 and 2.3 present corresponding statistics for (i) the fraction of subjects who accepted / followed their Sender recommendations in the deception game, (ii) the initial (pre-experiment) mood report of participants (on a scale of zero to five, from "bad" to "great"), and (iii) the fraction who (after the deception game was complete and the treatments received) reported mood increases and mood decreases, respectively. A check for random assignment of our treatments is provided by comparison of the accept / follow decisions and initial moods of the subjects across treatment groups. If we have random assignment, there should be no significant differences between these indicators across the treatments, and indeed, Table 2.3 reveals no significant differences. Overall, approximately 62 percent of our student participants chose to accept / follow their Sender recommendation in the Deception game, with only slight variation from one treatment group to another.

Treatments did, however, affect mood *changes* in predictable directions. LT subjects were significantly more likely to experience mood decreases relative to either the Control or

TT participants. This means that some of the treatment effects on trust and trustworthiness (LT versus Control and TT) may in principle be attributable to resulting impacts on mood. We return to this issue in a moment.

Table 2.4 supplements Table 3 by reporting Probit regression results for Sender and Returner decisions, controlling for the subject's gender, initial mood, and fixed course effects. While the added correlates increase precision in estimated treatment effects, the broad conclusions of Tables 2.2-2.3 are upheld. For example, the Lied To treatment is estimated to reduce the probability of trust (Send1) by 19.3 percent, and the probability of a trustworthy and truthful strategy by 19.8 percent, both effects statistically significant.

Arguably surprising (and related) features of our baseline results are the large fraction of subjects who choose to Send when no Message is received (Send2) and the non-negligible fraction of subjects who are trustworthy but send no Message indicating their choice (TWBNM). In all treatments, over a third of subjects elect to Send when no Message is received.[14] On the Returner side, 11.8 percent of subjects choose TWBNM, which is roughly one-third of subjects who send no Message; corresponding fractions are highest for Lied To participants (19.4 percent of whom choose TWBNM, out of 47.2 percent who send no Message), next highest for participants who are Told the Truth (9.7 percent choosing TWBNM out of 23.6 percent who send no Message), and lowest for Control participants (5 percent choosing TWBNM out of 31.7 percent who send no Message).

Charness and Dufwenberg (2010) observe similar fractions in a similar experiment, but with different payoffs and much smaller subject numbers than we have; in their experiment, two of seven Senders who received no Message chose to Send (Send2), and three of seven Returners who sent no Message chose the generous option (TWBNM). We find that this phenomenon is more general and cannot be explained by randomness in participant choices. Perhaps the TWBNM strategy is motivated by psychic rewards to perceived acts of virtue untainted by a self-interested Message. If so, a reduced saliency of Messages might be expected to tip this calculus in favor of the TWBNM strategy, as we observe for the Lied To subjects.

2.4.2. Decomposing Treatment Effects

From our experiment, we are interested not only in identifying broad effects of our treatments – exposure to dishonesty or honesty – on behavior in trust relationships, but also the impact of this specific experience, as separate and distinct from mood effects and impacts of being burned or not in the first experiment (which can also affect mood and behavior). Our baseline comparisons (in Tables 2.2-2.4 above) conflate these phenomena. When a subject accepts his Sender recommendation in our Deception game (Experiment 1), he is hurt / burned when Lied To and benefited / not-burned when Told the Truth; conversely, when a subject rejects his Sender recommendation, he is not burned when Lied To and burned when Told the Truth. Now, if acceptance and rejection of Sender recommendations were equi-probable, there would be no differences in propensities to be burned or not burned across the two (LT and TT) treatment groups; in this case, cross-group differences could not be explained by a burned composition effect. However, in our

[14] On the sender side in our experiment, a risk neutral subject purely interested in his own payoffs would choose the Send2 strategy only if the probability of a generous Returner (given that no Message is sent) is 40 percent or higher. This condition is violated in the Control group and in our overall sample (the relevant benchmark given random matching across all subject participants).

experiment, roughly 62 percent of participants accepted their Sender recommendations, meaning a much higher fraction of burned subjects in the LT treatments than in the TT treatments. Our baseline results could therefore be explained by treatment effects on being burned, rather than a pure experience effect of being Lied To. For example, Lied To subjects may be less trusting because they are more likely to have been burned.

To disentangle these effects, we first break down our subject decisions by both treatment group and Acceptance / Rejection decisions from Experiment 1. The decomposed summary statistics from the experiment are given in Table 2.5. In principle, one way to net out burn effects would be to compare LT Accepters to TT Rejecters, both of whom are burned, and LT Rejecters to TT Accepters, neither of whom are burned. However, this comparison conflates potentially different drivers of Receiver Acceptance / Rejection decisions; accepters may be different types of people than rejecters. Indeed, the last column of Table 5 reports difference statistics for behavior of Control Accepters and Control Rejecters in our experiment. Because all Controls are equally likely to have been lied to or told the truth, there is no differential burned effect for Control Accepters vs. Rejecters. However, there are several key differences in behavior. The Control Accepters are significantly less responsive to communication in their trust decision, significantly more likely to choose the generous option, significantly less likely to be deceitful, and significantly more likely to be trustworthy without sending a Message. In sum, these statistics indicate that Accepters come from a different population than Rejecters, meaning that the comparisons proposed above would conflate the experience effect of being Lied To (vs. Told the Truth) with differences between Accepters and Rejecters.

We overcome this confound by constructing difference-in-difference statistics that exploit the Control subjects to adjust LT-versus-TT differences for burned and not-burned subjects, respectively; this is done by netting out corresponding differences between Control Accepters and Rejecters. For burned subjects, the difference-in-difference takes the difference between LT Accepters (LTA) and TT Rejecters (TTR), and subtracts out the corresponding difference between Control Accepters (CA) and Control Rejecters (CR). This difference-in-difference gives us a pure experience (vs. burned) effect of being Lied To (vs. Told the Truth). Similarly, for not-burned subjects, the difference-in-difference compares LT Rejecters (LTR) to TT Accepters (TTA), and subtracts out the corresponding difference between Control Rejecters and Control Accepters. Parallel difference-in-difference statistics give the pure burned effect for LT subjects, (LTA-LTR)-(CA-CR), and for TT subjects, (TTR-TTA)-(CR-CA).

Table 2.6 presents a first set of these decompositions. The first two columns give the pure Lied To (vs. TT) experience effect for burned and not-burned subjects, respectively; the third and fourth columns give pure burned effects for TT and LT subjects, respectively. z-statistics for the difference-in-differences are given in parentheses.[15] Columns (1) and (3) (and columns (2) and (4)) add up to the joint LT (vs. TT) effect for accepters (LTA-TTA), combining the experience and burned effects of the different treatments; this joint effect is presented in column (5).

---

[15] The z-statistics are calculated as z = D/se, where D=difference in difference=$(p_1-p_2)-(p_3-p_4)$ and se=$[\sum_{i=1}^{4} v_i/n_i]^{1/2}$ where $v_i= p_i(1-p_i)$, p_i=proportion in sample i, and n_i=size of sample i. For Send1-Send2, sample variances are used for the variance estimates v_i.

At the bottom of Table 2.6, we find that the propensity for a negative mood change is significantly raised by the LT (vs. TT) experience (for the not burned) and by the burned experience (for the TT subjects). The propensity for a positive mood change is significantly reduced by the burned experience (for the LT subjects). Being burned thus worsens our subjects' moods. Being Lied To also worsens mood, at least for those not burned. Because of the latter effect, we want to construct difference-in-difference statistics for trust outcomes that control for mood changes directly.

Table 2.7 presents generalized difference-in-difference statistics for pure LT (vs. TT) experience and pure burned effects, respectively, that control for gender, course effects, initial mood, and mood changes (positive and negative). These statistics are constructed from robust tests of coefficient differences in linear probability (OLS) estimations; p-values for the test statistics are reported in parentheses.

2.4.3 Main Results

Tables 2.6 and 2.7 reveal broadly similar experience and burned effects, and give us the main conclusions from our experiment:

First, we find that *being Lied To (versus Told the Truth) erodes trust* both for burned and not-burned subjects. However, for the not burned, trust is eroded when a Message is sent (Send1), whereas for the burned, trust is eroded when a Message is not sent (Send2). The LT experience effects are large. For the not-burned, the LT experience reduces the propensity for trust Send1 by an estimated 33.5 percent (Table 2.7), compared with an average rate of trust for Control subjects of 58.3 percent (Table 2.2). For the burned, the LT experience reduces the propensity for trust Send2 by an estimated 39.7 percent (Table 2.7), compared with a Control subject propensity of 36.8 percent (Table 2.2). These numbers capture the intrinsic effects of lies that we discussed at the start of the paper, separate from any treatment effects on mood and/or being burned or not in Experiment 1.

We expect effects of being burned to be different for TT and LT subjects. TT subjects are burned when *not following* their Sender recommendation in Experiment 1; we expect being burned to motivate different (*more* trusting) behavior in the trust game and/or to make the signal of truthfulness more salient, again favoring more trusting choices. Conversely, LT subjects are burned when following their Sender recommendations in Experiment 1; we therefore expect being burned to motivate *less* trusting behavior in the trust game. For Send1, however, we find no significant burned effect, leading to a joint (experience and burned) effect of the LT treatment that is negative and significant (column (5), Tables 2.6-2.7). For Send2, we find a significant positive burned effect on TT subjects, consistent with expectations, but no burned effect on LT subjects. The joint (experience and burned) effect combines the negative experience effect (of LT) on Send2 (column (1), Tables 2.6-2.7) with the positive burned effect (for the TT subjects, column (3), Tables 2.6-2.7), for a net null effect.

Second, *being Lied To (versus Told the Truth) and being burned interact to erode trustworthiness*. The experience effect of being Lied To (for the burned), and the burned effect (for the Lied To), are to reduce trustworthiness by a statistically and economically significant fraction. The LT experience reduces the propensity for overall trustworthiness (OptGen) by an estimated 51.5 percent and the propensity for both truth and trustworthiness (TWTruth) by an estimated 45.1 percent (Table 2.7, column (1)). The burned effect (for the Lied To) reduces overall trustworthiness (OptGen) by an estimated 57.2 percent and

TWTruth by an estimated 35.8 percent (Table 2.7, column (4)). However, we find no significant LT experience effect on trustworthiness for the not-burned, and no significant burned effect for the TT subjects. Hence, being Lied To and being burned each reduce trustworthiness in our experiment, but only when the other is present.

This general conclusion is reinforced by two more nuanced results. Being Lied to (for the burned) is estimated to raise the likelihood of untrustworthiness with no deceit (UTBND); and being burned (for the Lied To) is estimated to lower the likelihood of trustworthy behavior with no Message (TWBNM). Interestingly, however, we find no significant Lied To effects on the propensity for Deceit (an untrustworthy choice together with a deceitful Message indicating the opposite).[16]

Third, the propensity to send a Message is reduced both by the pure experience effect of being Lied To (versus Told the Truth) and by being burned in Experiment 1 (columns (1)-(4), Tables 2.6-2.7). However, none of these effects is statistically significant individually and the estimated impact of being burned is particularly small (see p-values in columns (3)-(4) of Table 2.7). Combining Lied To and burned effects on the Accepters (column (5) of Tables 2.6-2.7) therefore reflects primarily the experience effect. And the combined effect – which is measured with more precision – is statistically significant and negative.

We conclude that *the Lied To experience reduces communication*. For example, the Lied To experience, for burned subjects, reduces the fraction of Returners who send a Message by 26.8 percent (Table 2.7, column (1)), compared with an overall propensity to send Messages of 68.3 percent among Control participants (Table 2.2). The reduced reliance on communication to promote trust generally contributes to reduced trustworthiness, as described above.

In summary, we find that the pure experience effect of being Lied To (versus Told the Truth) erodes trust, trustworthiness, and communication in our experiment. These effects are distinct from (and control for) treatment impacts on mood and being burned.

2.4.4. The Role of Beliefs

Sapienza, Toldra-Simats and Zingales (STZ, 2013) suggest that the best measure of trust that one can obtain from the Berg, et al. (1995) game is one based on expectations: For a given amount of money sent to a Returner, how much money does a Sender expect to be returned? For our simplified trust game, this question is addressed with a measure of how likely a subject believes it is that a Returner will choose the generous return strategy. Given communication in our experiment, this question is well posed only when conditioned on the receipt of a Message. In order to construct this modified STZ (2013) measure of trust, we used an incentive compatible approach to elicit this belief – along with three others – from subjects in our experiment.

Specifically, we asked subjects to predict four outcomes from the experiment, paying $1 for each prediction that was within 5 percent (plus or minus) of the true percentage (using 5 percentage point bands). The outcomes for which we solicited predictions are:

Q1. The fraction of Senders who Send when receiving a Message (Send1).

Q2. The fraction of Returners who, if Sending a Message, choose the generous option.

[16] Burns appear to increase deceit for the Lied To subjects, although the effect is not quite significant (p-value of .11).

Q3. The fraction of Returners who send a Message.

Q4. The fraction of participants who indicate (in a separate "yes" or "no" question) that a Returner *should* choose the generous option if sending a Message.

Q2 gives the STZ measure of trust for our game. Table 2.8 provides summary and difference-in-difference statistics for the four beliefs, as well as whether subjects think a Returner should choose the generous option if sending a Message (Q5, 1 for "yes," 0 for "not necessarily").

Broadly, we find two main differences between subject answers on the belief and norm questions across the treatment groups (LT, TT, and Control). Being Lied To has a significant negative effect on the perceived likelihood of trust Send1 (Q1) and the perceived norm on whether Returners *should* be trustworthy when promising to be so (Q4). These results likely reflect false consensus effects – that is, subject beliefs that conform to the subjects' own choices (Ross, et al., 1977; Ellingsen, et al., 2010).

Decomposing the treatment effects further, with difference-in-difference statistics, we find only one significant effect. The Lied To experience, for burned subjects, significantly reduces the predicted frequency with which Returners will be generous when sending the Message (Q2). In other words, *we find a negative Lied To effect on the STZ measure of trust*. Note, however, that we find no significant Lied To experience effects on beliefs and norms for the not-burned. Hence, the experience effects that we identify in column (2) of Tables 2.6-2.7 cannot be attributed to treatment effects on beliefs, at least not the ones that we measure. Most importantly, the significant Lied To experience effect in eroding trust (Send1, for the not-burned) does not appear to be attributable to beliefs.

2.5. Conclusion

We find that being on the receiving end of a lie (vs. a truth) leads to an erosion of trust, even in interactions with those who have nothing to do with the initial deception and even though the deceptive act is known to have no bearing on the overall propensity for dishonesty among experimental participants. Given the central role that trust is known to play in promoting economic interchange and growth, this conclusion suggests that social institutions that deter dishonesty and promote norms of truthfulness are of potential economic value.

A key feature of the analysis is the identification of an individual experience effect of the "Lied To" and "Told the Truth" treatments, controlling for mood, the impact of being burned or not, and overall Sender propensities for honesty. Separate from everything else, the individual experience alters behavior. These results expose a potentially general link between individual experience and behavior in social interchange. However, our conclusions are admittedly preliminary in the sense that they do not speak to the mechanisms by which our treatments have the effects that they do in our sample. A great deal of research studies what drives or deters trust and trustworthiness, including (among others) expectations (Sapienza, et al., 2013), reciprocity (Charness and Rabin, 2002)), and guilt aversion (Charness and Dufwenberg, 2006). One possible interpretation of our results is that reciprocal preferences that drive trust are determined by a broad social context and specific experiences in a compendium of social interactions, including experiences of lies and truths; lies may reduce the positive reciprocity and/or the extent of guilt aversion that lead to trustworthy choices. Although this interpretation is plausible, our coarse examination of the

trust game does not decompose treatment effects on *different* drivers of trust per se; this is a subject that we believe merits further study.

Table 2.2
Sample Summary Statistics

	All Observations (n=204)	Treatments		
		Control (n=60)	Lied To (n=72)	Told the Truth (n=72)
Trust				
Send1 (Trust When Message Rec'd)	0.539	0.583	0.431	0.611
Send2 (Trust When No Message Rec'd)	0.368	0.333	0.361	0.403
Send1 – Send2 (Effect of Message on Trust)	0.172	0.250	0.069	0.208
Trustworthiness				
OptGen (Trustworthy: Generous Option Chosen)	0.525	0.483	0.458	0.625
MessGen (Message Sent)	0.659	0.683	0.528	0.764
Deceitful (Message Sent, but Untrustworthy)	0.250	0.250	0.264	0.236
UTBND (Untrustworthy, but no deceit/no Message)	0.225	0.267	0.278	0.139
TWTruth (Trustworthy & Truthful Message)	0.407	0.433	0.264	0.528
TWBNM (Trustworthy, No Message)	0.118	0.050	0.194	0.097
Deception Game Decision and Mood				
Accept (in Deception Game)	0.623	0.617	0.597	0.653
Initial Mood	2.936	2.783	2.958	3.042
Positive Mood Change^{+}	0.127	0.100	0.153	0.125
Negative Mood Change^{+}	0.113	0.067	0.222	0.042

+ Positive (Negative) Mood Change = 1 if Mood Change (post-treatment minus pre-treatment) > (<) 0, 0 otherwise.

Table 2.3.
Difference Statistics Across Treatments

	Lied To vs. Control (z-statistic)	Told the Truth vs. Control (z-statistic)	Lied To vs. Told the Truth (z-statistic)
Trust			
Send1	-1.77*	0.32	-2.20**
Send2	0.35	0.83	-0.52
Send1 – Send2	-3.05***	-0.57	-2.46**
Trustworthiness			
OptGen	-0.30	1.65	-2.04**
MessGen	-1.93*	1.03	-3.06***
Deceitful	0.19	-0.19	0.39
UTBND	0.15	-1.82*	2.08**
TWTruth	-2.17**	1.08	-3.36***
TWBNM	2.71***	1.05	1.67*
Deception Decisions and Mood			
Accept	-0.24	0.43	-0.69
Initial Mood	0.92	1.35	-0.50
Positive Mood Change	0.95	0.45	0.48
Negative Mood Change	2.70***	-0.63	3.32***

*,**,*** Significant at 10%, 5%, 1% (two-sided).

Table 2.4.
Probit Regressions

Dependent Variable		Marginal Effect (Robust t-statistic)					LT vs. TT
		Lied To (LT)	Told the Truth (TT)	Male Gender	Initial Mood	Course Effects	Difference (p-value)[+]
Send1 (Trust When Message Rec'd)	Model 1	-0.193 (-2.117)**	-0.007 (-0.081)	0.065 (0.895)	0.089 (2.535)**	Yes	-0.186 (0.028)**
	Model 2	-0.160 (-1.823)*	0.019 (0.218)	0.072 (1.015)	No	No	-0.141 (0.032)**
Send 2 (Trust When No Message Rec'd)	Model 1	0.042 (0.481)	0.084 (0.975)	0.058 (0.842)	-0.014 (-0.431)	Yes	-0.042 (0.604)
	Model 2	0.043 (0.508)	0.083 (0.971)	0.052 (0.764)	No	No	-0.040 (0.627)
Send1-Send2 (Effect of Message on Trust)[++]	Model 1	-0.226 (-2.039)**	-0.090 (-0.838)	0.005 (0.053)	0.021 (2.328)**	Yes	-0.136 (0.205)
	Model 2	-0.201 (-1.812)*	-0.063 (-0.575)	0.018 (0.211)	No	No	-0.138 (0.197)
MessGen (Message Sent)	Model 1	-0.155 (-1.830)*	0.081 (0.961)	0.136 (1.977)**	0.010 (0.320)	Yes	-0.236 (0.004)***
	Model 2	-0.162 (-1.940)*	0.075 (0.889)	0.145 (2.152)**	No	No	-0.087 (0.003)***
OptGen (Trustworthy: Generous Option Chosen)	Model 1	-0.043 (-0.476)	0.133 (1.508)	0.059 (0.816)	0.037 (1.05)	Yes	-0.176 (0.037)**
	Model 2	-0.014 (-0.163)	0.151 (1.730)*	0.049 (0.695)	No	No	-0.165 (0.047)**
Deceitful (Message Sent, Untrustworthy)	Model 1	0.047 (0.597)	0.005 (0.066)	0.045 (0.727)	-0.039 (-1.360)	Yes	0.042 (0.573)
	Model 2	0.013 (0.168)	-0.017 (-0.221)	0.053 (0.870)	No	No	0.030 (0.684)
UTBND (Untrustworthy, No Message)	Model 1	0.004 (0.056)	-0.132 (-1.892)*	-0.101 (-1.679)*	-6.5e-05 (-0.002)	Yes	0.136 (0.047)**
	Model 2	0.004 (0.052)	-0.132 (-1.87)*	-0.101 (-1.711)*	No	No	0.136 (0.046)**
TWBNM (Trustworthy, No Message)	Model 1	0.184 (2.664)***	0.091 (1.356)	-0.026 (-0.665)	-0.008 (-0.408)	Yes	0.093 (0.098)*
	Model 2	0.200 (2.783)***	0.101 (1.457)	-0.040 (-0.998)	No	No	0.099 (0.099)*
TWTruth (Trustworthy & Truthful Message)	Model 1	-0.198 (-2.255)**	0.070 (0.807)	0.088 (1.213)	0.044 (1.245)	Yes	-0.268 (0.001)***
	Model 2	-0.179 (-2.078)**	0.087 (1.012)	0.089 (1.266)	No	No	-0.266 (0.001)***

*,**,*** Significant at 10%, 5%, 1% (two-sided).

[+] p-value for test of equal coefficients on Lied To and Told the Truth (heteroskedasticity-robust). [++]OLS for the difference between the 0-1 choices to Trust with a Message and to Trust without a Message. All other models are Probit estimations.

Table 2.5.
Decomposed Summary Statistics Across Treatments

	Control Accept (n=37)	Control Reject (n=23)	LiedTo Accept (n=43)	LiedTo Reject (n=29)	Told the Truth Accept (n=47)	Told the Truth Reject (n=25)	Control Accept – Control Reject Difference (z-statistic)
Trust							
Send1	0.514	0.696	0.442	0.414	0.617	0.600	-0.182 (-1.44)
Send2	0.405	0.217	0.349	0.379	0.340	0.520	0.188 (1.59)
Send1 – Send2	0.108	0.478	0.093	0.034	0.277	0.080	-0.370 (-3.19)***
Trustworthiness							
OptGen	0.568	0.348	0.349	0.621	0.638	0.600	0.220 (1.71)*
MessGen	0.649	0.739	0.488	0.586	0.745	0.800	-0.090 (-0.75)
Deceitful	0.162	0.391	0.279	0.241	0.191	0.320	-0.227 (-1.93)*
UTBND	0.270	0.261	0.372	0.138	0.170	0.080	0.009 (0.08)
TWTruth	0.486	0.348	0.209	0.345	0.553	0.480	0.138 (1.08)
TWBNM	0.081	0.000	0.140	0.276	0.09	0.120	0.081 (1.81)*
Mood							
Initial Mood	2.838	2.696	2.91	3.034	3.085	2.96	0.142 (0.508)
Positive Mood Change	0.054	0.043	0.070	0.276	0.170	0.04	0.011 (0.19)
Negative Mood Change	0.135	0.044	0.302	0.103	0.000	0.120	0.091 (1.29)

*,**,*** Significant at 10%, 5%, 1% (two-sided).

Table 2.6.
Difference-in-Difference Statistics

	Lied To Effect		**"Burned" Effect**		**Total Lied To and Burned Effect for Accepters**
	For the "Burned"	For the "Not Burned"	For the Told-Truth	For the Lied To	
	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)	Difference (z-statistic)
	(1) (LTA-TTR) - (CA-CR)	(2) (LTR-TTA) - (CR-CA)	(3) (TTR-TTA) - (CR-CA)	(4) (LTA-LTR) - (CA-CR)	(5) LTA-TTA
Trust					
Send1	0.024 (0.13)	-0.385 (-2.25)**	-0.199 (-1.12)	0.210 (1.24)	-0.175 (-1.69)*
Send2	-0.359 (-2.68)***	0.227 (1.39)	0.368 (2.14)**	-0.218 (-1.35)	0.009 (0.08)
Send1-Send2	0.383 (1.72)*	-0.612 (-2.81)***	-0.567 (-2.61)***	0.429 (1.93)*	-0.184 (-1.16)
Trustworthiness					
OptGen	-0.471 (-2.62)***	0.202 (1.17)	0.181 (1.02)	-0.492 (-2.91)***	-0.290 (-2.87)***
MessGen	-0.221 (-1.33)	-0.249 (-1.52)	-0.035 (-0.22)	-0.007 (-0.04)	-0.256 (-2.58)***
Deceitful	0.188 (1.11)	-0.179 (-1.16)	-0.101 (-0.61)	0.267 (1.74)*	0.087 (0.98)
UTBND	0.283 (1.89)*	-0.023 (-0.16)	-0.081 (-0.57)	0.225 (1.51)	0.202 (2.20)**
TWTruth	-0.409 (-2.31)**	-0.070 (0.40)	0.065 (0.36)	-0.274 (-1.67)*	-0.344 (-3.60)***
TWBNM	-0.062 (-0.63)	0.272 (2.65)***	0.116 (1.28)	-0.217 (-2.00)**	0.054 (0.82)
Mood					
Initial Mood	-0.195 (-0.51)	0.092 (-0.23)	0.017 (0.05)	-0.270 (-0.74)	-0.178 (-0.85)
Positive Mood Change	0.019 (0.240)	0.116 (1.02)	-0.120 (-1.35)	-0.217 (-2.03)**	-0.101 (-1.49)
Negative Mood Change	0.091 (0.75)	0.195 (2.15)**	0.211 (2.16)**	0.108 0.95	0.302 (4.32)***

*,**,*** Significant at 10%, 5%, 1% (two-sided).
Legend: CA = Control Accept, CR = Control Reject, LTA = Lied To Accept, LTR = Lied To Reject, TTA = Told the Truth Accept, TTR = Told the Truth Reject.

Table 2.7.
Difference-in-Difference Statistics Controlling for Course Effects, Gender, Initial Mood, and Mood Change

	Lied To Effect		**"Burned" Effect**		**Total Lied To and Burned Effect for Accepters**
	For the "Burned"	For the "Not Burned"	For the Told-Truth	For the Lied To	
	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)	Difference (p-value)
	(1) (LTA-TTR) - (CA-CR)	(2) (LTR-TTA) - (CR-CA)	(3) (TTR-TTA) - (CR-CA)	(4) (LTA-LTR) - (CA-CR)	(5) LTA-TTA
Trust					
Send1	-0.031 (0.867)	-0.335 (0.062)*	-0.163 (0.379)	0.141 (0.449)	-0.194 (0.083)*
Send2	-0.397 (0.028)**	0.25 (0.139)	0.381 (0.030)**	-0.266 (0.132)	-0.016 (0.881)
Send1-Send2	0.366 (0.109)	-0.585 (0.007)***	-0.544 (0.014)**	0.407 (0.078)*	-0.178 (0.214)
Trustworthiness					
OptGen	-0.515 (0.006)***	0.279 0.118	0.221 (0.227)	-0.572 (.002)***	-0.294 (0.882)
MessGen	-0.268 (0.126)	-0.214 (0.201)	-0.048 (0.772)	-0.102 (0.564)	-0.316 (0.003)***
Deceitful	0.183 (0.269)	-0.231 (0.124)	-0.159 (0.325)	0.256 (0.111)	-0.024 (0.798)
UTBND	0.332 (0.041)**	-0.048 (0.748)	-0.063 (0.677)	0.317 (0.059)*	0.269 (0.005)***
TWTruth	-0.451 (0.013)**	0.017 (0.925)	0.111 (0.228)	-0.358 (0.047)**	0.340 (0.002)***
TWBNM	-0.063 (0.523)	0.262 (0.014)**	0.111 (0.548)	-0.214 (0.052)*	0.048 (0.507)

Legend: CA = Control Accept, CR = Control Reject, LTA = Lied To Accept, LTR = Lied To Reject, TTA = Told the Truth Accept, TTR = Told the Truth Reject. Difference-in-difference (and difference) statistics are obtained from OLS regressions that include treatment dummies, course effects, gender, initial mood, positive mood change, and negative mood change. p-values are presented for heteroskedasticity-robust test statistics on linear restrictions of zero difference-in-difference (zero difference in column (5)).

Table 2.8.
Subject Beliefs about Behavior and Norms

Participant Predictions →	Q1 (% choosing Send1)	Q2 (% choosing OptGen when choosing MessGen)	Q3 (% choosing MessGen)	Q4 (% saying "yes" on Q5)	Q5 (Norm) (1 = "yes" = *should* choose OptGen if send MessGen)
Summary Statistics: Mean (Standard Deviation)					
All Obs. (n=204)	51.005 (25.318)	47.857 (24.736)	55.025 (25.700)	53.719 (22.201)	42.365 (49.536)
Control (n=60)	55.583 (23.526)	46.583 (23.926)	56.667 (25.926)	54.915 (20.605)	55.932 (50.073)
Lied To (LT) (n=72)	45.278 (25.604)	44.577 (26.949)	53.611 (26.474)	48.333 (23.271)	33.333 (47.471)
Told Truth (TT) (n=72)	52.917 (25.739)	52.153 (22.764)	55.069 (24.972)	58.125 (21.533)	40.278 (49.390)
Difference Statistics (z-statistics)					
LT – Control	-10.305 (-2.41)**	-2.006 (-0.45)	-3.056 (-0.67)	-6.582 (-1.72)*	-22.599 (-2.66)***
TT – Control	-2.666 (-0.62)	5.570 (1.36)	-1.598 (-0.36)	3.210 (0.87)	-15.654 (-1.81)*
LT – TT	-7.639 (-1.79)*	-7.576 (-1.82)*	-1.458 (-0.34)	-9.792 (-2.62)***	-6.945 (-0.87)
Difference-in-Difference [+] (p-values)					
Lied To Effect for "Burned"	0.069 (0.994)	-17.675 (0.045)**	6.483 (0.491)	-8.318 (0.314)	-1.787 (0.924)
Lied To Effect for "Not Burned"	-9.036 (0.272)	-0.455 (0.959)	-5.605 (0.525)	-10.601 (0.162)	-11.985 (0.340)
"Burned" Effect for TT	-10.475 (0.220)	11.772 (0.158)	-12.149 (0.181)	0.792 (0.919)	-1.842 (0.920)
"Burned" Effect for LT	-1.37 (0.875)	-5.448 (0.566)	-0.061 (0.995)	3.076 (0.703)	8.356 (0.661)

*,**,*** Significant at 10%, 5%, 1% (two-sided).

[+] Difference in difference statistics are calculated from heteroskedasticity-robust OLS regressions that control for course effects, gender, initial mood, and mood changes. Lied To Effect for "Burned" = (LTA-TTR)-(CA-CR), Lied To Effect for "Not Burned" = (LTR-TTA)-(CR-CA), "Burned" Effect for TT = (TTR-TTA)-(CR-CA), "Burned" Effect for LT = (LTA-LTR)-(CA-CR).

Chapter 3
Pure Lying Aversion

3.1 Introduction

A recent literature has evolved concerning people's propensity to lie or tell the truth and the motivations behind such decisions. Pure economic man concerned only with his own material payoff is thought not to suffer an intrinsic cost of lying. Thus, the decision to lie or not should be based on the probability of being found out and the consequences resulting from that (Lewicki 1983). Yet, in practice, it is well-documented that a sizeable proportion of people do in fact tell the truth even when doing so is to their material disadvantage, suggesting that many people face an intrinsic cost of lying.

A natural distinction in the literature can be made between situations where there is a direct economic consequence of the lie to others, which might cause an aversion to lying stemming from guilt aversion, shame aversion or altruism, (for example, Battigalli and Dufwenberg (2007), Battigalli, Charness and Dufwenberg (2012), Charness and Dufwenberg (2006, 2010), Gneezy (2005), Gneezy et al. (2013), Greenberg et al. (2015), Lundquist et al. (2009)) versus where a lie has no direct negative consequences for others (Mazar et al. (2008), Shalvi et al. (2011a,b), Lopez-Perez and Spiegelman (2012), Fischbacher and Föllmi-Heusi (2013), Gibson et al. (2013), Utikal and Fischbacher (2013), Abeler et al. (2014)), and thus the aversion to lying arguably stems from the act of lying itself rather than its impact on others.

This second situation might seem particularly puzzling. If no-one is directly negatively affected by the lie, why would people not lie if it is to their material advantage? Yet this second strand of the literature has established that many people are not willing to lie to benefit themselves monetarily, even when no-one is directly adversely affected by the lie. Indeed, Abeler et al. (2014) in a telephone field study with a random sample of the general public, and Utikal and Fischbacher (2013), using a sample of nuns, find reverse lying; that is, people reported fewer than would be randomly observed of the financially rewarding coin flips (in the Abeler et al. study) and die rolls (in the Utikal and Fischbacher study). In studies utilizing students, lying is generally found in the expected direction but a substantial proportion of participants report truthfully. Additionally, several papers have found evidence of incomplete lying, including Mazar et al. (2008), Fischbacher and Föllmi-Heusi (2013), Shalvi et al. (2011 a.,b.), Utikal and Fischbacher (2013) and Hao and Houser (2011). In these studies, participants have to make a report following an action which no-one else observes, such as reporting the outcome of a die roll, or reporting the number of matrices they were able to solve in a given time; thus, if they were payoff maximizing they would report the highest paying outcome, while if they were honest, the distribution of reports would be expected to be uniform in the case of the single die roll, or the same as the distribution of a control group in the case of the matrix task. However, these studies find that a

substantial proportion of participants report more than the expected true value, but less than the payoff maximizing value. This is explained by Mazar et al. (2008) in terms of maintaining a self-concept of honesty. Shalvi et. al (2011a,b) develop this theory further by focusing on justified versus unjustified lies. Their experiments allow some participants to roll the die several times, but they are asked to report the outcome of the first roll. They find that lying is more common when participants have the opportunity to report another number that they have actually rolled, than when, in order to lie, they have to choose a number that they have not rolled. They also find that people consider the latter to be a worse lie than the former. An alternative explanation for incomplete lying is proposed and tested by Hao and Houser (2011), who have participants first predict the number they will roll with a die, then roll and report the unobserved result. By allowing participants to express preferences for appearing honest and being honest separately, they find that 95% express a preference for appearing honest, while only 44% are actually honest in their actions when they have an opportunity to cheat. Further, they find that participants give up on average 25% of the maximum payment in order to maintain an honest appearance.

The present paper focuses on the single choice of whether or not to lie, that is ,"clear-cut lies……where the liar knows that what he is communicating is not what he believes, and where he has not deluded himself into believing his own deceits" (Bok, 1978), similar to Lopez-Perez and Spiegelman (2013) and Gibson et al. (2013), rather than where there is room for self-deception as in the Shalvi et al. (2011a,b) studies. Lopez-Perez and Spiegelman (2013) discuss five different possible motivations for lying aversion; selfishness (through reputation building), altruism, guilt aversion, belief-dependent lying aversion, and pure lying aversion. They design a simple experiment to detect the existence of pure lying aversion, controlling for the other motivations, and find that approximately one third of participants exhibit pure lying aversion, given the economic incentive in their experiment of a one euro gain to lying. The Gibson et al. (2013) study, in a within subject design, varies the economic gain to lying, and finds that people respond to economic incentives monotonically but with a diminishing marginal effect over their range of variation.

In this paper we report the results of several classroom experiments which study how different factors affect pure lying aversion. We vary the economic gain to lying, extending the range of variation relative to the Gibson et al. (2013) design, and find that lying does not increase monotonically as the economic gain increases. We also vary the strength of the statement one is making when telling the truth or lying, and find that this has a large effect on truthfulness that is surprisingly constant as the economic incentives change. In a separate two stage experiment, we find that exposure to information on whether one has been lied to or told the truth in an earlier interaction has a significant effect on truthfulness. Thus the paper contributes to the literature by extending the study of potential determinants of pure lying aversion in several directions. While the situation where one person's lying has a direct negative effect on another's payoff, as captured in the Sender-Receiver[17] games, might seem more natural and important, the pure lying

[17] A game where a Sender has information about the payoff structure, while the other participant, the Receiver does not. The Sender must send a message to the Receiver about which Option choice will benefit him/her, after which, the Receiver, based only on the message, must choose which Option should be implemented.

aversion scenario is of interest for at least two reasons. If we are interested in how much of one's behavior stems from one's attitude toward lying itself, studying this in a context where there is no potential victim allows us to isolate this effect from other motivations, such as altruism or guilt aversion. Secondly, the situation of no-one being immediately negatively affected arguably reflects situations where we are faced with lying or telling the truth in interactions with large bureaucracies such as the government or large corporations. In such situations, one might perceive that one's own lack of honesty would have an insignificant financial effect. However, while, for example, the Allingham and Sandmo (1972) model of tax compliance views the taxpayer as a rational maximizer of the expected utility of the tax evasion gamble, this does not explain why a substantial portion of taxpayers pay all taxes owed all of the time (Alm 2012, Lewis et. al. 2009). Similarly, Pruckner and Sausgruber (2013) found that nineteen percent of people paid full-price, and a further forty two percent paid a positive amount, for a newspaper, when it was sold on the honor system with no monitoring.

In our experiments, based on the procedure used in Lopez-Perez and Spiegelman (2013), participants are asked to identify the color of a dot, and then to send a message to someone about the color. There is no payoff consequence for the person receiving the message, only the participant sending the message; thus, this experiment attempts to measure lying aversion in its purest sense, distinct from any possible effect due to other motivations such as guilt aversion or altruism. The experimental procedure here differs from that in Lopez-Perez and Spiegelman (2013) however, in that they do have a receiver who receives a payout, even though the sender's decision does not affect the amount of the receiver's payout. We do not have a paid receiver, although the participant is asked to choose a message to send to someone about the color of the dot. In this, our experimental set-up is more similar to Gibson et al. (2013) in which participants, in the role of CEOs, have to announce earnings for their companies, and can choose a truthful or untruthful announcement. No-one is directly affected by the announcement. However, because the participants are role-playing and both the "truthful" and "untruthful" earnings announcements are fictitious, this could potentially give a less accurate measure of how many people are averse to lying than an experiment such as ours, which does involve a factual statement about the color of an actual dot on the paper. For example, Gibson et al. report that only 82 percent of their participants told the truth when there was no economic cost of doing so, whereas in our different sessions we found from 86 to 96 percent of participants told the truth when there was no economic cost of doing so, with only one session having less than 91 percent truthful.

3.2 The Experiments

3.2.A Strength of Message and Economic Incentive Experiment.

3.2.A.1 Introduction and Description of Experiment

In our first set of experiments, we varied the strength of the message the participants were sending to verify the color of the dot, and the size of the economic incentive to lie that the participants faced. A green or blue dot was placed at the top of the page in the experimental questionnaire. Participants, who were students in Economics classes at UC Merced, were told that they were to choose a message to send

to another person about the dot, but there was no economic consequence for the recipient of the message. We randomized the actual dot color across questionnaires, and participants were asked to write down the color of the dot, to check for color-blindness.

Each participant was faced with two scenarios, one where there was no economic gain to lying, and the other where there was a gain. Additionally, each participant was treated with either weak or strong statements. Thus the design is between-subject except that everyone faced the zero gain scenario as well as one other. The weak message statements were: "The dot is blue" and "The dot is green". The strong message statements were: "I solemnly swear that the dot is blue" and "I solemnly swear that the dot is green".

Lundquist et al. (2009) find that strong messages reduce lying relative to weaker messages, in an experiment consisting of participants taking a cognitive test, then reporting their score to a potential partner to try to persuade them to enter into a contract. However, the effect was not quite statistically significant for their whole sample, though it was significant for the subsample of students who were not attending the Stockholm School of Economics. While they did not find a significant difference overall between strong and weak messages generated by the experimenter, Lundquist et al. (2009) did find that free-form messages were significantly more effective in reducing lying than either weak or strong messages generated by the experimenter. Charness and Duwfenberg (2010) also find that experimenter generated messages are only weakly significant in promoting trustworthy behavior in their variant of a trust game[18], while free-form messages were strongly significant. Capellen et al.(2013) also study the effect of the message on lying aversion, in the context of a White Lies experiment based on the Erat and Gneezy (2012) design. This is a Sender-Receiver game in which the lying choice benefits both the Sender and Receiver. In the Capellen et al. (2013) study, Senders personalize the message by writing an anecdote about the number they are choosing to send to their paired Receivers. They found that using personalized messages reduced lying from 69% to 55% in their sample. Serra-Garcia et al. (2013), in a public good game[19], find that there is less lying and less free-riding when participants can send a message about their own contribution, versus a message about the return to a contribution.

Unlike our study, the experiments in the above studies all involved sending a message to a recipient who was directly affected by the outcome of the experiment; thus the messages involved persuading a receiver to take an action, possibly through altering beliefs or preferences (DellaVigna and Gentskow 2010). A priori, it seems reasonable to expect less of an effect of a stronger message in the context we are using here, where there is no question of persuasion and no-one is directly affected monetarily by the lie. Yet, on the contrary, we find an effect that is statistically significant and larger in

[18] In a trust game, one participant, the Sender, must decide whether or not to send their initial endowment to the other participant, the Returner, in which case the experimenter increases the size of the endowment. The Returner then decides how to split this enlarged pie between him/herself and the Sender. In the literature, sending is viewed as trusting behavior, while returning a generous portion to the Sender is considered trustworthy behavior.

[19] In a public goods game subjects secretly choose how much of their initial endowment to put into a public pot. The contributions to this pot are multiplied by a factor (greater than one and less than the number of players, N) and this "public good" payoff is evenly divided among players. Each subject also keeps the portion of their endowment that they do not contribute.

magnitude than the effects in the Lundquist et al. (2009) and Cappelen et al. (2013) papers. In a related treatment in one of their experiments on cheating and self-concept, Mazar et al. (2008) look at the effect of signing an honor code statement, "I understand that this short survey falls under _____'s honor code." They found that including this statement eliminated cheating, with the effect being the same at universities that did not in fact have an explicit honor code and at one which did. However, invoking a university's honor code may have different effects on lying aversion than altering the strength of a simple statement about a dot, as it may invoke feelings of identity with the university, and being dishonest about one's own performance on a task may be different than making a report about an external factual object.

For the economic incentive variation in our experiment, in the second scenario some subjects received \$2 if they lied and \$1 if they told the truth, while others received \$3 if they lied and \$1 if they told the truth, and yet others received \$4 if they lied and \$1 if they told the truth. Thus the gain from lying was \$1, \$2 and \$3 respectively. Gneezy (2005), in the Sender-Receiver game where the interests of the Sender and Receiver were not aligned, found that, for a given hurt to the Receiver, Senders lie more as the economic gain to themselves increases. Gibson et al. (2013), in their CEO announcement design where no-one is harmed by the lie, find a monotonically increasing function of lying as the economic gain increases. To our knowledge, no-one has investigated the joint effect of the strength of the message and increasing economic gain from lying on truthfulness.[20] With two strengths of message and 3 payout variations, we have 6 different treatment groups.

3.2.A.2 Logistics

The experiment was conducted in several Economics classes at the University of California, Merced. Treatments were randomly assigned across classes, resulting in a sample of 302 subjects. One hundred and Sixty one subjects were treated with weak messages, and one hundred and forty one with strong messages. Questionnaires were identified by ID numbers, with tags attached that had matching ID numbers. Students retained these tags and used them to collect their payments at the beginning of the following week's class. The experimental instructions are attached in Appendix B.

3.2.A.3 Results

Table 3.1 shows the proportions telling the truth for the six treatment groups and the number of participants in each group, along with the proportions telling the truth when there is no economic gain from lying.

As can be seen in the table, there is considerable variation in truthfulness as the economic incentive varies, and between the strong and weak message treatments. Over ninety percent of participants reported the true color of the dot when there was no economic cost of doing so, in both the strong and weak message treatments. This

[20] The Mazar et al. (2008) study did look at the effect of signing a statement acknowledging an honor code when the incentive for cheating was varied. However, as discussed above, invoking an honor code is not strictly comparable to varying the strength of a message, and since it entirely eliminated cheating, there was no variation with respect to the different economic incentives.

suggests at least a small preference for honesty, since if participants were indifferent to being honest or lying per se, as assumed by the standard economic model of the self-interested utility maximizer, when the economic outcome was the same we would expect to see truthfulness of around fifty percent.

Table 3.2 gives differences in proportions with z- statistics for Strong vs Weak Messages. Unlike Lundquist et al (2009), we find a significant difference for the (experimenter-generated) strength of message, and it holds for all positive economic incentive levels. The difference in truth-telling when there is no economic incentive for lying is not significant, differing by less than three percentage points. In contrast, at all positive economic incentive levels, the stronger message "I solemnly swear that the dot is _____" significantly increases the probability of sending a truthful message compared to the weaker message "The dot is _____", with the difference ranging from twenty seven percentage points for the $2 gain, to thirty one percentage points for the $1 gain to thirty four percentage points for the $3 gain. Thus, the stronger message has a fairly consistent effect, and does not change in a monotonic way as the economic incentive changes.

When we consider how truthfulness varies with the size of the economic gain to lying, unlike Gibson et al. (2013), we do not find that truthfulness monotonically decreases as the economic gain to lying increases. However, although Gibson et al. (2013) vary the economic incentive in four steps, the largest gain to lying in their experiment is approximately $1. Our results are thus consistent with their results because, like them, we find that in moving from a $0 gain to a $1 gain the level of truthfulness significantly declines. In fact, we find a very similar magnitude of the decline of truthfulness moving from $0 to $1. In the Gibson study, truthfulness declined from 82 to 21.1 percent when moving from a gain to lying of zero to 1.2 Swiss francs which they report was just over one dollar at the prevailing exchange rate at the time of their experiment. In our experiment, truthfulness in the weak message sessions declined from 93.2 to 32.4 percent when moving from a gain to lying of zero to one dollar. Further, we also find a significant decline in the level of truthfulness when moving from a $1 gain to a $2 gain. However, as we move to a higher economic incentive, specifically from a $2 gain to a $3 gain, the level of truthfulness increases for both strong and weak messages, though the difference is not statistically significant. Table 3 shows differences in truthful proportions and z-statistics for the various economic incentive levels for both strong and weak messages. The reduction in truthfulness that occurs when comparing a zero economic gain to lying to a positive economic gain is highly significant for all six treatment groups (within subject group comparison). The reduction in truthfulness between subject groups that occurs when the economic gain increases from $1 to $2 is also significant, at the 1% level for strong messages and at the 5% level for weak messages. In contrast, he difference in truthfulness between the $2 gain and $3 gain and between the $1 gain and $3 gain is not significant for either weak or strong messages.

In order to analyze our results further, we utilize and expand on the theoretical specification developed in Gibson et al. (2013). They posit a utility function for an individual:

$$V_i(T) = \begin{cases} b(m - ECOST) & \text{if } T = 1 \\ bm - C_i & \text{if } T = 0, \end{cases}$$

where $V_i(1)$ is the utility the individual attains when telling the truth, $V_i(0)$ is the utility the individual attains when he/she lies, m is the payment received when lying, ECOST is the monetary loss experienced when telling the truth, and C_i is the total cost of lying. Given the relatively low level of the stakes involved, the marginal utility of wealth is assumed to be a constant, denoted by b. The difference between the utility from telling the truth and from lying is thus given by:

$$Y_i^* = C_i - bECOST$$

An individual will tell the truth if $Y_i^* > 0$. The C_i term, called the total cost of lying in the Gibson et al. paper, is proposed to consist of an intrinsic cost of lying and the external cost of truth-telling. Their model incorporates the effect of the external cost of truth-telling linearly, so that it is not possible to distinguish between two opposing effects; 1) the negative direct incentive effect of the external cost of truth-telling on truth-telling and 2) its positive effect on the total cost of lying and, therefore, its positive effect on truth-telling. However, their model also allows for the external cost of truth-telling effect to enter non-separably with the intrinsic cost of lying. Since the level of truthfulness shows a declining marginal decrease as the economic incentive to lie increases in both the Gibson et al. data (Table 2 in Gibson et al. (2013)) and our data, we suggest that incorporating the square of the economic cost can better pick up the impact of the two separate and opposing effects of the economic cost. As the economic cost of truthfulness, or the economic gain from lying, increases, we hypothesize that it acts as a "size of lie effect" as in Lundquist et al. (2009), who find that the larger the lie needed to reach a desired outcome, the less people will lie. Here, the size of the lie itself does not change, as it consists of simply stating the color of the dot, but an increasing economic gain achieved by lying can be considered to increase the salience of the lie, which may trigger a stronger commitment to the truth. Therefore, the size of the gain to lying could actually have an increasing impact on truthfulness through its effect on the intrinsic cost of lying, working in the opposite direction to its economic incentive effect. This would result in the observed non-linear relationship between truthfulness (or lying) and the economic gain to lying.

Another important trigger of commitment to the truth that we incorporate in our experiments is the strength of the message, which, as seen in Tables 3.1 and 3.2, has a strong effect on truthfulness. Thus we posit an intrinsic cost of lying function $C_i = C(S, ECost^2, z_i,$ interaction terms), where S is a dummy variable indicating strength of message and z_i are individual characteristics which might affect the intrinsic cost of lying. In our experiment, the only individual characteristic variable we have data on is gender. Potentially, the strength of the message and gender could also interact with the economic cost variable.

Based on the above discussion, we examine two possible specifications of the total cost of lying:

a. $C = a_o + a_1 S + a_2 ECost^2 + a_3 Gender$

b. $C = a_o + a_1 S + a_2 ECost^2 + a_3 Gender + a_4 S*Ecost^2 + a_5 Gender*Ecost^2$

implying an empirical specification for a. of:

$$Y^*_{ij} = Ci - bECostj + \varepsilon_{ij} = \beta_0 - bECost_j + \beta_1 S + \beta_2 ECost_j^2 + \beta_3 Gender + \varepsilon_{ij}$$

and for b.:

$$Y^*_{ij} = Ci - bECostj + \varepsilon_{ij} = \beta_0 - bECost_j + \beta_1 S + \beta_2 ECost_j^2 + \beta_3 Gender + \beta_4 S*ECost^2 + \beta_5 Gender*ECost^2 + \varepsilon_{ij}$$

In the regressions of the simpler specification, reported in Columns (1) and (2) of Table 3.4, the economic cost, economic cost squared and strength of message variables are all highly significant in the expected direction. The coefficient on the gender variable indicates males are less truthful than females but the effect is not statistically significant. When the interaction terms are included, the same variables remain significant, while the interaction terms are not significant. That the strength of message interaction term is not significant is not too surprising, given the uniformity of the impact of the strong message in the raw data results in Table 3.1 above, where we see that the effect of the strong message is to increase truthfulness by approximately thirty percentage points at all positive levels of economic gain to lying. When the gender interaction term is included, its coefficient is negative, while the coefficient on the direct gender term becomes positive, suggesting that males are more truthful when the economic cost is low, but suffer less increase in their intrinsic cost of lying as the size of the economic cost increases. However, the coefficients are not statistically significant. Several previous studies such as Dreber and Johannsen (2008), Erat and Gneezy (2012), Houser et al. (2011) and Gibson et al. (2013) all find that males are more likely to lie when it results in a higher monetary payout to themselves. In contrast, Childs (2012) finds no difference in lying behavior by gender in a similar sender-receiver game as used by Dreber and Johannsen (2008), but with higher stakes, and others, including Aoki et al. (2010), Cappelen et. al. (2013) and Gylfason et. al. (2013) also find no significant difference by gender in various contexts. Thus caution seems warranted in making conclusions about lying based on gender**.**

Since the economic cost squared term in the regression could simply be picking up curvature in the distribution of lying aversion stemming from heterogeneous lying costs between individuals, rather than the intrinsic cost of lying varying with the gain to lying, we also calculated the total marginal effect of ECost on truthfulness at the different positive ECost levels in our data, of \$1, \$2 and \$3[21], based on the estimated coefficients in Column (2) of Table 3.4. These are reported in Table 3.4b, along with p-values from linear restrictions testing the null hypothesis that the marginal effect is zero. We find that the marginal effect of the economic gain to lying on truthfulness becomes significantly positive at ECost=3, which would not arise from a heterogeneous (between individuals) but fixed (within individual) intrinsic lying cost, and thus lends support to the proposition that intrinsic lying costs do increase as the economic gain to lying increases.

[21] ECost is \$1 when a participant could receive \$2 for lying and \$1 for telling the truth, \$2 when the payment is \$3 for lying and \$1 for telling the truth, and \$3 when the payment is \$4 for lying and \$1 for telling the truth.

As a further robustness check of our modification of the Gibson et al. (2013) model, we repeat the regressions in their paper, including the square of the economic gain to lying, both with and without the interaction terms of their intrinsic cost of lying measures with the economic gain to lying. Table 3.5 reports the results, and for comparison includes a replication of their Columns 3 and 4 from their Table 3[22]. Their main results from these regressions are that Protected Values (a measure of non-consequentialist moral values) and the interaction of Protected Values with the economic gain to lying are positively significant in determining the level of truthfulness, whereas their other two candidate measures of the intrinsic cost of lying, impression management (ExtDeceit) and self-image (SelfDeceit), are not significant. They include a measure of altruism (35Hurts) which is significant in their regression when they do not include the interaction terms, but the significance goes away when the interaction terms are included. Additionally, they find that women are more truthful than men, but the gender effect becomes insignificant in their regressions that include the values measures. When we rerun their regressions with the additional term of the square of the economic gain to lying, the protected values variables, both separately and interacted with the economic gain to lying, remain significant, and the ECostSquared variable is also highly significant. This suggests that the curvature of the effect of the economic gain to lying persists even when the values interaction terms are included, and so including the ECostSquared term can better capture the countervailing effect of the economic cost of truthfulness, through its size of lie effect, on the intrinsic cost of lying.

Utilizing a simple experiment, consisting of sending a message about the color of a dot, we find that the strength of the message being sent is a very significant determinant of truth-telling. This is consistent with earlier findings, such as Lundquist et al. (2009), though their results did not quite attain statistical significance for the difference in effect between experimenter-generated strong and weak messages. We also find that the size of the economic gain when telling a lie reduces truth-telling over some range, consistent with findings by Gneezy (2005) and Gibson et al. (2013). However, unlike these papers, we find that the decrease in truth-telling is not monotonic as the economic gain from lying increases. This holds true for both strong and weak messages. Modifying Gibson et al.'s theoretical framework on the cost of lying to incorporate both the changing economic incentives and the strength of the message, we find that the non-monotonicity is consistent with an increasing effect of the economic gain from lying on the cost of lying, resulting in more truthfulness, and thus acting as a countervailing force to its direct economic incentive effect in reducing truthfulness. However, we did not find a significant interaction effect between the strength of the message and the size of the economic gain from lying. The stronger message had a quite consistent effect on truthfulness of close to thirty percentage points at all positive economic incentive levels.

[22] The coefficients and standard errors for the constant and the ECost variable are slightly different than in the Gibson et al. paper, but the others are identical.

3.2.B. Prior Experience Experiment.

3.2.B.1. Introduction

In a separate set of experiments we consider the question of whether this pure form of lying aversion is affected by prior experiences of being lied to or told the truth. We are all involved in many economic and social interactions every day, which may involve being told the truth or being lied to. Do consecutive dishonest interactions have a cascading effect, decreasing the general level of honesty in society? Or can being lied to sometimes have a correcting effect on our own level of honesty?

Innes and Mitra (2013) found that lying responds to social cues on how often others lie. In their experiment, students randomly drew five slips of paper with information about the honesty of participants in previous sessions of the experiment. In sessions in both the US and India, students lied more when they drew more slips indicating other students were untruthful. Here, we focus on the direct experience of knowing one has been lied to, rather than knowledge of how many others lie in general. To investigate this issue, we first had our participants play the Receiver role in a Sender-Receiver game modeled on Gneezy (2005), in which the interests of the Senders and Receivers are not aligned. The participants were students in various Economics and Sociology classes at UC Merced and CSU East Bay. Senders in this experiment were informed of two possible payouts, one of which would be chosen by their paired Receiver. In our game, the payoff options, following Innes and Mitra (2013) were as follows:

Option C: $6 to the Sender and $3 to the Receiver.
Option D: $4 to the Sender and $6 to the Receiver.

The Sender had to choose one of two Messages to deliver to the Receiver, one *truthful* (Message D) and the other *untruthful* (Message C). The two possible Messages were:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

After receiving a message from their paired Sender, the Receiver had to decide whether to follow the Sender's message and choose the recommended option, or reject the recommendation and choose the other option. The only information the Receivers had about the payouts was the message sent by the Sender, and the statement that one of the options was better for the Receiver, and the other option was better for the Sender. After making their decision on whether or not to follow the Sender's message, and turning in their questionnaires for Experiment 1, Receivers were randomly assigned to three Treatment groups. The first group was the set of Control subjects who were exposed only to common information about the Sender-Receiver game – that is, information that was given to all Receivers. The purpose of this information was to control for subject beliefs about behavior in the Sender-Receiver experiment. The specific information given to all subjects after their experiment 1 questionnaire was turned in, was as follows:

"In Experiment 1, roughly 5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED."

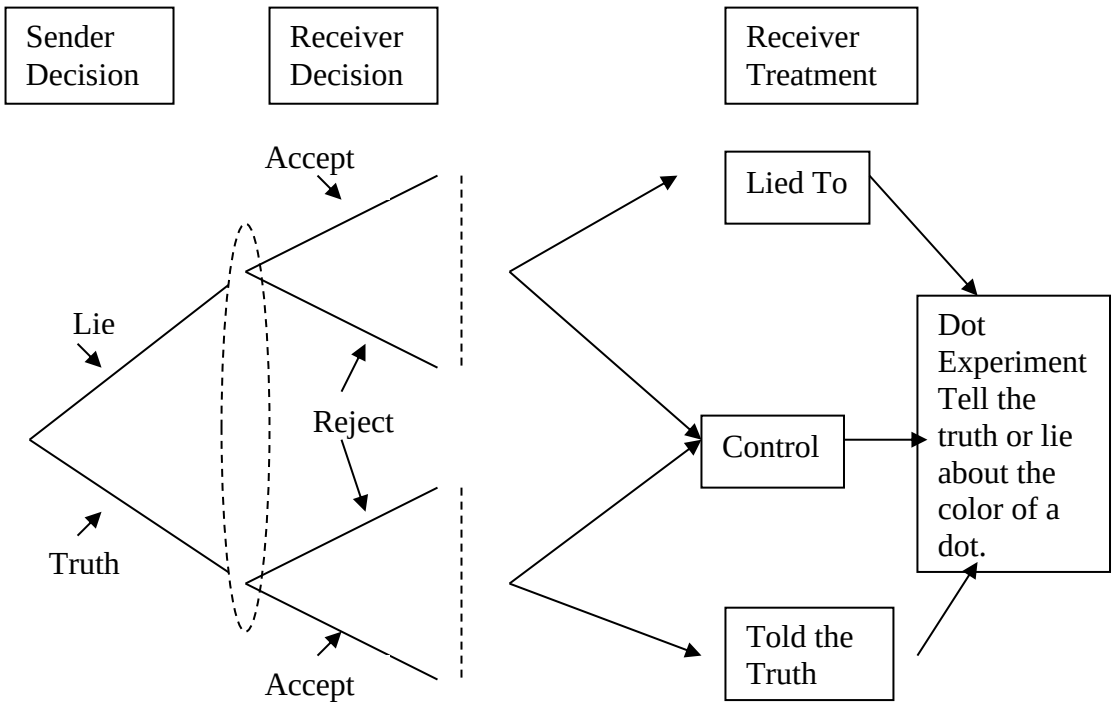
The statement reports the approximate percentage of truthful subjects from the Sender side of our experiment. The precise percentage of truthful Senders was 47.8 percent.

Each subject in the second and third Treatment groups was told whether his or her own matched Sender *lied* (Treatment LT) or *told the truth* (Treatment TT) in the Message that was sent. Next, the participants were alerted to a colored dot at the top of the page in the questionnaire, and asked to choose a message to send about the dot. The possible messages were "The dot is blue" and "The dot is green". Again, we randomized the actual dot color across questionnaires, and asked participants to write down the color of the dot, to check for color-blindness. Table 6 shows the mean proportions and standard errors of those telling the truth when reporting the color of the dot in each of two payout scenarios, both for the total sample, and divided by experience in the first experiment. In the first scenario, the participant receives the same amount ($1) whether he/she tells the truth or lies. In the second case, the payout is $2 if the participant lies and $1 if he/she tells the truth. The table also shows the difference statistics for the Lied To versus Control, Told the Truth versus Control, and Lied To Versus Told The Truth.

Figure 1 summarizes the structure of this two-part experiment.

.

Figure 3.1 Gneezy Game, Receiver Treatment and Dot Experiment



3.2.B.2 Logistics

The experiment was conducted in Economics and Sociology classes at the University of California, Merced, and Cal State East Bay. All three Treatments were conducted in all classes, resulting in a sample of 204 subjects. Sixty subjects were exposed to the Control

Treatment; 72 subjects were exposed to the "Lied To" Treatment; and 72 subjects were exposed to the "Told the Truth" Treatment. Questionnaires were identified by ID numbers, with tags attached that had matching ID numbers. Students retained these tags and used them to collect their payments at the beginning of the following week's class. The experimental instructions are attached in Appendix B.

3.2.B.3. Results

For the first scenario, when there was no difference in payment, over 94% of the total group told the truth about the color of the dot. This confirms that most people tell the truth if there is no reason not to, indicating an intrinsic preference for honesty. The rate of truth-telling was less for both those who had found out they had been lied to and those who had found out they had been told the truth relative to the controls, but the differences are not significant. For the second scenario when there was a monetary incentive to lie, almost 60% of the control group still told the truth about the color of the dot. We found that again those who were told whether they had been lied to or told the truth in the earlier experiment were less truthful in this experiment than the controls, with 35% of the Lied To group and 44% of the Told the Truth group telling the truth. The difference statistic for the Lied To group versus the Control group is significant at the 1% level and that for the Told the Truth group versus the Control group is significant at the 10% level. Thus we see that people's responses to incentives even in the case when the only consequence of the action is for their own payout are indeed affected by knowledge of prior experience.

Table 3.7 gives the results of a Probit Regression where the dependent variable is a zero-one variable, with one representing telling the truth when the payout is \$2 for lying and \$1 for telling the truth, regressed on the treatment variables, gender and course dummy variables. This regression confirms that both treatments, being lied to and told the truth, reduce truthfulness in the dot experiment, which is a rather surprising result. The decomposition below will disentangle several confounding factors contributing to this result. We also find that males are more truthful than females in this experiment, but not significantly so. Unlike Frank et al. (1993), we do not find that students studying economics are significantly less honest than those studying other subjects.[23]

Since, in the prior experiment, the experience of each participant depends both on whether or not they were lied to or told the truth and whether they accepted or rejected the sender's recommendation, their decision to lie in the dot experiment may be influenced by whether or not they received a low payment in the first experiment. In addition, there may be inherent differences in people who reject versus those who accept others' recommendations. Table 3.8 shows the proportion telling the truth for the six subgroups: Control Accepters, Control Rejecters, Lied To Accepters, Lied To Rejecters, Told the Truth Accepters and Told the Truth Rejecters. Table 3.9 presents difference-in-difference statistics, which, while controlling for baseline differences between Accepters and Rejecters, show Lied To effects for those receiving lower and higher payoffs (labeled Burned and Not-Burned, respectively), and the effect of receiving a lower payoff for those who were Lied To and those who were Told the Truth.

[23] However, we do not have data on subjects' majors, only on the class in which they participated in the experiment.

We find that among those who received lower payouts in the first experiment, those who were lied to in that first experiment are significantly more likely to lie in the dot experiment than those who were told the truth. In contrast, those who were lied to in the first experiment, but because they rejected the recommendation came out with the better payoff, are significantly more truthful in the dot experiment than those who were told the truth and accepted the recommendation, when controlling for the underlying differences between accepters and rejecters. This is somewhat surprising, since if someone was told the truth and, by believing the message, they came out with a good payoff, one might think that they would in turn tend to be more truthful. However, perhaps for those who had a lucky escape, in the sense that their rejection of the untruthful recommendation gave them the better payoff, the salience of the lie committed by someone else reinforced their aversion to lying even more strongly than the experience of truthfulness did for those who were told the truth and accepted. This is consistent with results in Chapter 2 where subjects who had similarly been lied to but who had rejected the Sender's recommendation, and subsequently played a trust game, were less trusting and less likely to communicate, but not less trustworthy or more deceitful, than those who received the same good payoff through acceptance of a truthful recommendation.

For those who were told the truth in the first experiment, the effect of being burned, (that is, because they rejected the truthful recommendation rather than accepting it), is to increase truthfulness in the dot experiment. Again, it appears that the salience of truthfulness of the Sender's statement is highlighted by the active decision of the Receiver to reject the statement which resulted in causing themselves harm, which in turn encourages the Receiver to be truthful in the dot experiment. Finally, for those who were lied to in the earlier experiment, the burned effect acts to reduce truthfulness in the dot experiment. Again, these results are all controlling for baseline differences in behavior between those who accepted their partner's recommendation in the first experiment and those who rejected it. The significant difference in truth-telling between Control Accepters (69%) and Control Rejecters (43%) suggests that there are indeed differences in this underlying behavior. Perhaps it is not surprising that the effect of being lied to or told the truth on subsequent truthfulness also seems to work differently for those who accepted or rejected their Sender's recommendation in the first experiment, in the sense that rejecting the Sender's recommendation seems to increase the salience of the lie or truth-telling of the Sender in the Receiver's decision on whether or not to lie about the color of the dot. In both cases, this leads to increased truth-telling by Rejecters relative to Control Rejecters (those who are not informed whether they were lied to or told the truth in the earlier experiment). In contrast, knowledge of either having been lied to or having been told the truth leads to a lower level of truthfulness for Accepters, relative to Control Accepters. Recent research in moral psychology, such as Greene and Paxton (2004) and Greene et al. (2009), emphasizes that moral decision-making reflects a tension between competing sub-systems in the brain, namely cognitive and emotional processes. It is possible that these processes are active to a different relative extent for those who have rejected versus accepted the Sender's recommendation in the first experiment, and that this contributes to different decisions of truthfulness in response to being lied to or told the truth, relative to the Control Accepters and Rejecters.

Difference-in-Difference statistics based on OLS regressions also controlling for course effects, gender, initial mood and mood changes are reported in Table 3.10, and are

very similar in magnitude and significance to the raw Difference-in-Difference statistics in Table 3.9.

The raw data given in Table 3.6 indicate that being lied to or told the truth in the prior experiment are both associated with significantly lower truth-telling in the dot experiment, relative to the controls who were not informed whether or not the Sender lied in the first experiment. However, breaking the results down by whether the person received a lower or higher payout in the first experiment as a result of being lied to or told the truth, as well as controlling for any inherent differences in behavior between people who tend to follow others' recommendations versus those who reject, gives a more nuanced story of behavior in response to finding out whether one has been lied to or told the truth.

3.3 Conclusion

While much research on lying aversion has understandably focused on situations where someone is directly affected by the lie, untangling lying aversion per se from other motivations such as guilt aversion and altruism is difficult in these settings. In the present paper, we contribute to the newer strand of literature that looks at lying aversion in its purest form, when no-one is directly adversely affected by the lie. While this method is primarily utilized in order to isolate the effect of lying aversion from other motivations, the situation of no-one being immediately negatively affected arguably reflects situations where we are faced with lying or telling the truth in interactions with large bureaucracies such as the government or large corporations.

In two distinct sets of experiments, we examine several factors affecting pure lying aversion. In the first set of experiments, we find that making a stronger statement significantly promotes truthfulness in a very consistent way across economic incentive levels, while changing the economic incentive has a non-linear effect on truthfulness. This is consistent with an intrinsic cost of lying that is increasing in the size of the economic incentive, counteracting the negative direct effect of the economic incentive on truthfulness.

In the second set of experiments, we look at the effect on pure lying aversion of the experience of being on the receiving end of a lie or a truth in a prior interaction. In this context, the interplay of the material outcome of the first experience and the contribution of the participant's own decision as Receiver in determining that outcome, play a role in the effect of being lied to on subsequent honesty. Specifically, for those who ended up with the lower payoff in the first experiment, getting that payoff through being lied to increased lying in the dot experiment, while for those who received the higher payoff in the earlier interaction, being lied to and having rejected that recommendation actually decreased subsequent lying relative to those who received the same payoff by accepting a truthful recommendation. Thus, being lied to, in some situations appears to encourage truthfulness in subsequent interactions, while promoting further lying in others.

Table 3.1
Proportion of Participants Telling the Truth by Strength of Statement and Size of Gain for Lying

Gain for untruthful statement vs. truthful statement	Strength of Statement	
	"The dot is _____."	"I solemnly swear that the dot is_____."
$0 (All subjects)	0.907 n=161	0.930 n=141
$1	0.324 n=74	0.633 n=60
$2	0.136 n=44	0.415 n=41
$3	0.209 n=43	0.550 n=40

Table 3.2
Difference (z) Statistics for Effect of Strong versus Weak Messages on Truthfulness for Different Levels of Economic Incentive to Lie

Economic Incentive to Lie	Difference in Truthful Proportions for Strong versus Weak Message (z-statistics)
$0 Gain	0.023 (0.726)
$1 Gain	0.309 (3.739)***
$2 Gain	0.270 (2.926)***
$3 Gain	0.341 (3.407)***

Table 3.3
Difference in Truthfulness and z-Statistics by Economic Gain to Lying and Type of Message

Comparison of Truthfulness across size of economic gain from lying		Weak Message	Strong Message
Within Subject Group Comparisons	\$0 vs \$1 Gain	0.608 (9.848)*** (n=74)	0.300 (4.282)*** (n=60)
	\$0 vs \$2 Gain	0.778 (11.450)*** (n=44)	0.512 (5.957)*** (n=41)
	\$0 vs \$3 Gain	0.651 (7.991)*** (n=43)	0.375 (4.297)*** (n=40)
Between Subject Group Comparisons	\$1 vs \$2 Gain	0.188 (2.226)**	0.218 (2.648)***
	\$2 vs \$3 Gain	-0.076 (-0.903)	-0.136 (-1.245)
	\$1 vs \$3 Gain	0.115 (1.394)	0.083 (0.837)

, * denotes significance at 5%, 1% level, respectively.

Table 3.4
Determinants of Truthful Behavior

	(1) Model Without Interaction Terms Probit	(2) Model Without Interaction Terms OLS	(3) Expanded Model Probit	(4) Expanded Model OLS
Economic Cost of Telling the Truth	-1.922 (-2.723)***	-0.642 (-2.865)***	-1.857 (-2.653)***	-0.636 (-2.839)***
Strength of Message (S)	0.886 (5.720)***	0.318 (6.002)***	0.775 (3.219)***	0.290 (3.329)***
Economic Cost Squared($ECost^2$)	0.441 (2.483)***	0.147 (2.610)***	0.446 (2.452)**	0.157 (2.743)***
Gender (Male = 1)	-0.081 (-0.529)	-0.026 (-0.498)	0.169 (0.708)	0.062 (0.717)
S*$ECost^2$			0.026 (0.546)	0.006 (0.346)
Gender *$ECost^2$			-0.063 (-1.342)	-0.021 (-1.332)
Constant	1.019 (1.735)*	0.826 (4.228)***	0.911 (1.529)	0.791 (3.974)***

*, **, *** denotes significance at 10%, 5%, 1% level, respectively

Table 3.4b.
Marginal Effect of ECost

ECost ($Gain to Lying)	1	2	3
Total Marginal Effect of ECost (=coeff(ECost)+2*ECost*Coeff($ECost^2$)	-0.348	-0.054	0.240
(p-value)	(0.002)***	(0.094)*	(0.047)**

*, **, *** denotes significance at 10%, 5%, 1% level, respectively

Table 3.5

Replication and Extension of the Results in Gibson et al. (2013) Table 3 (Logit Regressions with Dependent Variable = Proportion Truthful)

	(3)	(3) with ECostSquared	(4)	(4) with ECostSquared
Ecost (standard errors in parenthenses)	-2.749*** (0.197)	-6.22*** ((0.518)	−6.852*** (1.503)	-8.449*** (1.393)
EcostSquared		2.906*** (0.337)		2.458*** (0.327)
Gender (Female = 1)	0.304 (0.276)	0.313 (0.287)	0.297 (0.281)	0.306 (0.289)
Age (in years)	-0.027 (0,031)	-0.028 (0.032)	-0.027 (0.030)	-0.028 (0.031)
Psychology (=1)	0.026 (0.312)	0.034 (0.324)	0.047 (0.313)	0.046 (0.324)
Other Studies (=1)	0.242 (0.401)	0.261 (0.409)	0.272 (0.398)	0.276 (0.324)
35Hurts	0.216 (0.109)	0.224 (0.114)	0.122 (0.112)	0.159 (0.128)
ExtDeceit	0.505 (0.972)	0.522 (1.011)	-0.071 (0.873)	-0.014 (1.017)
SelfDeceit	0.019 (1.027)	-0.049 (1.067)	0.083 (0.973)	-0.012 (1.122)
Protected Values	0.735*** (0.139)	0.764*** (0.145)	0.361*** (0.135)	0.494*** (0.157)
ECost*35Hurts			0.178 (0.209)	0.111 (0.189)
ECost*ExtDeceit			1.182 (1.592)	1.006 (1.439)
ECost*SelfDeceit			-0.659 (1.741)	-0.384 (1.598)
ECost*PV			0.748*** (0.221)	0.488** (0.202)
Constant	-2.297* (1.23)	-1.821 (1.247)	-0.196 (1.176)	-0.364 (1.326)
Pseudo R^2	0.248	0.270	0.261	0.275
Pseudo Log Likelihood	-667.4	-648.0	-655.7	-643.2

*, **, *** denotes significance at 10%, 5%, 1% level, respectively.

Table 3.6
Proportions Reporting Truthfully and Differences in Proportions between Treatment Groups in Prior Experience Experiment

Economic Gain to Lying	All (n=204)	Control (n=60)	Lied To (n=72)	Told the Truth (n=72)
$0 Gain	0.946	0.966	0.931	0.944
$1 Gain	0.460	0.593	0.380	0.431
Difference in Proportions and z-Statistics				
	Lied To - Control	Told the Truth - Control	Lied To – Told the Truth	
$0 Gain	-0.025 (-1.127)	-0.022 (-1.110)	-0.013 (-0.023)	
$1 Gain	-0.187 (2.912)***	-0.162 (-1.777)*	-0.051 (-1.189)	

Table 3.7

Probit Regression of Proportion Telling the Truth when Economic Gain is $1, on Treatment Effects, Gender and Course Effects

	Coefficient (z-statistic)
Lied To	-0.537 (-2.356)**
Told Truth	-0.417 (-1.862) *
Male	0.223 (1.224)
ECONUCM	-0.141 (-0.590)
ECONCSU	0.299 (1.138)
Constant	0.122 (0.616)

Table 3.8
Proportions Telling the Truth By Accept/Reject Decision and Treatment Group

	$0 Gain From Lying	$1 Gain From Lying
Control Accept	0.972	0.694
Control Reject	0.957	0.435
Lied To Accept	0.930	0.326
Lied To Reject	0.931	0.464
Told the Truth Accept	0.936	0.383
Told the Truth Reject	0.960	0.520

Table 3.9
Difference-In-Difference Statistics

	Lied To Effect		"Burned" Effect	
	For the "Burned"	For the "Not Burned"	For the Told-Truth	For the Lied To
	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)	Diff-in-Diff (z-statistic)
	(1) (LTA-TTR) - (CA-CR)	(2) (LTR-TTA) - (CR-CA)	(3) (TTR-TTA) - (CR-CA)	(4) (LTA-LTR) - (CA-CR)
Tell the Truth in Dot Experiment when payoff was $1 if tell the truth, $2 if lie	-0.454 (-2.52)***	0.341 (1.97)**	0.397 (2.21)**	-0.398 (-2.36)**

Legend: CA = Control Accept, CR = Control Reject, LTA = Lied To Accept, LTR = Lied To Reject, TTA = Told the Truth Accept, TTR = Told the Truth Reject.

, * denotes significance at 5%, 1% level, respectively.

Table 3.10
Difference-in-Difference Statistics Controlling for Course Effects, Gender, Initial Mood, and Mood Change

	For the "Burned"	For the "Not Burned"	For the Told-Truth	For the Lied To
	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)	Diff-in-Diff (p-value)
	(1) (LTA-TTR) - (CA-CR)	(2) (LTR-TTA) - (CR-CA)	(3) (TTR-TTA) - (CR-CA)	(4) (LTA-LTR) - (CA-CR)
Tell the Truth in Dot Experiment when payoff was $1 if tell the truth, $2 if lie	-0.419 (0.0243)**	0.379 (0.0368)**	0.408 (0.0380)**	-0.391 (0.0257)**

Legend: CA = Control Accept, CR = Control Reject, LTA = Lied To Accept, LTR = Lied To Reject, TTA = Told the Truth Accept, TTR = Told the Truth Reject. Difference-in-difference (and difference) statistics are obtained from OLS Regressions that include treatment dummies, course effects, gender, initial mood, positive mood change, and negative mood change. Robust p-values are presented for test statistics on linear restrictions of zero difference-in-difference . ** denotes significance at 5% level.

Chapter 4
Lying Through Others[24]

4.1. **Introduction**

A recent segment of "60 Minutes" concerned the US flooring company Lumber Liquidator's outsourcing of production of laminated flooring to companies in China. Apparently, these companies were stamping the laminate as complying with California's formaldehyde regulations, even though it did not in fact comply. This allowed the Chinese companies, and thus Lumber Liquidator, to keep costs low by using cheaper glues that contain high levels of formaldehyde, a known carcinogen. This report highlights a potential benefit of outsourcing. By delegating production to companies in countries with less stringent labor, environmental or safety regulations and inspections, Western corporations may not only benefit from lower labor costs, their ostensible reason for contracting with producers there, but can also separate themselves physically, psychologically and morally from questionable practices which contribute to lower costs, and thus increase their profits further. Incidents with tragic consequences abound, where workers in factories producing for Western corporations have died because of shoddy construction or inadequate fire escape facilities. Consumers in Western countries may be exposed to unsafe products such as in the Lumber Liquidator case, with formaldehyde present in the laminate at levels likely to cause cancer. Such behavior on the corporate level gives rise to the question of whether individuals are more willing to employ others to lie on their behalf than they would be willing to do directly. This paper uses several experiments to study this question. We detect a significant effect of delegation on lying aversion; that is, we find significantly more people will delegate to someone they think is likely to lie than will choose to lie directly themselves.

A considerable body of recent research has documented that a substantial proportion of people are unwilling to lie to achieve a superior monetary outcome for themselves, both when that lie involves another person and affects that other person's payoff as in Gneezy (2005) (and the large number of papers stemming from that research in various settings), and even when another is not directly hurt by the lie, as in Lopez-Perez and Spiegelmann (2013) and Gibson et al. (2013). Gneezy (2005) found that the preference for truthfulness was affected by the size of the gain to the Sender and the cost to the Receiver. Other factors have also been found to affect the propensity for truthfulness, such as the size of the lie (Lundquist et al. 2009), the strength of the message (Lundquist et al. (2009), and individual's value systems (Gibson et al. (2013). The question we focus on here is, if people prefer a higher material outcome for themselves, but incur a cost if they have to lie in order to secure that higher payoff, is the cost as high if they assign someone else to tell the lie that secures them the higher payoff? In order words, is lying aversion lower when the lying is delegated?

[24] This essay is based on joint work with Robert Innes, UC Merced.

The literature on delegation has traditionally focused on how to design an efficient mechanism to obtain optimal effort on the part of an agent (e.g. Bolton and Dewatripont 2005). More recently, however, other aspects of delegation have been emphasized, including the incentive effects of giving agents more control (Charness et al. 2012), and the use of delegation to avoid punishment (Coffman 2011, Bartling and Fischbacher 2012, Oexl and Grossman (2013)). In these latter papers, the authors utilize dictator games with punishment to study how the Receivers (those who cannot make a decision on the allocation) attribute responsibility to those who do make that decision both directly and by delegating the decision to a second person. They find that responsibility is shifted in the delegation treatments, as more punishment is meted out to the person who directly made the decision, that is, the delegate in the delegation treatment, rather than to the principal who chose to delegate the allocation decision. In Bartling and Fischbacher (2012), for example, because Senders anticipate this effect, there is three times as much delegation in the punishment as in the non-punishment treatments. The authors find that the punishment behavior is stable over time in their repeated game treatment, while the Senders, learning that Receivers do not increase their punishment of the dictators over time, delegate more as time goes on. Another treatment, which allows for the unfair allocation only when the Sender delegates, also finds that the blame is shifted to the delegate. There is also some, though less, blame shifting when the delegation is to a random device rather than to another player.

In the Bartling and Fischbacher (2012) study, delegating involved giving up control over the outcome, as the delegate was free to make any decision. However, they report that enough delegates chose the unfair allocation that it was the payoff maximizing choice for dictators. In another important paper in this literature, Hamman et al (2010) look at the behavior of dictators when they can choose which agent to delegate to. They find that dictators switch from more generous to less generous agents. They attribute this delegating behavior to responsibility diffusion. Even in the absence of punishment, an extrinsic motivation, as in Bartling and Fischbacher (2012), dictators feel less personally responsible when they delegate than when they make the allocation directly, even though in choosing a less generous agent they are in fact choosing a less generous allocation.

While all the above experimental literature on delegation utilizes the dictator game, in a paper more directly related to the present study, Erat (2013) looks at the delegation of deception, an inherently immoral action, which may have different implications for responsibility shifting or diffusion. He conducts an experiment modeled on the Erat and Gneezy (2012) design, where a person sends a message about the outcome of the roll of a die to a Receiver who must decide what number they think was actually rolled. In this case, the Sender can either send the message him/herself, or delegate to an agent to undertake the task. The study utilizes 2 treatments. In both, the A option gives $10 to each of the three participants in the group, and the B option gives identical payments to the Sender and Agent of $15, but in treatment T[-2], the receiver gets $8, and in treatment T[-6] the receiver gets $4. The paper has three main results: 1) 30% of Senders overall delegate. Since participants believe that no more than 30% of Agents will tell the truth, and this is not significantly different for those who delegate (35%) and those who don't delegate (30%), those delegating are doing so with the expectation that the agent will most likely lie. 2) Senders are more likely to delegate when the harm to the Receiver is greater. The delegation breaks down to 25% in

treatment T[-2] and 34% in treatment T[-6]. 3) Women are more likely to delegate than men.

Erat argues that this gender effect is consistent with previous literature that finds women less likely to lie when the hurt to others is greater, such as Dreber and Johannsen (2008). However, he also acknowledges that women delegating more could be a gender effect on willingness to make a moral decision that affects others, rather than a difference in lying aversion. This is a general feature in the design, where delegation is optional. A lot of the delegation may be picking up unwillingness to make a decision, akin to an exit choice in a dictator game (Dana et al., 2006; Broberg et al., 2007; Lazear et al., 2012), rather than the effect on lying aversion per se. Additionally, with no control treatments, it is impossible to tell who is choosing to delegate – those who would otherwise choose to lie or those who would otherwise tell the truth. In order to fully isolate the effect of delegation on lying aversion, it is necessary to distinguish between four different factors; the disutility from lying itself, the change in this disutility due to delegation, the baseline level of caring about others' outcomes, and the change in caring about others' outcomes due to delegation.

In order to address these issues, in the present paper we utilize a design where everyone has to delegate. The choice is in who to delegate to, someone more likely to lie or someone more likely to tell the truth. Comparing this to playing the deception game without delegation creates a clean comparison on the effect of delegation on lying aversion. In addition, we also conduct direct and delegated dictator games to control for effects on preferences over allocations. We find significantly more people will delegate to someone they think is likely to lie than will choose to lie directly themselves, and this effect is more pronounced in the deception game than in the dictator game, suggesting that it cannot be attributed to the effect of delegation on preferences over allocations.

We construct a very specific test in this paper in order to pinpoint a preference-driven effect of delegation. Delegation can affect the decision environment in two inter-related ways: it can reduce both (1) control over a decision and outcomes; and (2) attribution of responsibility for a decision. Often these two are considered one and the same. For example, philosophers argue that individuals are responsible for outcomes only if they can control them (Nelkin, 2004; Gurdal et al., 2014). Bartling and Fischbacher (2012) construct a measure of responsibility driven by the extent to which a principal is perceived to be able to affect the outcome. Indeed, existing literatures on delegation and other mechanisms for attenuation of social preferences mostly embed reductions in the principal's control over outcomes. For example, in Erat's (2013) key paper, delegation cedes control over the decision. A distinct literature on responsibility alleviation (Charness, 2000) and hidden costs of control (Falk and Kosfeld, 2006) documents how reducing an agent's control can elevate moral hazard (Charness et al., 2012). Similarly, when negative outcomes for a matched player can be due to either nature or a dictator's decision, thus reducing the probability that the dictator's decision is implemented, dictators tend to "hide behind nature" and act more selfishly (Dana et al., 2007; Andreoni and Bernheim, 2009). An interesting paper by Haisley and Weber (2010) shows that subjects may also "hide behind ambiguity"; adding ambiguity to probability distributions of outcomes , even when ultimate probabilities are the same, leads to more selfish decisions . While ambiguity does not reduce control objectively, it may be perceived as doing so. Decisions in groups, or based on team incentives, can also make

subjects more selfish (Charness and Sutter, 2012) and more willing to lie (Conrads et al., 2013; Cadsby et al., 2010; Muehlheusser et al., 2015). Other recent research shows that, if subjects can be "willfully ignorant" by opting not to know about the adverse consequences of increasing their own payoff on the payoff of someone else, then they tend to be more selfish (Dana et al., 2007) and less subject to punishment (Bartling et al., 2014). In contrast, Grossman (2015) finds that willful ignorance declines to almost zero when subjects must deliberately choose to be ignorant. While willful ignorance does not cede control per se, it does cede knowledge of consequences.

While a reduction in control surely implies a reduction in responsibility, delegation can affect responsibility attribution even when it has no effect on control. For example, Oexl and Grossman (2013) find that delegation reduces the extent to which dictators are blamed and punished for selfish allocations, even when no control is relinquished by the delegation. However, there is also scope for *internal* responsibility attribution – assignment of responsibility to oneself for purposes of evaluating moral trade-offs. This is the central focus of this paper: How does delegation affect lying aversion when there is no loss of control over decisions and no scope for external responsibility attribution or punishment?[25] Figure 1 frames this focus by decomposing responsibility attribution along two dimensions, "Attributor" (external vs. internal) and "Driver" (knowledge of consequences, control over decisions and outcomes, and intermediation/delegation).

[25] In this respect, our study is similar to Drugov et al. (2014) where the intermediary is transparent and does not make any decision.

Figure 4.1 Responsibility Attribution: A Decomposition

		ATTRIBUTOR		
		External (Do others hold you responsible?)		***Internal*** (Do you hold yourself responsible?)
DRIVER	***Knowledge of Consequences*** (Less knowledge → less responsible)	Willful ignorance[A]	↔	Willful ignorance
	Transparency of Decision (Less transparent → less responsible)	Hiding behind nature[B] ↕	↔	Hiding behind nature ↕
	Control over Decision & Outcome (Less control → less responsible)	Hiding behind nature[B] Hidden cost of control[C] Shifting the blame[E] ↕	↔ ↔	Hiding behind nature Hidden cost of control Team/group incentives[D] Delegation[F] ↕
	Delegation (with No Loss of Control) (Delegation → less responsible)	Shifting the blame		Delegation **OUR FOCUS (re: Deception)**

[A] Dana et al. (2007), Bartling et al. (2014), Grossman (2015).
[B] Dana et al. (2007), Andreoni & Bernheim (2009), Bartling & Fischbacher (2012), Haisley & Weber (2010).
C Charness (2000), Charness et al. (2012), Falk and Kosfeld (2006).
[D] Conrad et al. (2013), Cadsby et al. (2010), Muehlheusser et al. (2015).
[E] Bartling & Fischbacher (2012), Coffman (2011), Oexl & Grossman (2013), Gudal et al. (2013).
[F] Hamman et al. (2010), Erat (2013), Drugov et al. (2014).

The remainder of the paper is as follows. Section 4.2 describes our first experiment. Several potential issues arise from the results of the first round experiments, as discussed below, so we modified the design and conducted further sessions of experiments. Section 4.3 follows with our second experiment, which represents a more controlled approach to determining delegation effects, and consists of two sets of treatments, one where the Receiver is harmed by a lie and the second where the lie benefits both the Sender and the Receiver. Section 4.4 concludes.

4.2 The First-Round Experiment

4.2 A. Direct Treatment.

Our deception experiment is modeled after the Gneezy (2005) design. In the direct treatment, participants called Senders each observed 2 possible payoffs to themselves and another person called the Receiver, who did not observe the payoffs. The Sender had to send a message to the Receiver about which option was better for the Receiver, and the Receiver, based only on that message, had to choose which option should be used for the payoffs. The Options we used are:

Option C: $5 to the Sender and $6 to the Receiver

Option D: $7 to the Sender and $3 to the Receiver.

The messages that the Sender could send were:

Message C: Option C is better for you (the Receiver) than Option D; or

Message D: Option D is better for you (the Receiver) than Option C.

Since one payoff was better for the Sender, and the other for the Receiver, if the Sender sent an untruthful message, this potentially harmed the Receiver. Thus following the Gneezy (2005) nomenclature, this can be classified as a Selfish Black Lie. In all cases, Senders were given general information on the propensity of Receivers to accept their recommendations. Based on results from Gneezy's (2005) experiments (where 78 percent of Receivers followed the Sender recommendations), we told all Senders the following:

> "In past experiments like this one, roughly 8 out of 10 Receivers chose the Option recommended by their Senders."

Receivers were not given this information, and Senders were so informed. To verify that Senders generally believed that Receivers would accept their recommendations, we asked them to predict their Receiver's choice and paid them $1 for a correct prediction. Overall, 72 percent of Senders predicted that their Receiver would choose the recommended option, indicating that Senders generally expect their recommendations to be followed; hence, Sender choices likely reflect a concern for the "fairness" / morality of lying, rather than strategic motives. As it turned out, 72.9 percent of our Receivers followed their Sender recommendations.

In order to control for preferences over allocations, subjects played both a Gneezy deception game and a parallel dictator game, with one of the games randomly selected for payment (each with 50 percent probability). This approach mimics Hurkens and Kartik (2009) (see also Innes and Mitra, 2013). When the dictator game was selected for payment, the chosen dictator allocation was implemented with 80 percent probability in order to replicate deception game payoffs (again as in Gneezy, 2005). Senders were informed of this procedure.

4.2. B. Delegation Treatment. In the Delegation treatment, the Sender was told that a message would be sent to a Receiver, but that he/she would not send the message him/herself. Instead, he/she would choose a Sender from one of two possible groups. These groups consisted of people sorted by their choice in a prior experiment (Experiment 0) which was explained to the Sender in the current experiment. In Experiment 0, participants observed a colored dot at the top of the page in their

questionnaire. They had to send a message about the color of the dot to someone, but the message did not affect anyone else's payment, only their own. The possible messages were "The dot is blue" and "The dot is green". Reporting the color of the dot truthfully earned the participant $1, while reporting untruthfully earned the participant $2. We labeled people who earned $1 in this experiment "$1 Senders" and those who earned $2, "$2 Senders". The participants in the Dot experiment also took part in a deception experiment direct treatment as described above. In the delegation treatment of the current experiment, we asked the Sender to choose the group, either $1 Senders or $2 Senders, from which they wished a student to be randomly selected to send the message on their behalf to their Receiver. The message sent on behalf of the Sender to his/her Receiver was the same message that the chosen delegated Sender sent to their own Receiver in their own session of the experiment. Thus selecting a $1 Sender was delegating to a relatively truthful person, while delegating to a $2 Sender was delegating to a less truthful person, in their choice in the Dot Experiment. We tested for understanding of this by asking the Senders "Which person earned more money by sending an untruthful message about the dot, a $1 sender or a $2 sender?" Over eighty three percent of respondents answered correctly. We also elicited the students' beliefs about probabilities of the different Sender types choosing the self-interested Message and dictator allocation, respectively. Each correct prediction, within the correct ten percentage point band, was rewarded with a $1 payment. Responses indicated general, but far from perfect understanding of differential behavior of the $1 and $2 Senders. Thirty percent of delegating subjects predicted a *lower* probability of a truthful choice by a $1 versus $2 Sender, contrary to expectations. For the dictator game, twenty-five percent predicted a lower probability of a generous choice by a $1 versus $2 Sender. These percentages are significantly lower than 50 percent, indicating broad success in conveying the nature of the delegation choice. As a robustness check on our results below, we consider a sample of delegators that excludes the anomalous predictors whose delegation decisions potentially do not reflect the intended correspondence between more and less truthful agents.[26]

4.2. C. Logistics. The experiment was conducted in upper division undergraduate economics classes at U.C. Merced. In total, there were 142 Sender/Receiver pairs, with 72 Senders in the direct / control treatment and 70 Senders in the delegated treatment. Receivers were in different classes than paired Senders. Participation was purely voluntary. Subjects were instructed to communicate only with the experimenter and were carefully monitored to this end. Control and delegated Sender treatments were implemented with equally mixed questionnaires in each of three sessions, randomly distributed to students.[27] Each session lasted approximately twenty five to thirty minutes. Questionnaires were identified by ID numbers, with tags attached that had

[26] We do not want to over-emphasize responses to the belief questions. Prior work in psychology and economics documents that subjects often make decisions that reflect a subconscious understanding of the situation even when they cannot explain this understanding (see, for example, the celebrated "red card/blue card" paper by Bechara et al., 1997, and a recent paper by Friedman et al., 2015 indicating that large numbers of repetitions can produce cooperation even though participant answers reveal that they do not understand the economic environment underpinning the interactions).

[27] Our experiments are conducted in classrooms with limited time and subject anonymity. As a result, we have only one observable variable to judge cross-treatment balance, namely, gender (Male). The proportions of male subjects in the direct and delegated samples (full and restricted) are, respectively, 40.0%, 36.23% (full delegated) and 35.42% (restricted delegated); differences are not significant.

matching ID numbers. Students retained these tags and used them to collect their payments at the beginning of the following week's class.

4.2. D. Results. Table 4.1 presents the results for the deception and dictator games. In the deception game there is a significant delegation effect in the expected direction. Over fifty eight percent of participants told the truth in the direct deception treatment, that is, when they had to send a message directly to their Receiver. In contrast, in the delegated treatment, when Senders were choosing which group they wanted their delegated Sender to be chosen from, less than thirty six percent of Senders chose to delegate to a Sender from the more truthful group. This difference is significant at the one percent level (2-sided test). In contrast, in the dictator game, there was no significant effect of delegation. A difference-in-difference z-statistic calculated across the treatments was found to be significant at just under the five percent level (z=1.919).

In order to reduce effects of participants misunderstanding the instructions in the delegation treatment, we repeated the analysis restricting the delegated group to include only those whose answers to the prediction questions indicated that they believed people in the $1 Sender group would be more truthful than those in the $2 Sender group. Table 4.2 shows the results of this analysis. The results are broadly similar, with just under thirty three percent of the delegated group of the restricted sample delegating to the more truthful Sender, so that the z-statistics are slightly more significant with the difference-in-difference statistic (z= 2.039) now significant at the five percent level (2-sided test).

Hurkens and Kartik (2009), in their critique of the Gneezy (2005) paper, point out that it is important to condition on preferences over allocations when making conclusions about lying aversion. That is, rather than simply comparing, as Gneezy does in his paper, the proportions who choose to lie when the gain to the Sender or the loss to the Receiver is varied, it is more valid to look specifically at the lying behavior of those who state a preference in the dictator game for the allocation that is potentially obtained when the Sender lies. The Gneezy experiment used a between subject design with regard to the deception and dictator games for each set of allocations, so it is not possible to examine the lying behavior of those who preferred the selfish outcome in the dictator game at an individual level, only at an aggregate level. In contrast, similarly to Hurkens and Kartik's own experiments, here we have Senders make both deception and dictator choices, so we can look at the lying behavior of those Senders who, in the dictator game, choose the selfish allocation. Table 4.3 presents the proportions of Senders who made truthful choices in the direct and delegated treatments. The z-statistic for the difference between the treatments is significant at the five percent level.

Table 4.4 presents regressions that control for course and gender effects.[28] Coefficients for the Delegate dummy indicate the effect of delegation in the deception game, first on the overall propensity for truthfulness (Model 1), second on the preference for truthfulness among the selfish (the Hurkens-Kartik Model 2), and third on the differential preference for truthfulness over generosity (the difference-in-difference of Models 3 and 4). In all cases, the difference results of Tables 4.1 and 4.3 are confirmed, with strikingly consistent parameter estimates. Overall, delegation is estimated to reduce truthfulness by roughly 22 to 25 percent.

[28] Coefficients are from OLS regressions with robust standard errors, indicating marginal effects of delegation. Similar results are obtained from qualitative dependent variable models such as Probit.

The first-round experiment produces two tentative conclusions. First, *delegation reduces lying aversion for a significant fraction of subjects*. Second, however, *delegation does not eliminate lying aversion*. The second conclusion is indicated by the Hurkens-Kartik statistics in Table 4.3. Among the delegating selfish, a positive proportion are truthful and this proportion is significantly different from zero in both delegation samples; corresponding z statistics (p values) are 3.76 (p=0.0006) and 2.51 (p=0.0186).

Potential limitations of first-round experiment. In the Hurkens and Kartik (2009) experiment, only four out of ninety participants lied in the deception game but were generous in the dictator game. Table 4.5 shows the breakdown by choice in the two games for the two treatments in our experiment. In contrast to the very small proportion of Generous and Liar pairings in Hurkens and Kartik (2009), we find overall 35 out of 142 participants chose that pairing, breaking down by treatment to fourteen out of seventy two (0.194) in the direct treatment and twenty one out of seventy (0.300) in the delegation treatment. The presupposition in Gneezy (2005) and Hurkens and Kartik (2009) is that the dictator game cleanly measures preferences over allocations, and thus can be used as a yardstick against which lying aversion can be measured. Thus, under the hypothesis that some people are lying averse, the proportion lying is expected to be less than the proportion choosing the selfish option. However, the significant fraction of subjects in our experiment choosing to be generous in the dictator game but lying in the deception game suggests that it is difficult to measure true preferences in the two games using a within-subject design. If so, the Hurkens and Kartik critique of Gneezy's paper, namely that it is more appropriate to use a within-subject design for the deception and dictator games, may not be valid.

To some extent in our delegation treatments, the occurrence of a large number of Lying and Generous choices may also be an indication that participants did not necessarily associate truthfulness with generosity, so that having to pick from groups based on their truthfulness in a prior experiment rather than their generosity may have not seemed very meaningful to at least some participants.[29] In order to address this concern, we redesigned the delegation choice in the dictator game in the next round of experiments reported below. However, this criticism cannot be applied to our direct treatment, in which almost twenty percent of participants also chose this pairing. One possible reason for this apparent conundrum is that what we are picking up is that, when subjects participate in multiple experiments, they compensate for less pro-social behavior in one game with more pro-social behavior in another. This is consistent with findings by Gneezy et al (2011), who found that participants who had lied in a sender-receiver game were more likely to contribute to charity than those who had told the truth, and with Ploner and Regner (2013), who found that subjects who had the opportunity to cheat in one game tended to be more generous in a subsequent dictator game. Thus, a within

[29] To some extent, responses to the belief questions suggest otherwise, with only 25 percent of delegating subjects indicating that $1 Senders are less likely to be generous than $2 Senders, and less than 20 percent indicating so in the restricted sample. However, subjects may nevertheless find it more difficult to align preferences over allocations with agent behavior in the dot experiment. For example, even among participants with beliefs consistent with more truthful behavior by $1 Senders (our restricted sample), mean predicted probabilities of selfish behavior is very similar for the $1 Senders (58 percent selfish) as for $2 Senders (66 percent selfish), while mean predicted probabilities of deceit in the Gneezy game are significantly different for the $1 Senders (45 percent) than for the $2 Senders (70 percent). Perhaps these difficulties compromise the use of the dictator game in controlling for preferences over allocations.

subject design may suffer from a lack of independence of behavior in the deception and dictator games, which might compromise the use of the dictator game behavior as the yardstick by which preferences over allocation are measured. If this is the case, then the between-subject design of Gneezy (2005) would seem to be a cleaner method of controlling for preferences over allocations using the dictator game, rather than the within-subject design used here and advocated by Hurkens and Kartik (2009), even though the latter design does have the advantage of being able to compare the behavior in the deception and dictator games at the individual level.

A further factor contributing to the surprising combination of Lying and Generous might be that in the dictator game, where the Receiver makes no decision, the Sender feels more responsibility for the outcome than in the deception game where the allocation is ultimately chosen by the Receiver. Thus, some or all participants may perceive even the direct deception game as involving some degree of delegation relative to the dictator game, a factor that might encourage lying even if one would make the generous choice in the dictator game. Additionally, some participants might have a negative cost of lying, that is, they prefer to lie when they can, and the utility they gain from lying in the deception game outweighs their preference for the generous allocation, resulting in the lying and generous pairing. The Hurkens and Kartik (2009) analysis assumes that the cost of lying is non-negative.

A final, and possibly the most important, concern with the design of our first round experiment concerns the probability with which an action is carried out. In the direct treatment, decisions made by subjects are implemented with 100 percent probability, while in the delegated treatment, choice of a $2 versus a $1 Sender leads to a different probability that the deceptive message is sent. For example, suppose that a subject believes that there is a 70 percent probability that $2 Senders are deceptive in the Gneezy game and a 40 percent probability that $1 Senders are deceptive. Then, by choosing a $2 Sender rather than a $1 Sender, the subject increases the probability of deception by 30 percentage points. In an expected utility framework in which a Sender's utility is a function of payoff allocations and whether a lie has been sent on behalf of the Sender, then all that matters for comparison between delegated and direct treatments is that the perceived probability of deception by $2 Senders is greater than by $1 Senders.[30]

In this case, the comparisons we have made above are clean tests of whether delegation affects lying aversion, assuming our dictator controls reflect preferences over

[30] Let U(A,L,D) = utility index with allocation A when lying or not (L=0 or 1) and when delegating or not (D=0 or 1), q_i = probability that a $i Sender chooses a deceptive message (i=1 or 2), A_j = allocation benefiting the Sender (j=S) or Receiver (j=R), p = probability of Receiver acceptance (e.g., 80%), and r_i = probability that a $i Sender chooses the selfish allocation A_S in the dictator game. Further, let individual utilities equal "base"/index utilities plus an idiosyncratic error, following common discrete choice logic. Then for D=1, deception is advantageous when (and only when) $p(q_2-q_1)(\Delta U(1)+\varepsilon_1)>0$, where $\Delta U(D)=U(A_S,1,D)-U(A_R,0,D)$ and ε_1 is an individual-specific error. For D=0, deception is advantageous when $p\,(\Delta U(0)+\varepsilon_2)>0$. Similarly, for D=1, the selfish allocation A_S is advantageous in the dictator game when (and only when) $(r_2-r_1)(\Delta V(1)+\varepsilon_3)>0$, where $\Delta V(D)=U(A_S,0,D)-U(A_R,0,D)$; when D=0, allocation A_S is advantageous when $\Delta V(0)+\varepsilon_4>0$. Our null hypothesis is that the base utility difference-in-difference is zero, $(\Delta U(1)-\Delta U(0)) - (\Delta V(1) - \Delta V(0)) = [U(A_S,1,1) - U(A_S,0,1)\ ] - [U(A_S,1,0)\ - U(A_S,0,0)] = 0$. Provided $(q_2-q_1)>0$ and $(r_2-r_1)>0$, utility changes and corresponding decisions are invariant to the probability levels and we can test the null hypothesis with estimations of the difference-in-difference in base utilities using observable decisions and attributes.

allocations. However, perhaps lying aversion depends to some extent on the probability with which a lie is implemented. For example, could aversion to a lie that occurs with low probability be more than proportional to the probability? This would be true if any positive probability of a lie, however small, leads to some non-negligible cost to the Sender. Alternately, perhaps a probabilistic implementation of a lie leads to a less-than-proportionate aversion cost, relative to the certain implementation of a lie. In either of these cases, the delegation effects we find could potentially be explained by the probabilistic differences between delegated and direct treatments. While these probabilistic differences are features of actual delegation situations, they may cloud the identification of pure delegation effects on lying aversion.

In our next round of experiments, we redesigned several elements in order to achieve a cleaner examination of the delegation effect. Specifically, we implemented a: (1) between-subject design; that (2) explicitly controls for probabilities that choices are implemented; and (3) a framework for delegation in the dictator game that is explicitly tied to agent behavior in that game.

4.3. The Second Round Experiments

4.3.1. The Deception Treatments

4.3.1. A. Direct Treatment: In the direct treatment, Senders were asked to send a message to a Receiver about the color of a dot which they observed at the top of the page in the questionnaire. The possible messages were:

Message GREEN: I solemnly swear that the dot is GREEN

Message BLUE: I solemnly swear that the dot is BLUE.

The Sender was informed that the Receiver would then have to report the color of the dot, based only on the Sender's message. To minimize strategic considerations, Senders were told that Receivers usually reported the color that Senders told them. This statement was based on Erat (2013) in which fifteen out of sixteen Receivers reported according to the message they received. In our experiment, 80.4 per cent of Receivers reported according to the message they received. There are two different treatments with respect to payouts. The experiment was designed as a between-subject design, so that each Sender in the deception experiment only made a choice for one of these payouts.

a. *Selfish Black Lie Treatment* – the Sender is benefited and the Receiver hurt if the Receiver reports the color of the dot incorrectly. If the Receiver reports the correct color of the dot, both the Sender and the Receiver are paid $5. If the Receiver reports the dot color incorrectly, the Sender is paid $7 and the Receiver $3.

b. *Pareto White Lie Treatment* – both are benefited by the lie, if the Receiver reports according to the message. If the Receiver reports the correct color of the dot, both the Sender and the Receiver are paid $5. If the Receiver reports the incorrect color of the dot, both the Sender and the Receiver are paid $6.

4.3.1.B. Delegation Treatment.

Agents. We first ran a modified session of the direct deception experiment with a group of graduate students, whom we designated as Agents. This session differed from the description above in two respects: 1) The Agents made choices for both of the above

payouts; 2) The payments to the Senders and Receivers were described to the Agents, but they were told that their own payment would be based both on how many Senders select the message they chose, and how many other agents selected the same message. For example, if, for one of the payout scenarios, there were three agents who chose the BLUE message, and thirty student senders, then each of the three agents would be one of the alternatives offered to ten students. Agents were paid $.50 for each student selecting them as their agent. From this session, we obtained at least one agent for each possible choice that the Senders could make.

Senders. The Senders were shown a colored dot as in the direct treatment above, and told that a message would be sent to their Receiver, who would then, based only on that message, have to make a report about the color of the dot that would determine payouts to both the Sender and Receiver. Again, each Sender participated in one of the payout treatments. Instead of choosing the message themselves, the Senders were told that previously Agents had made choices about which message to send to a Receiver. Each Agent was associated with three messages. Messages One and Two were the Agent's own choice and Message Three was the correct color of the dot. So if the correct color was blue, and an Agent chose Message Blue, all three messages associated with that Agent were Blue, whereas if the Agent chose Message Green, the messages associated with that Agent were two Green messages and one Blue Message. If an Agent was chosen, one of the three associated messages would be randomly selected to be sent to the Receiver. The Sender's task was to decide which Agent they would prefer to be selected – Agent 1 who is associated with two Green and one Blue Message, or Agent 2 who is associated with three Blue messages. Thus, if a Sender chose Agent 1, an incorrect message would be sent with two thirds probability, and if Agent 2 was chosen, the correct message would be sent with one hundred percent probability. Figure 2 illustrates the treatments, decisions and probabilities of implementation for the deception treatments of this experiment.

Receivers. As in the direct treatment, Receivers had to make a report on the color of the dot, with their only information being the message they received.

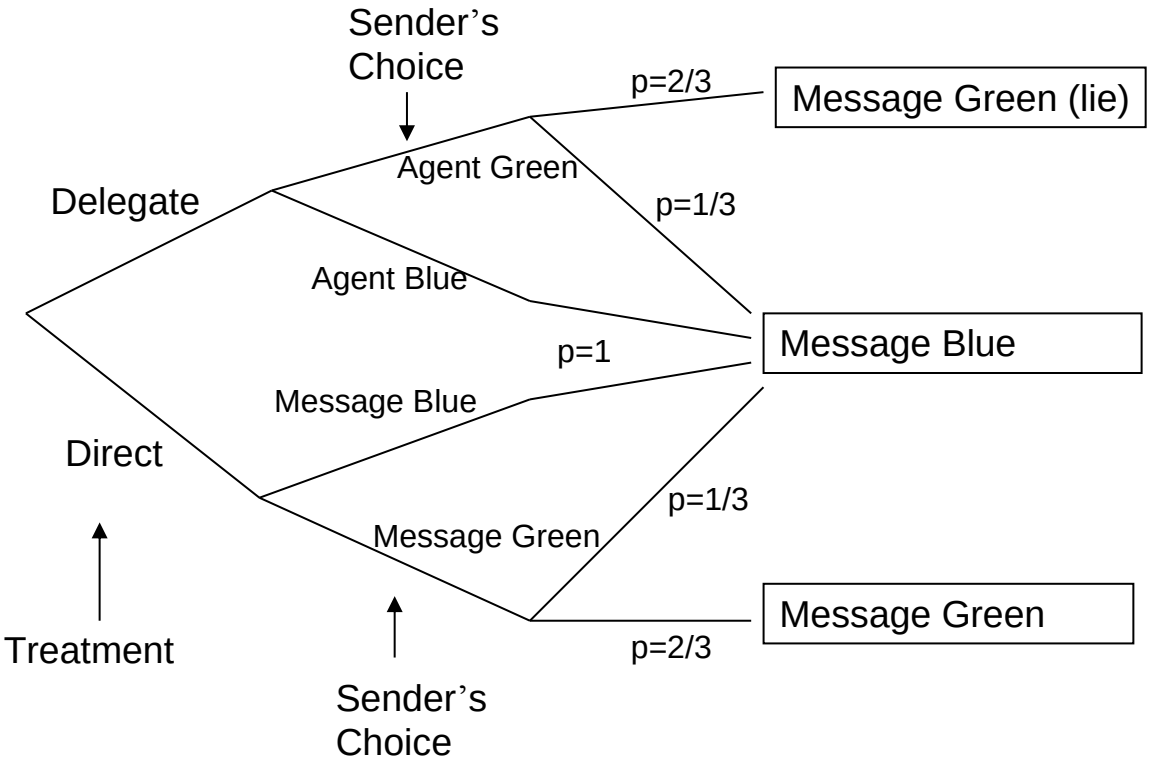


Figure 4.2 Deception Treatments Decision Tree

4. 3.2. The Dictator Treatments.

4. 3.2. A. ***Direct Treatment***: In the direct treatment, participants are asked to make allocation choices that have the same payoffs implemented with the same probabilities as in the Black Lie deception treatments, as detailed above.

4. 3.2. B. Delegation Treatment: In the delegation treatment, participants were told that an allocation would be made with payments to be made to themselves and their assigned receiver. They were told that they would not make the allocation decision themselves, but rather, would choose an agent from one of two groups. One group consisted of people who previously participated in a similar dictator game and chose the generous allocation, and the other of people who in a previous similar dictator game chose the selfish allocation.

4. 3.3. Logistics

Participants were students in economics classrooms at UC Merced and the sessions lasted twenty to thirty minutes. Since the experiment was conducted utilizing a between subject design with respect to both payoffs and treatment, there were eight different types of treatment groups (four treatments times two different payoff sets). We also varied the dot color, so we had a total of sixteen different questionnaires, randomly assigned within classrooms. Anonymity was preserved by having participants retain the number tags attached to the questionnaires matching the Registration number on the questionnaire. Participants then presented these number tags at the beginning of class the following

week to receive their payments which had been placed in envelopes, and thus were not visible at the time they were handed to participants.

4.3. 4. Results

We present first the Black Lie results, followed by the White Lie results.

4.3.4.A. Black Lie Results

In total, in the Black Lie sessions, there were 365 Sender-Dictator/Receiver pairs, with 87 Senders in the direct deception game, 94 in the delegated deception, 88 in the direct dictator, and 87 in the delegated dictator. Table 4.6 presents proportions in the four different categories, for all subjects and broken down by gender, along with difference statistics and difference-in-difference statistics.

Overall, the effect of delegation in the deception game is to lower the proportion sending the truthful message by approximately twenty six percentage points, from 74.7% to 48.9%. This is significant at the 1% level ($z = 3.712$). Females were more truthful than males, with the difference not significant in the direct treatment, but significant at ten percent in the delegated treatment. This is consistent with the mixed findings in the literature on lying by gender (see e.g. Dreber and Johannsen (2008), Childs (2012)). Erat (2013) found that women delegate more, but, as noted above, his design does not allow one to conclude whether it is liars or truth-tellers (in a direct game) who are choosing to delegate. Here, we see a smaller difference for women in the proportion telling the truth directly and delegating to the more truthful agent. Thus, men seem more likely to delegate to the selfish agent, both absolutely and relative to what they do when making the decision directly, as measured by the difference between truth-tellers in the direct and delegated treatments (.3 for men versus .213 for women). While this difference is not significant in itself, combined with the lower level of truth-telling by men in the direct treatment, it leads to the significant difference by gender in the delegated treatment. By contrast, while we find that delegation reduced the proportion being generous in the dictator game, the difference was not significant, overall, or for either gender. Interestingly, however, the difference between the genders was significant at the five percent level for both the direct and delegated treatment, with women making the generous choice approximately twenty three percent more often than men in the direct treatment and twenty two percent more in the delegated treatment. Earlier studies have found women to be more generous than men in direct dictator games (Eckel and Grossman 1998, Croson and Gneezy 2008), but this is, to our knowledge, the first study to show that the significant difference in generosity holds whether the allocation decision is made directly or through delegation.

Our main result is that the effect of delegation on lying aversion is significant even when controlling for possible changes in preferences over allocations. The difference-in-difference statistic is significant at the five percent level ($z=2.021$). This confirms the result in the first-round experiment, while improving the design of the delegated dictator game to both make the choice of agent more natural, and to control for probabilities of choice implementation, as well as utilizing a between subject design for the deception and dictator games to preclude spillover effects from behavior in one game to the other.

We next calculate the Hurkens-Kartik (HK) statistic for this experiment. For the between subject design of the games, this means looking at the difference in the ratio of the truthful to selfish in the direct and delegated treatment. The appropriate z-statistic is quite complex and is derived in Appendix D. Table 4.7 presents the ratios and the difference statistic. The z-statistic is significant at the five percent level ($z = 2.242$). Thus the result above is confirmed and strengthened. *Delegation decreases lying aversion when controlling for preferences over allocations.*

Table 4.8 confirms conclusions of Table 4.6 with regressions that control for course effects and gender. The regressions include all data from all four treatments. The endogenous variable is the zero-one indicator for a truthful or generous choice in, respectively, a deception and dictator game. The key regressors are the dummies for deception, delegation, and their interaction. The interaction captures the difference-in-difference effect of delegation on truthfulness vs. generosity. The models add controls as one moves from left to right, starting with the pure difference-in-difference model 1, adding course effects in model 2, gender effects in model 3, and gender-treatment interactions in model 4.[31] In all models, the difference-in-difference interaction is statistically significant at five percent, with coefficients that are remarkably stable across models. The estimated delegation effect on truthfulness, over and beyond its effect on the propensity for generosity, is a reduction of between 20.7 percent (model 1) and 21.7 percent (model 2). Male gender has a significantly negative effect on truthfulness and generosity, but there is no evidence of significant differences in the gender effect across games and treatments.

The propensity for truthfulness is significantly greater than is the propensity for generosity, as indicated by coefficients on the Deception dummy. These coefficients confirm prior work documenting lying aversion in direct games. However, we cannot reject the hypothesis of a zero deception effect on the choices of subjects in the delegation treatments. Test statistics for the null of a zero sum of coefficients on the Deception dummy and the Deception-Delegation interaction have p-values ranging from 0.38 (model 4) to 0.71 (model 1). In fact, comparing our delegation samples, the propensity for truthfulness is less than the corresponding propensity for generosity, although the two are very close and not different in a statistical sense.

Overall, the second-round Black Lie results confirm the conclusion of the first-round experiment that *delegation reduces lying aversion*. However, contrary to the first round experiment, the second-round Black Lie experiment does not provide positive evidence that there is lying aversion in the delegated treatment. The White Lie experiment provides some more direct evidence on both conclusions.

4.3.4.B. White Lie Results

While the Black Lies experiment seems to provide compelling evidence for the impact of delegation on lying aversion when controlling for preferences over allocation, we also implemented a Pareto White Lies design where the lying outcome benefitted both participants, so would presumably be the preferred allocation for all participants. In these sessions there were 78 Sender/Receiver pairs, 39 in each of the direct and delegated treatments. Table 4.9 presents the results.

[31] The regressions are OLS with robust errors, so that coefficients estimate marginal effects. Similar results are obtained from qualitative dependent variable models (such as Probit).

Again, there is a significant effect of delegation in reducing truthfulness, with the difference being significant at the five percent level. Yet, even when the lie is benefiting both the Sender and the Receiver, we see evidence of a considerable amount of lying aversion in both treatments. Approximately sixty seven percent of Senders tell the truth when sending the message themselves directly. This proportion is much higher than the rate of truth-telling in the two Pareto White Lie treatments in Erat and Gneezy (2012) who found that forty nine percent lied when the gain was 1 to the Sender and 10 to the Receiver, and sixty five percent lied when the gain from the lie was 10 to the Sender and 10 to the Receiver. However, our utilization of the fairly strong message "I solemnly swear that the dot is _____" may account for the higher direct level of truth-telling. Both males and females show a reduction in truth-telling between the direct and delegated treatment, but the relatively small numbers mean that neither reaches significance at the ten percent level when analyzed separately. Females are more truthful than males, but again the low numbers mean that the difference fails to reach statistical significance. Similarly to the Black Lie results above, the effect of delegation is slightly stronger for males, but not significantly so.

Panel B of Table 4.9 presents regressions that control for course and gender effects, confirming the evidence from the coarse statistics. The main conclusion of the White Lies treatment is that the earlier results are confirmed. *When the lie benefits both participants, delegation significantly reduces lying aversion.* However, it does not eliminate lying aversion. Even with delegation, over 43 percent of subjects are truthful to their detriment, a percentage that is significantly different from zero (z-statistic 5.49, p-value < 0.001).

A Potential Criticism of the White Lie experiment. We do not control for preferences over allocations in the White Lie experiment because of the natural presumption that all will prefer the ($6,$6) payoff to the truthful ($5,$5) counterpart. However, perhaps delegation affects the strength of relative preference for the former over the latter. If subjects prefer to directly choose the superior allocation, rather than have it chosen indirectly via an Agent, then our experimental results will understate the extent to which delegation reduces lying aversion; some subjects may then opt to tell the truth under delegation when they would not do so directly because they gain relatively less from the superior (untruthful) allocation. Alternatively, if subjects have a stronger preference for the superior allocation when it is obtained by a delegated choice than when it is obtained by a direct choice, then our experimental results would *overstate* the extent to which delegation reduces lying aversion.

We investigate the latter possibility with an additional experiment, the *Plus-One* design. This design borrows from the Bartling and Fischbacher (2012) experiments in which there are two matched sets of Sender-Receiver pairs. In the first pair (players B and D), the Sender (player B) chooses either an initial equal-split allocation or a second / plus-one allocation that adds $1 to the payoffs of both Sender (player B) and Receiver (player D). In the second pair of participants (players A and C, our interest), the Sender (player A) has three options: (1) implement the initial equal split, (2) implement the decision of the matched Sender B, or (3) directly add $1 to their own and their Receiver's (player C's) payoff, with the change implemented with the same probability as all player

B's in the experiment choose the plus-one allocation.[32] The order of the option choices is randomly varied.

In this experiment, indifference between the direct (option 3) alternative and the delegated (option 2) alternative would produce an equal split between the option selections. If more subjects prefer the delegated implementation to the direct implementation of the $1 change (the conjecture motivating the experiment), then we should see a larger proportion of subjects choosing the delegated option 2 than the direct option 3. In contrast, if more subjects prefer the direct to the delegated implementation, we should observe a larger proportion choosing the direct option 3 versus the delegated option 2.

The Plus-One experiment was implemented in one undergraduate economics class at U.C. Merced. Out of 48 A players, 26 chose to implement the plus-one change directly and 11 chose to delegate. The proportions (54.2% and 22.9%, respectively) are significantly different, indicating that more subjects prefer direct to delegated implementation – exactly the opposite of the conjectured objection to our Experiment 3 results.[33] This preference cannot be explained by probability differences, as the direct choice is implemented with exactly the same probability as the delegated choice and choosers are informed of this probabilistic equivalence.

These results suggest that the White Lies treatment of the second-round experiment *understates* the impact of delegation on lying aversion. They also reflect the intuition that delegation is advantageous when it distances individuals from decisions that are harmful to others (and beneficial to oneself), but disadvantageous when it distances individuals from decisions that are *beneficial* to others. In the Plus-One case, the decision is good news and the direct choice is preferred on average. This is the insight of Machiavelli (2003): "Princes should delegate to others the enactment of unpopular measures and keep in their own hands the means of winning favors" (see Bartling and Fischbacher, 2012; Eisenkopf and Fischbacher, 2014).

4.4. Discussion and Conclusion

In our experiments, subjects are more willing to lie when the lie is made with an agent's message than when the deceptive message is sent directly by the subjects themselves. This is true even though the outcomes of the choices in the two treatments – deceptive agent vs. truthful agent in the delegation treatment and deceptive message vs. truthful message in the direct treatment – are exactly the same, and even though we control for delegation effects on preferences over the payoff options.

These results add to the growing literature on what determines individuals' propensities for truthfulness. To date, scholars have shown that lying aversion is affected by the consequences for both sides of the interaction (Gneezy, 2005; Gibson et al., 2013),

[32] To implement the Plus-One experiment, all players are given the three Options. Delegating (Option 2) players are each randomly matched with a non-delegating (Option 1 or 3) player and the delegate's (Option 1 or 3) choice is implemented for the delegating player. The set of non-delegating players represents the set of all "player B's" in the experiment for purposes of the probabilistic implementation of Option 3. Experimental instructions are consistent with this approach.

[33] Restricting the sample to the 37 subjects who did not opt for the initial equal split, the z-statistic for the difference in proportions is 2.70 (p-value < 0.01). Surprisingly, 11 of 48 subjects chose the initial (low) equal split.

a norm of honesty (Pruckner and Sausgruber, 2013), social cues on how often others lie (Innes and Mitra, 2013), gender (Dreber and Johanneson, 2008), the extent of the lie (Lundquist et al., 2009; Fischbacher and Heusi, 2013), team incentives (Conrads et al., 2013), and cooperation in prior play (Ellingsen et al., 2009) but not cooperative (vs. competitive) priming (Rode, 2010). Our results indicate that lying aversion is also sensitive to delegation, that is, whether the decision is made directly or indirectly via choice of an agent.

This conclusion has implications for the use of delegation in markets. For example, a rich literature identifies tradeoffs in the choice between vertical integration and vertical separation (e.g., see Bolton and Dewatripont, 2005). If preferences for costly honesty are reduced by delegation, there is an added (private) motive for firms to vertically separate using outsourced suppliers and subcontractors. From an empirical point of view, these benefits are likely to be particularly relevant in economic environments wherein deception is normal and advantageous, such as in cultures with weak moral institutions. In these corrupt environments, both private economic benefits of dishonesty to contracting firms and opportunities for contracting with dishonest agents, are likely to be greater. From a normative (social) point of view, economic effects of contractual relationships are also likely to be more pernicious than they would be absent their impacts on lying aversion.

In practice, the impact of agency on dishonesty might be greater than suggested by our analysis. In the context of dictator games, Lazear et al. (2012) show that the opportunity for exit, and the corresponding opportunity for selection into the dictator role, leads to selection in favor of more selfish dictators. While neither we nor they examine selection into the deception game, one might plausibly conjecture that self-selection will favor more dishonest agents. Our results suggest an additional mechanism for the promotion of dishonesty, namely, the selection of Agents by principals.

Another feature of our design may also underestimate agency effects on honesty. Hamman et al. (2010) suggest that, in order to "reintroduce social pressure or obligation to behave altruistically" (p. 1844), principals could be informed of their agents' decisions and required to certify or override them. Our experiments implement this prescription by confronting Senders with information about the Agents' decisions. Consistent with the Hamman et al. (2010) conjecture, this may explain why we find so little effect of delegation on generosity in our dictator games. Absent this set-up, with agents competing for the custom of principals as in the Hamman et al. (2010) experiments, we might expect to see more selfishness and more dishonesty by subjects in the delegation treatments.

On the flip side, our results suggest that first party interactions are likely to be more truthful than are second party interactions. For example, private party sellers of used cars are more likely to be honest than are sales representatives in used car dealerships. This may help to explain (and justify) Gneezy's (2005) survey results indicating that students overwhelmingly have this belief.

Overall, our results speak to the importance of responsibility to social behavior, at least in the context of deception. Bartling and Fischbacher (2012) show that a measure of a subject's responsibility for a decision better explains punishment behavior of others at the receiving end than do outcome-based or intention-based competitors; that is, how

others respond to our actions depends upon how responsible they think we are for those actions. Our results suggest that if a subject perceives a reduced responsibility for a lie, even if *true* responsibility is not changed in an objective way, their *own* moral compass is altered – even absent any role for response or even full understanding by those at the receiving end of the lie. Acting through another person, as in our delegation treatments, appears to have this effect. In this sense, the results support the non-consequentialist view that preferences depend upon process as much as payoffs, and that this is important in economic interactions (Charness and Dufwenberg, 2006; Charness and Rabin, 2002) and that self-concept and self-justification are key drivers of moral decisions (Mazar et al., 2008).

Table 4.1
Summary Statistics and Difference and Difference-in-Difference Statistics for Direct and Delegated Treatments

Game	Direct or Delegated	Proportion Truthful or Generous
Deception	Direct Told Truth	.583 42/72
Delegated to More Truthful Sender	.357 25/70	
	z-stat for difference between direct and delegated	2.722***
Dictator	Direct Generous	.528 38/72
	Delegated to More Truthful Sender	.514 36/70
	z-stat for difference between direct and delegated	0.167
	z-stat for diff-in-diff between the games across treatments	1.919*

*,**,*** denotes significance at 10%,5%,1% respectively (2-tailed)

Table 4.2
Summary Statistics and Difference and Difference-in-Difference Statistics for Direct and Restricted Delegated Sample

Game	Direct or Delegated	Proportion Truthful or Generous
Deception	Direct Told Truth	.583 42/72
	Delegated to More Truthful Sender	.327 16/49
	z-stat for difference between direct and delegated	2.886***
Dictator	Direct Generous	.528 38/72
	Delegated to More Truthful Sender	.510 25/49
	z-stat for difference between direct and delegated	0.195
	z-stat for diff-in-diff between the games across treatments	2.039**

Table 4.3
Truthfulness in Deception Game Conditional on Selfish Choice in Dictator Game (Hurkens-Kartik statistic)

Treatment	Proportion Truthful (Full Sample)	Proportion Truthful (Restricted Sample for Delegated)
Direct	0.529 18/34	0.529 18/34
Delegated	0.294 10/34	0.208 5/24
z-stat for difference between direct and delegated	2.026**	2.695**

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).

Table 4.4
Regression Results

	Overall Truthfulness	Hurkens-Kartik Percent Truthfulness for Selfish	Difference-in-Difference	
Dep. Variable / **Regressor**	**Truth (1) Marg. Effect (n=142)**	**Truth (2) Marg. Effect (n=69)**	**Truth – Generous (3) Marg. Effect (n=142)**	**Truth – Generous (4) Marg. Effect (n=142)**
Delegate	**-0.2524 (0.0833)*** **	**-0.2181 (0.1200)***	**-0.2241 (0.1093)** **	**-0.2549 (0.1098)** **
Gender	-0.0461 (0.0856)	-0.0885 (0.1215)	**No**	-0.0401 (0.1104)
Course Effects	Yes	Yes	Yes	Yes
R^2	0.0703	0.0639	0.0517	0.0615

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).
[A] Restricted Sample: Restricted to subjects with reported beliefs consistent with more truthful behavior by $1 (vs. $2) Senders.
[B] OLS regressions with robust standard errors in parentheses.

Table 4.5
Distribution of Choice Pairs in the Deception and Dictator Games by Treatment

Choice in dictator and deception games	Direct Treatment	Delegated Treatment	Total
Selfish and Liar	16	24	40
Selfish and Truth	18	10	28
Generous and Liar	14	21	35
Generous and Truth	24	15	39
Total	72	70	142

Table 4.6
Summary Statistics and Difference and Difference-in-Difference Statistics for Direct and Delegated Treatments in Black Lie Second-Round Experiment

Treatment		Mean Truthful (Deception) or Generous (Dictator)			z-stat for difference between male and female
		All	Male	Female	
Deception Game	Direct Told Truth	.7471 65/ 87	.696 32/46	.805 33/41	0.906
	Delegated to More Truthful Agent	.4894 46/94	.396 19/48	.587 27/46	1.886*
	z-stat for difference between direct and delegated	3.712***	3.065***	2.285**	
Dictator Game	Direct Generous	.5682 50/88	.471 24/51	.703 26/37	2.261**
	Delegated to More Generous Agent	.5172 45/87	.413 19/46	.634 26/41	2.114**
	z-stat for difference between direct and delegated	0.678	0.385	0.699	
	z-stat for diff-in-diff between treatments in the two games	2.021**	1.723*	1.043	

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).

Table 4.7
Hurkens-Kartik Statistic

	Proportion of Liars divided by Proportion of Selfish
Direct	.5857
Delegated	1.058
z-stat for difference	2.242**

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).

Table 4.8.
Regressions[A]. (Dependent Variable: Truth)

	Model			
	(1) Marginal Effect (Standard Error)	(2) Marginal Effect (Standard Error)	(3) Marginal Effect (Standard Error)	(4) Marginal Effect (Standard Error)
Deception Dummy	0.1789 (0.0708)**	0.1830 (0.0692)***	0.1733 (0.0685)**	0.1296 (0.0871)
Delegation Dummy	-0.0509 (0.0756)	-0.0498 (0.0741)	-0.0596 (0.0726)	-0.0506 (0.0918)
Deception*Delegation	**-0.2068 (0.1030)****	**-0.2167 (0.1011)****	**-0.2101 (0.0994)****	**-0.2079 (0.0998)****
Course Effects	No	Yes	Yes	-0.2128 (0.0911)**
Gender(Male)	No	No	-0.1823 (0.0514)***	0.0795 (0.1006)
Male*Deception	No	No	No	0.0795 (0.1006)
Male*Delegation	No	No	No	-0.0200 (0.1003)
R^2	0.0410	0.0880	0.1200	0.1217

Notes:
[A] OLS regressions with robust standard errors.
*,**,*** denotes significant at 10%, 5% or 1% (two tail).

Table 4.9
Proportions selecting the truthful choice by treatment and gender with difference statistics in White Lies Treatment

	All	Male	Female	z-stat for difference between genders
Direct	0.66667 26/39	0.6111 11/18	0.71429 15/21	0.681
Delegated	0.43590 17/39	0.35294 6/17	0.50000 11/22	0.934
z-stat for difference between Direct and Delegated Treatments	2.107**	1.582	1.476	

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).

B. Regressions

	Dependent Variable: Truth N=78	
	(1) Marginal Effect (Standard Error)	(2) Marginal Effect (Standard Error)
Delegate	**-0.2308** **(0.1110)****	**-0.2343** **(0.1105)****
Gender (Male)	No	-0.1388 (0.1100)
Course Effects	Yes	Yes
R^2	0.0770	0.0954

Table 4.10
Plus-One Experiment

Choice	Number	Proportion
No Change	11/48	.229
Implement Plus One Change Directly	26/48	.542
Delegate the Implementation of the Plus One Change	11/48	.229
Difference in Proportion between Direct and Delegated Implementation (z-stat)		.313 (2.26)**

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail).

References

Abeler, J., A. Becker, & A. Falk. 2014. "Representative Evidence on Lying Costs." *Journal of Public Economics*, 113, 96-104.

Allingham, M. G. and S. Agnar. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1, 323–38.

Alm, J. 2012. "Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Field Studies, and Experiments." *International Tax and Public Finance*, 19(1), 54-77.

Al-Ubaydli, O., D. Houser, J. Nye, M. Paganelli, & X. Pan. 2013. "The Causal Effect of Market Priming on Trust." *PLOS ONE*, 8(3), 1-8.

Altmann, S., T. Dohmen, & M. Wibral. 2008. "Do the Reciprocal Trust Less?" *Economics Letters*, 99(3), 454-457.

Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Wlow Giving." *Economic Journal*, 100(401), 464-477.

Andreoni, J. & B. D. Bernheim. 2009. "Social Image and the 50-50 Norm." *Econometrica*, 77(5): 1607-1636.

Aoki, K., K. Akai, & K. Onoshiro. 2010. "Deception and Confession: Experimental Evidence from a Deception Game in Japan." Osaka University Discussion Paper No. 786. Available at SSRN: http://ssrn.com/abstract=1677773.

Ashraf, N., I. Bohnet, & N. Piankov. "Decomposing Trust and Trustworthiness." *Experimental Economics*, 9(3), 193–208..

Aumann, Robert. 1990. "Nash Equilibria Are Not Self-Enforcing." In *Economic Decision Making: Games, Econometrics, and Optimization*, ed. J. Gabszewicz, J. F. Richard, and L. Wolsey, 201–06. New York: Elsevier Science Publishers.

Bardsley, N. 2008. "Dictator Game Giving: Altruism or Artifact?" *Experimental Econ.*, 11, 122–33.

Bartling, B., F. Engl, & R. Weber. 2014. "Does Willful Ignorance Deflect Punishment? – An Experimental Study," *European Economic Review*, 70, 512-24.

Bartling, B., & U. Fischbacher. 2012. "Shifting the Blame: On Delegation and Responsibility." *Review of Economic Studies* 79(1), 67–87.

Battigalli, P. & M. Duwfenberg. 2007. "Guilt in Games." *American Economic Journal*, 97(2), 170-176.

Battigalli, P., G. Charness & M. Dufwenberg. 2013. "Deception: The Role of Guilt," *Journal of Economic Behavior & Organization*, 93, 227-233.

Bechara, A., H. Damasio, D. Tranel & A. Damasio. 1997. "Deciding Advantageously before Knowing the Advantageous Strategy." *Science*, 275, 1293-5.

Bellemare, C. & S. Kroger. 2007. "On Representative Social Capital." *European Economic Review*, 51(1), 183–202.

Ben-Ner, A., L. Putterman, F. Kong & D. Magan. 2004. "Reciprocity in a Two-Part Dictator Game." *Journal of Economic Behavior & Organization*, 53, 333–35

Berg, J., J. Dickhaut, & K. McCabe. 1995. "Trust, Reciprocity, and Social-history." *Games & Economic Behavior*, 10(1), 122–142.

Bloom, N., R. Sadun & J. Van Reenen. 2009. "The Organisation of Firms across Countries." Working Paper. *CEP Discussion Papers*, 937. Centre for Economic Performance, London School of Economics and Political Science.

Bok, S. 1978. *Lying: Moral Choice in Public and Private Life*. New York.Pantheon Books.

Bolton, P. & M. Dewatripont. 2005. *Contract Theory*. MIT Press.

Brandts, J. & G. Charness. 2003. "Truth or Consequences: An Experiment." *Management Science*, 49(1), 116-130.

Brandts, J. & G. Charness. 2011. "The Strategy versus the Direct-response Method: a First Survey of Experimental Comparisons." *Experimental Economics*, 14(3), 375-398.

Broberg, T., T. Ellingsen, & M. Johannesson. 2007. "Is Generosity Involuntary?" *Economics Letters* 94(1): 32–37.

Burks, S., J. Carpenter, & E. Verhoogen. 2003. "Playing both Roles in the Trust Game." *Journal of Economic Behavior & Organization*, 51, 195–216.

Cadsby, C.B., F. Song & F. Tapon. 2010. "Are you Paying your Employees to Cheat? An Experimental Investigation." *B.E. Journal of Economic Analysis & Policy*, 10(1): Article 35.

Cappelen, A., E. Sorenson & B. Tungodden. 2013. "When Do We Lie?" *Journal of Economic Behavior and Organization*, 93, 258-265.

Capra, C. M. 2004. "Mood-driven Behavior in Strategic Interactions." *AEA Papers & Proceedings, American Economic Review*, 95(2), 367-372.

Charness, G. 2000. "Responsibility and Effort in an Experimental Labor Market," *Journal of Economic Behavior and Organization*, 42, 375-384.

Charness, G., R. Cobo-Reyes, N. Jiménez, J. Lacomba & F. Lagos. 2012. "The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game." *American Economic Review,* 102(5), 2358–2379.

Charness, G. & M. Dufwenberg. 2006. "Promises and Partnership." *Econometrica,* 74(6), 1579-1601.

Charness, G. & M. Dufwenberg. 2010. "Bare Promises: An Experiment." *Economic Letters,* 107, 281-283.

Charness, G. and M. Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117, 817-868.

Charness, G. and M. Rabin. 2005. "Expressed Preferences and Behavior in Experimental Games." *Games & Economic Behavior*, 53, 151-169.

Charness, G. & M. Sutter. 2012. "Groups Make Better Self-Interested Decisions." *Journal of Economic Perspectives,* 26(3): 157-76.

Chaudhuri, A. & L. Gangadharan. 2007. "An Experimental Analysis of Trust and Trustworthiness." *Southern Economic Journal,* 73(4), 959-985.

Childs, J. 2012. "Gender Differences in Lying." *Economics Letters*, 114, 147–149.

Coffman, L. 2011. "Intermediation Reduces Punishment (And Reward)." *American Economic Journal: Microeconomics,* 3(4), 77- 106.

Conrads, J., B. Irlenbusch, R.M. Rilke & G. Walkowitz. 2013 . "Lying and Team Incentives." *Journal of Economic Psychology* 34, 1–7.

Costa-Gomez, M., S. Huck & G. Weizsacker. 2010. "Beliefs and Actions in the Trust Game: Creating Instrumental Variables to Estimate the Causal Effect." *IZA Disc. Paper* 4709.

Cox, J., D. Friedman & V. Sadiraj. 2008. "Revealed Altruism." *Econometrica*, 76(1), 31-69.

Crawford, V. & J. Sobel. 1982. "Strategic Information Transmission." *Econometrica,*, 50, 1431–1451.

Croson, R. & U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2): 448–474.

Dana, J., D. Cain & R. Dawes. 2006. "What You Don't Know Won't Hurt Me: Costly (But Quiet) Exit in Dictator Games." *Organizational Behavior and Human Decision Processes*, 100(2), 193-201.

Dana, J., R. Weber & J. Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory,* 33(1), 67-80.

DellaVigna, S. & M. Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics*, 2, 643–669.

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

Deneckere, R. & S. Severinov. 2007. "Optimal Screening with Costly Misrepresentation," Mimeo.

DePaulo, B.D., S. Kashy, N. Kirkendol, M. Wyer & J. Epstein, J. 1996. "Lying in Everyday Life." *Journal of Personality & Social Psychology,* 70(5), 979-95.

Dreber, A. & M. Johannesson. 2008. "Gender Differences in Deception." *Economics Letters,* 99, 197–199.

Drugov, M., J. Hamman & D. Serra. 2014. "Intermediaries in Corruption: An Experiment." *Experimental Economics* 17: 78-99.

Dufwenberg, M. & U. Gneezy. 2000. "Measuring Beliefs in an Experimental Lost Wallet Game." *Games & Economic. Behavior*. 30, 163–182.

Dufwenberg, M., U. Gneezy, W. Güth, & E. van Damme. 2001. "Direct vs Indirect Reciprocity: an Experiment." *Homo Oeconomicus,* 18, 19–30.

Eckel, C. & P. Grossman. 2008. "Are Women Less Selfish than Men: Evidence from Dictator Games." *Economic Journal*, 108: 726-735.

Eisenkopf, G. & U. Fischbacher. 2014. "Naïve Responses to Kind Delegation." *Managerial and Decision Economics*, doi: 10.1002/mde.2688

Ellingsen, T., M. Johannesson, J. Lilja & H. Zetterqvist. 2009. "Trust and truth." *Economic Journal,* 119, 252-276.

Ellingsen, T., M. Johannsesson, S. Tjotta & G. Torsvig. 2010. "Testing guilt aversion." *Games & Economic Behavior,* 68, 95-107.

Ellingsen, T. & R. Östling. 2010. "When Does Communication Improve Coordination?" *American Economic* Review, 100, 1695–1724.

Erat, S. 2013. "Avoiding Lying: The Case of Delegated Deception." *Journal of Economic Behavior and Organization,* 93, 273-278.

Erat, S. & U. Gneezy. 2012. "White lies." *Management Science* 58 (4), 723–733.

Falk, A., M. Kosfeld.2006. "The Hidden Costs of Control." *American Economic Review,* 96 (5), 1611–1630.

Fehr, E., U. Fischbacher, B. von Rosenbladt, J. Schupp & G. Wagner. 2003. "Nationwide Laboratory Examining Trust and Trustworthiness by Integrating Behavioural Experiments into Representative Surveys." *CEPR Disc. Paper No. 3858*.

Fehr, E. & K.M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114, 817–868.

Fershtman, C., & U. Gneezy. 2001. "Strategic Delegation: An Experiment." *Rand Journal of Economics*, 32(2):352-368.

Fershtman, C., K. L. Judd & E. Kalai. 1991. "Observable Contracts: Strategic Delegation and Cooperation," *International Economic Review*, 32: 551–559.

Fischbacher, U. & F. Föllmi-Heusi. 2013. "Lies in Disguise – An Experimental Study on Cheating." *Journal of the European Economic Association*, 11(3), 525-547.

Frank, R.H., T. Gilovich & D.T. Regan. 1993. "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives*, 7(2), 159-171.

Freeman, R. & A. Gelber. 2010. "Prize Structure and Information in Tournaments: Experimental Evidence." *American Economic Journal: Applied Economics*, 2(1), 149-164.

Friedman, D., S. Huck, R. Oprea & S. Weidenholzer. 2015. "From Imitation to Collusion: Long-run Learning in a Low-Information Environment." *Journal of Economic Theory* 155: 185-205.

Gibson, R., C. Tanner & A.F. Wagner. 2013. "Preferences for Truthfulness: Heterogeneity among and within Individuals." *American Economic Review*, 103(1), 532-48.

Glaeser, E., D. Laibson, J. Scheinkman & C. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics*, 65(3), pp. 811–846.

Gneezy, A., A. Imas, A. Brown, L.D. Nelson & M. I. Norton. 2011. "Paying to Be Nice: Consistency and Costly Prosocial Behavior." *Management Science*: 1–9.

Gneezy, U. 2005. "Deception: The Role of Consequences." *American Economic Review*, 95, 384–94.

Gneezy, U., B. Rockenbach & M. Serra-Garcia. 2013. "Measuring Lying Aversion." *Journal of Economic Behavior & Organization*, 93, 293-300.

Greenberg, A., M. Smeets & L. Zhrakhovska. 2015. "Lying and Shame Aversion." SSRN-id2587104.

Greene, J.D., L. Nystrom, L. Engell & J. Darley. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgement." *Neuron*, 44, 389–400.

Greene, J.D. & J. Paxton. 2009. "Patterns of Neural Activity Associated with Honest and Dishonest Moral Decisions." *Proceedings of the National Academy of Sciences*, 106, 12506–12511.

Gurdal, M.Y., A. Ozdogan & I. Saglam. 2014. "Truth-telling and Trust in Sender-Receiver Games with Intervention: an Experimental Study." *Review of Economic Design*, 18(2): 83-103.

Greiner, B., & V. M. Levati. 2005. "Indirect Reciprocity in Cyclical Networks: an Experimental Study." *Journal of Economic Psychology*, 26, 711–731.

Grossman, Z. 2015. "Strategic Ignorance and the Robustness of Social Preferences." *Management Science*, 60(11), 2659-2665.

Guiso, L., P. Sapienza & L. Zingales. 2004. "The Role of Social Capital in Financial Development." *American Economic Review*, 94, 526–556.

Guiso, L., P. Sapienza & L. Zingales. 2008. "Trusting the Stock Market." *Journal of Finance*, 63, 2557–2600.

Guiso, L., P. Sapienza & L. Zingales. 2009. "Cultural Biases in Economic Exchange." *Quarterly Journal of Economics*, 124(3), 1095–1131.

Gylfason, H., A. Arnardottir & K. Kristinsson. 2013. "More on Gender Differences in Lying." *Economics Letters* 119, 94-96

Haisley, E. & R. Weber. 2010. "Self-Serving Interpretations of Ambiguity in Other-Regarding Behavior." *Games and Economic Behavior*, 68(2): 614–625.

Hamman, J., G. Loewenstein, & R. Weber. 2010. "Self-Interest Through Delegation: An Additional Rationale for the Principal–Agent Relationship." *American Economic Review* 100 (4): 1826–1846.

Hao, L. & D. Houser, D. 2011. "Honest Lies." Discussion Paper, Interdisciplinary Center for Economic Science, George Mason University.

Herne, K., O. Lappalainen & E. Kestilä-Kekkonen. 2013. "Experimental Comparison of Direct, General, and Indirect Reciprocity." *Journal of Socio-Economics*, 45, 38-46.

Houser, D., D. Schunk & J. Winter. 2010. "Distinguishing Trust from Risk: an Anatomy of the Investment Game." *Journal of Economic Behavior & Organization*, 74(1–2), 72–81.

Houser, D., S. Vetter & J. Winter. 2012. "Fairness and Cheating." *European Economic Review*, 56, 1645–1655.

Hurkens, S. & N. Kartik, N. 2009. "Would I Lie to You? On Social Preferences and Lying Aversion." *Experimental Economics*, 12(2), 180-92.

Innes, R. & A. Mitra. 2013. "Is Dishonesty Contagious?" *Economic Inquiry*, 51(1), 722-734.

Johnson, N.D. & A.A. Mislin. 2011. "Trust Games: A Meta-analysis." *Journal of Economic Psychology*, 32, 865–889.

Josephson Institute of Ethics. 2012. *2012 Report Card on the Ethics of American Youth*. Los Angeles: Josephson Institute.

Kartik, N. 2009. "Strategic Communication With Lying Costs." *Review of Economic Studies*, 76 (4), 1359-1395

Kartik, N., M. Ottaviani & F. Squintanic. 2007. "Credulity, Lies and Costly Talk." *Journal of Economic Theory*, 134(1), 93-116.

Keizer, K., S. Lindenberg, & L. Steg. 2008. "The Spreading of Disorder." *Science*, 322, 1681-1685.

Kirchsteiger, G., L. Rigotti & A. Rustichini. 2006. "Your Morals Might be your Moods." *Journal of Economic Behavior & Organization*, 59, 155-172.

Knack, S. & P. Keefer. 1997. ''Does Social Capital have an Economic Payoff? A Cross-country Investigation." *Quarterly Journal of Economics*, 112, 1252–1288.

La Porta, R., F. Lopez de Silanes, A. Shleifer & R. Vishny. 1997. "Trust in Large Organisations." *American Economic Review*, 87(2), 333–338.

Lazear, E.; U. Malmendier & R. Weber. 2012. "Sorting in Experiments with Application to Social Preferences." *American Economic Journal: Applied Economics*, 4(1):136-163.

Lazzarini, S., R. Madalozzo, R. Artes & J. de Oliveira Siqueira. 2004. "Measuring Trust: an Experiment in Brazil." *Ibmec working paper* – WPE – 2004 – 1.

Lewicki, R. J. 1983. "Lying and Deception: A Behavioral Model." In *Negotiation in Organizations*. edited by Bazerman, M. H. & Lewicki, R. J.. Sage Publications, pp. 68–90.

Lewis, A., Carrera, S., Cullis, J., & Jones, P. 2009. "Individual, Cognitive and Cultural Differences in Tax Compliance: UK and Italy Compared." *Journal of Economic Psychology*, 30, 431 – 445.

List, J. 2007. "On the Interpretation of Giving in Dictator Games." *J. Political Econ.*, 115, 482–93.

Lopez-Perez, R. & E. Spiegelman.2013. " Why do People Tell the Truth? Experimental Evidence for Pure Lie Aversion." *Experimental Economics*, 16(3), 233-247.

Lundquist, T., T, Ellingsen, E. Gribbe & M. Johannesson. 2009. "The Aversion to Lying." *Journal of Economic Behavior & Organization*, 70 (1-2), 81-92.

Machiavelli, Niccolo. 2003. *The Prince*. London: Penguin Classics.

Mazar, N., O. Amir & D. Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self-concept Maintenance." *Journal of Marketing Research*, 45, 633-44.

Muehlheusser, G., A. Roider & N. Wallmeier. 2015. "Gender Differences in Honesty: Groups versus Individuals." *Economics Letters*, 128: (2015) 25–29.

Nelkin, D. K. 2004. "Moral Luck," in *The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta. Stanford University.

Oexl, R. & Z. Grossman, Z. 2013. "Shifting the Blame to a Powerless Intermediary." *Experimental Economics,* 16, 306–312.

Peeters, R., M. Vorsatz & M. Walzl. 2013. "Truth, Truth and Sanctions: On Institutional Selection in Sender-Receiver Games." *The Scandinavian Journal of Economics*, 115(2), 508-548.

Ploner, M. & T. Regner. 2013. "Self-Image and Moral Balancing: An Experimental Analysis." *Journal of Economic Behavior & Organization*, 93(C), 374-383.

Pruckner, G. & R. Sausgruber. 2013. "Honesty on the Streets: A Field Study on Newspaper Purchasing." *Journal of the European Economic Association*, 11(3), 661-79.

Rode, J. 2010. "Truth and Trust in Communication: Experiments on the Effect of a Competitive Context." *Games & Economic Behavior*, 68, 325–38.

Rosenbaum, S., S. Billinger & N. Stieglitz. 2014. "Let's be Honest: A Review of Experimental evidence of honesty and truth-telling." *Journal of Economic Psychology,* 45, 181–96.

Ross, L., D. Greene & P. House, P. 1977. "The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes." *Journal of Experimental Social Psychology*, 13, 279-301.

Sánchez-Pagés, S. & M. Vorsatz. 2007. "An Experimental Study of Truth-telling in a Sender–receiver Game." *Games & Economic Behavior,* 61, 86-112.

Sánchez-Pagés, S. & M. Vorsatz, M. 2009. "Enjoy the Silence: An Experiment on Truth-telling." *Experimental Economics,* 12 (2), 220-41.

Sapienza, P., A. Toldra-Simats & L. Zingales. 2013. "Understanding Trust." *Economic Journal*, 123, 1313–1332.

Shalvi , S., J. Dana, M. Handgraaf & C. De Dreu. 2011. "Justified Ethicality: Observing Desired Counterfactuals Modifies Ethical Perceptions and Behavior." *Organizational Behavior and Human Decision Processes,* 115, 181–190.

Shalvi, S., O. Eldar & Y. Bereby-Meyer. 2012. "Honesty Requires Time (and Lack of Justifications)." *Psychological Science*, 23(10), 1264-70.

Shalvi, S., M.J.J. Handgraaf & C.K.W. De Dreu. 2011. "Ethical Maneuvering: Why People Avoid both Major and Minor Lies." *British Journal of Management*, 22, S16–S27.

Stanca, L. 2009. "Measuring Indirect Reciprocity: Whose Back do We Scratch?" *Journal of Economic Psychology*, 30, 190–202.

Sutter, M. 2009. "Deception through Telling the Truth? Experimental Evidence from Individuals and Teams." *Economic Journal*, 119, 47–60.

Tyler, J.M., R.S. Feldman & A. Reichert, A. 2006. "The Price of Deceptive Behavior: Disliking and Lying to People who Lie to us." *Journal of Experimental & Social Psychology*, 42, 69–77.

Utikal, V. & U. Fishbacher. 2013. "Disadvantageous Lies in Individual Decisions." *Journal of Economic Behavior and Organization*, 85, 108-111.

Zak, P. & S. Knack. 2001. "Trust and Growth." *Economic Journal*, 111, 295–321.

60 Minutes. 2015. "Lumber Liquidators Linked to Health and Safety Violations." *CBS News*, March 1.

Appendices

Appendix A. Chapter 2 Experimental Instructions

A1. Experiment 1 Sender Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Receiver." For purposes of this experiment, you are a "Sender."

• Neither of you will ever know the identity of the other.

• The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

• Two possible monetary payments are available to you and your Receiver in this experiment:

Option C: $4 to you and $6 to the other student (your Receiver).
Option D: $6 to you and $3 to the other student (your Receiver).

• You know what the Options are, but your Receiver **DOES NOT**.

• You will choose between **ONE** of two possible messages, which will be sent to your Receiver:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

• Based **ONLY** on your message, your Receiver will **CHOOSE ONE** of the two Options. Your Receiver's choice will determine the payments in the experiment. Receivers will never know the true Options, or the sums to be paid to you under the different Options.

• **Information from Past Sessions**:

• In past sessions of this experiment, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• When participating in this experiment, your Receiver will have no information about decisions made by anyone in prior experiments.

Experiment (cont.)

• **YOUR DECISION**

****** I choose to send the following message (please circle ONE): ******

Message C **Message D**

• Please answer two questions for us:

1) Which option do you think your Receiver will choose based upon the message you sent? If you correctly predict your Receiver's choice, we will pay you an additional $1.

I believe my Receiver will choose (please CIRCLE ONE)

Option C Option D

2) What proportion of Senders in this session do you think will send Truthful messages? (A "Truthful" message indicates the Option that actually earns the Receiver more money.) If your prediction is correct (within five percentage points of the actual proportion, plus or minus), you will receive an additional $1 payment. CIRCLE ONE OF THE FOLLOWING PERCENTAGES.

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE AT YOUR STATION 1. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

A2. Experiment1 Receivers and Experiment 2 All Participant Instructions

Your Registration Number:_____

Introduction: In this experiment, you will make some decisions that will earn you some money.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER **ALL** QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM, AND DO NOT RETURN TO REVISE YOUR ANSWERS AT ANY LATER POINT IN THE EXPERIMENT.

Your registration number is written at the top of this page, and on the attached tag. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION to be announced by your experiment manager.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS. Even the experimenters will never know the person or name associated with a questionnaire.

BEFORE WE START: Please INDICATE YOUR **MOOD** right now. My **MOOD** is (CIRCLE ONE):

bad down so-so good very good great

Experiment 1

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Sender." For purposes of this experiment, you are a "Receiver."

• Neither of you will ever know the identity of the other.

• The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

• Two possible monetary payment Options (C and D) are available to you and your Sender in this experiment. One of the two Options is BETTER FOR YOU, and the other Option is BETTER FOR YOUR SENDER.

• We showed the two payment options to your Sender.

• **YOU** will choose **ONE** of the two options, which will determine the payments to the two of you. The ONLY information you will have is a message your Sender sent to you.

• Two possible messages could be sent to you:

Message C: "Option C will earn you more money than Option D."
Message D: "Option D will earn you more money than Option C."

• Your Sender decided to send you Message ____

• We now ask you to choose ONE of the two options. **(PLEASE CIRCLE ONE)**

Option C **Option D**

• Now we ask that you go to one of the **STATIONS** in the Room, where you will

• **TURN IN** the first two (stapled) pages of this questionnaire.

• Receive **FURTHER INSTRUCTIONS**.

• Please READ the FURTHER INSTRUCTIONS **AFTER** RETURNING TO YOUR SEAT and **BEFORE** completing the next experiment.

• PLEASE GO TO STATION 1 (Controls) / 2 (Treated).

FURTHER INSTRUCTIONS (Controls – Station 1)

• Please write down your Registration Number:_____

• We can now give you some information about Experiment 1. In Experiment 1, you chose a Payment Option for you and your Sender based on a Message/Recommendation of your Sender. The Message indicated the specific Payment Option that would earn you more money.

• In Experiment 1, roughly **5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED**.

• In prior sessions of Experiment 1, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• **Please INCLUDE THIS SHEET with your answers to the Experiment 2 questionnaire**.

FURTHER INSTRUCTIONS (Treated – Station 2)

• Please write down your Registration Number:_____

• We can now give you some information about Experiment 1. In Experiment 1, you chose a Payment Option for you and your Sender based on a Message/Recommendation of your Sender. The Message indicated the specific Payment Option that would earn you more money.

• In Experiment 1, roughly **5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED**.

• In prior sessions of Experiment 1, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• If your Registration number ENDS with a **V**,

Your Sender **TOLD YOU THE TRUTH** in Experiment 1 about the Option that earns you more money.

• If your Registration number ENDS with a **W**,

Your Sender **LIED TO YOU** in Experiment 1 about the Option that earns you more money.

• Based on this information, **PLEASE ANSWER THE FOLLOWING**:

Question: In experiment 1, my Sender (CIRCLE ONE):

TOLD ME THE TRUTH **LIED TO ME**

• **Please INCLUDE THIS SHEET with your answers to the Experiment 2 questionnaire**.

Your RegistrationNumber:____ _____

DO NOT START THIS SECOND EXPERIMENT UNTIL YOU HAVE GONE TO YOUR STATION AND RECEIVED FURTHER INSTRUCTIONS. THANK YOU!

Experiment 2

• This is a second short experiment that will earn you more money. We will call you **Participant 1** in this experiment.

• In this experiment, you are randomly matched with another student in another session, whom we will call **Participant 2**. **This is a different student than you are matched with in Experiment 1.** Neither of you will ever know the identity of the other.

• In this experiment, there are two types of participant, Senders and Returners.

• You will first be in the role of Returner and next be in the role of Sender. The same is true for Participant 2. Payments to the two of you will be determined either by

your strategy as Returner and Participant 2's as Sender (**Situation R**), OR
your strategy as Sender and Participant 2's as Returner (**Situation S**).

• Both of you will be paid for ONE of the two situations. The situation for which you will be paid is determined by a flip of a coin after all decisions have been made by all participants. YOU SHOULD THEREFORE MAKE YOUR DECISION IN EACH SITUATION AS IF IT IS THE ONE FOR WHICH YOU WILL BE PAID.

You and Participant 2 will be paid for **Situation R** if the coin toss is ***Heads***.
You and Participant 2 will be paid for **Situation S** if the coin toss is ***Tails***.

• Participant 2 will never be told what decisions you have made in this Experiment.

• Both you and Participant 2 will know the complete structure of this experiment before either of you makes any decisions.

• Please INDICATE YOUR **MOOD** right now. My **MOOD** is (**CIRCLE ONE)**:

bad down so-so good very good great

Situation R (You are Returner)

• You (the Returner) and Participant 2 (the Sender) start out with $4 each.

• Participant 2 (the Sender) has two choices:
KEEP. Keep the initial $4, implying that each of you will earn the initial $4 allocated to you.
SEND. Send his/her $4 to you, the Returner.

• If Participant 2 (the Sender) chooses SEND, the $4 sent will become $8, which combined with your initial $4, makes $12 available to you. In this case, you must choose between:
OPTION A. Return $7 to the Sender, so that **you receive $5 in total and your Sender receives $7 in total**.
OPTION B. Return $2 to the Sender, pay a fee of $2 and keep the remainder, so **you receive $8 in total and your Sender receives $2 in total**.
• The following table summarizes the payments. BE SURE YOU UNDERSTAND THESE PAYMENTS BEFORE PROCEEDING. Remember that **you are the Returner in Situation R**.

Table 1: Payments (in $)

		If Sender Chooses SEND		If Sender Chooses KEEP	
	Payment to	Payment to Returner	Payment to Sender	Payment to Returner	Sender
Returner's Option Choice	A	$5	$7	$4	$4
	B	$8	$2		

• Before you decide which Option to choose, you can either **send a message** to Participant 2 (the Sender) or **not send a message**. The message you can send, as the Returner, is:

MESSAGE A: I am going to choose Option A.

• Alternately you can choose to send **NO MESSAGE**. Note that **if you send Message A you are still permitted to choose Option B**.

• Participant 2 (the Sender) will make a choice (KEEP or SEND) for each possibility (if he/she receives Message A or No Message). Payments will then be determined (by Table 1) according to the decision made by the Sender for the **actual Message that you sent** (Message A or No Message), and (if the Sender chooses SEND) your choice of Option (Option A or Option B).

• Please make your decisions on the next page.

Situation R (You are Returner) (continued)

• For your reference, we reproduce the Payments Table 1 below.

• **MESSAGE CHOICE**. Please indicate the Message you will send to Participant 2 (the Sender).

** I choose the following (please CIRCLE ONE) **

MESSAGE A **NO MESSAGE**

• **OPTION CHOICE IN SITUATION R (AS RETURNER, if your Sender chooses SEND)**

** I choose the following (please CIRCLE ONE) **

OPTION A **OPTION B**

Table 1: Payments (in $)

		If Sender Chooses SEND		If Sender Chooses KEEP	
		Payment to Returner	Payment to Sender	Payment to Returner	Payment to Sender
Returner's Option Choice	A	$5	$7	$4	$4
	B	$8	$2		

Situation S (You are Sender)

• In Situation S, you are the Sender in the experiment described above, and Participant 2 is the Returner. Payoffs will be determined according to Table 1.

• You (the Sender) have two choices:

KEEP. You keep the $4, which will be your payment. (The Returner also keeps $4.)
SEND. You SEND $4 to your Returner. Your Returner then chooses between Option A and Option B.

• Your Returner has chosen EITHER to send you **Message A** ("I am going to choose Option A"), OR not to send you a message (**No Message**). You must choose whether to **KEEP** or **SEND** in either case. Payments will then be determined (by Table 1) according to the decision made by you (as Sender) for the **actual Message** chosen by your Returner, and (if you choose SEND) Participant 2's choice of Option (Option A or Option B).

YOUR DECISION IN SITUATION S (AS SENDER).

• **CASE 1: Suppose Participant 2 sends Message A** ("I am going to choose Option A").

• **DECISION FOR CASE 1**: **I choose the following (please CIRCLE ONE) **

KEEP **SEND**

• **CASE 2: Suppose Participant 2 sends NO Message**.

• **DECISION FOR CASE 2** ** I choose the following (please CIRCLE ONE) **

KEEP **SEND**

The following questions ask you to predict a percentage from Experiment 2. In each case, if your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment.

1) Of **SENDERS** who receive Message A, what proportion do you think will choose the **SEND** option? CIRCLE ONE OF THE FOLLOWING PERCENTAGES.

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

2) Of **RETURNERS** who send Message A, what proportion do you think will choose **OPTION A**? CIRCLE ONE OF THE FOLLOWING PERCENTAGES.

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) What proportion of **RETURNERS** in this experiment do you think will send **MESSAGE A**? CIRCLE ONE OF THE FOLLOWING PERCENTAGES.

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

4) If you could choose payments, which would you choose in each of the following choice situations? PLEASE CIRCLE ONE PAYMENT OPTION in EACH SCENARIO:

SCENARIO 1: Please **CIRCLE ONE** payment option (the one you prefer):

OPTION 1: $4 to you and $4 to Participant 2 OPTION 2: $2 to you and $8 to Participant 2

SCENARIO 2: Please **CIRCLE ONE** payment option (the one you prefer):

OPTION 1: $4 to you and $4 to Participant 2 OPTION 2: $8 to you and $2 to Participant 2

SCENARIO 3: Please **CIRCLE ONE** payment option (the one you prefer):

OPTION 1: $7 to you and $5 to Participant 2 OPTION 2: $2 to you and $8 to Participant 2

SCENARIO 4: Please **CIRCLE ONE** payment option (the one you prefer):

OPTION 1: $5 to you and $7 to Participant 2 OPTION 2: $8 to you and $2 to Participant 2

Final Questions (YOU ARE ALMOST DONE!)

5) a) In **Experiment 2**, do you think that someone who sends Message A *should* choose the indicated Option A?

Yes **Not Necessarily**

b) What proportion of participants in this experiment do you think will answer **Yes** to the last question (5a)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

6) Please tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE (INCLUDING THE FURTHER INSTRUCTIONS SHEET)</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. <u>PLEASE BRING YOUR TAG (WITH YOUR REGISTRATION NUMBER) TO THE CLASS FOR PAYMENT</u>.

Appendix B. Chapter 3 Experimental Instructions

General Instructions

Instructions **Your Registration Number:____ _____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENT ON THE NEXT PAGE…

B.1 Weak Message, $1 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.
There is a colored dot at the top of **THIS** page. What color is the dot?_____

Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: The dot is green.
Message BLUE: The dot is blue.
For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$2 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:

0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $2 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.2. Strong Message, $1 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____
Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: I solemnly swear that the dot is green.
Message BLUE: I solemnly swear that the dot is blue.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$2 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:
0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $2 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.3. Weak Message, $2 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____

Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: The dot is green.
Message BLUE: The dot is blue.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$3 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:
0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $3 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.4. Strong Message, $2 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____
Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: I solemnly swear that the dot is green.
Message BLUE: I solemnly swear that the dot is blue.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$3 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:

0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $3 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.5. Weak Message, $3 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____
Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: The dot is green.
Message BLUE: The dot is blue.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$4 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:

0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $4 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.6. Strong Message, $3 Gain To Lying

In this short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____
Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: I solemnly swear that the dot is green.
Message BLUE: I solemnly swear that the dot is blue.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$4 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:

0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $4 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B7. Prior Experience Experiment1 Sender Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Receiver." For purposes of this experiment, you are a "Sender."

• Neither of you will ever know the identity of the other.

• The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

• Two possible monetary payments are available to you and your Receiver in this experiment:

Option C: $4 to you and $6 to the other student (your Receiver).
Option D: $6 to you and $3 to the other student (your Receiver).

• You know what the Options are, but your Receiver **DOES NOT**.

• You will choose between **ONE** of two possible messages, which will be sent to your Receiver:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

• Based **ONLY** on your message, your Receiver will **CHOOSE ONE** of the two Options. Your Receiver's choice will determine the payments in the experiment. Receivers will never know the true Options, or the sums to be paid to you under the different Options.

• **Information from Past Sessions**:

• In past sessions of this experiment, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• When participating in this experiment, your Receiver will have no information about decisions made by anyone in prior experiments.

Experiment (cont.)

• **YOUR DECISION**

****** I choose to send the following message (please circle ONE): ******

Message C **Message D**

• Please answer two questions for us:

1) Which option do you think your Receiver will choose based upon the message you sent? If you correctly predict your Receiver's choice, we will pay you an additional $1.

I believe my Receiver will choose (please CIRCLE ONE)

Option C Option D

2) What proportion of Senders in this session do you think will send Truthful messages? (A "Truthful" message indicates the Option that actually earns the Receiver more money.) If your prediction is correct (within five percentage points of the actual proportion, plus or minus), you will receive an additional $1 payment. CIRCLE ONE OF THE FOLLOWING PERCENTAGES.

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE AT YOUR STATION 1. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

B.8. Prior Experience Experiment 1 Receivers and Experiment 2 All Participant Instructions

Your Registration Number:_____

Introduction: In this experiment, you will make some decisions that will earn you some money.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER **ALL** QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM, AND DO NOT RETURN TO REVISE YOUR ANSWERS AT ANY LATER POINT IN THE EXPERIMENT.

Your registration number is written at the top of this page, and on the attached tag. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION to be announced by your experiment manager.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS. Even the experimenters will never know the person or name associated with a questionnaire.

BEFORE WE START: Please INDICATE YOUR **MOOD** right now. My **MOOD** is (CIRCLE ONE):

bad down so-so good very good great

Experiment 1

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Sender." For purposes of this experiment, you are a "Receiver."

• Neither of you will ever know the identity of the other.

• The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

• Two possible monetary payment Options (C and D) are available to you and your Sender in this experiment. One of the two Options is BETTER FOR YOU, and the other Option is BETTER FOR YOUR SENDER.

• We showed the two payment options to your Sender.

• **YOU** will choose **ONE** of the two options, which will determine the payments to the two of you. The ONLY information you will have is a message your Sender sent to you.

• Two possible messages could be sent to you:

Message C: "Option C will earn you more money than Option D."
Message D: "Option D will earn you more money than Option C."

• Note that Sender Messages may or may not be truthful.

• Your Sender decided to send you Message ____

• We now ask you to choose ONE of the two options. **(PLEASE CIRCLE ONE)**

Option C **Option D**

• Now we ask that you go to one of the **STATIONS** in the Room, where you will

• **TURN IN** the first two (stapled) pages of this questionnaire.

• Receive **FURTHER INSTRUCTIONS**.

• Please READ the FURTHER INSTRUCTIONS **AFTER** RETURNING TO YOUR SEAT and **BEFORE** completing the next experiment.

• PLEASE GO TO STATION 1 (Controls) / 2 (Treated).

FURTHER INSTRUCTIONS (Controls – Station 1)

• Please write down your Registration Number:_____

• We can now give you some information about Experiment 1. In Experiment 1, you chose a Payment Option for you and your Sender based on a Message/Recommendation of your Sender. The Message indicated the specific Payment Option that would earn you more money.

• In Experiment 1, roughly **5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED**.

• In prior sessions of Experiment 1, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• **Please INCLUDE THIS SHEET with your answers to the Experiment 2 questionnaire**.

FURTHER INSTRUCTIONS (Treated – Station 2)

• Please write down your Registration Number:_____

• We can now give you some information about Experiment 1. In Experiment 1, you chose a Payment Option for you and your Sender based on a Message/Recommendation of your Sender. The Message indicated the specific Payment Option that would earn you more money.

• In Experiment 1, roughly **5 out of 10 Senders TOLD THE TRUTH and 5 out of 10 Senders LIED**.

• In prior sessions of Experiment 1, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• If your Registration number ENDS with a **Y**,

Your Sender **TOLD YOU THE TRUTH** in Experiment 1 about the Option that earns you more money.

• If your Registration number ENDS with a **Z**,

Your Sender **LIED TO YOU** in Experiment 1 about the Option that earns you more money.

• Based on this information, **PLEASE ANSWER THE FOLLOWING**:

Question: In experiment 1, my Sender (CIRCLE ONE):

TOLD ME THE TRUTH **LIED TO ME**

• **Please INCLUDE THIS SHEET with your answers to the Experiment 2 questionnaire**.

Experiment 2

In this second short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.
There is a colored dot at the top of **THIS** page. What color is the dot?_____

Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message BLUE: The dot is blue.
Message GREEN: The dot is green.

For each of the following scenarios, please circle the Message you choose to send.
SCENARIO 1: You will get **$1 if you send Message BLUE** and **$1 if you send Message GREEN.**
I choose to send (please CIRCLE ONE)

MESSSAGE BLUE MESSAGE GREEN

SCENARIO 2: You will get **$1 if you send Message BLUE** and **$2 if you send Message GREEN.**
I choose to send (please CIRCLE ONE)

MESSSAGE BLUE MESSAGE GREEN

Please answer the following questions for us:
1) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:
0-10% 10-25% 25-50% 50-75% 75-90 90-100%

2) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $2 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

3) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT

Appendix C. Chapter 4 Experiment Instructions

C.1 First-Round Experiment: Direct Treatment Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Receiver." For purposes of this experiment, **YOU ARE A "SENDER."**

• Neither of you will ever know the identity of the other.

• You and your Receiver will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. Both of you will be paid for ONE of the two situations. The situation for which you will be paid is determined by a flip of a coin after all decisions have been made by all participants. YOU SHOULD THEREFORE MAKE YOUR DECISION IN EACH SITUATION AS IF IT IS THE ONE FOR WHICH YOU WILL BE PAID.

Experiment Instructions (continued)

Situation K

• Two possible monetary payments are available to you and your Receiver in this situation:

Option C: $5 to you and $6 to the other student (your Receiver).
Option D: $7 to you and $3 to the other student (your Receiver).

• You know what the Options are, but your Receiver **DOES NOT**.

• You will choose between **ONE** of two possible messages, which will be sent to your Receiver:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

• Based **ONLY** on your message, your Receiver will **CHOOSE ONE** of the two Options. Your Receiver's choice will determine the payments in the experiment. Receivers will not know the true Options, or the sums to be paid to you under the different Options.

• **Information from Past Sessions**:

• In past sessions of this experiment, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• NOW PLEASE MAKE YOUR DECISION:

• **YOUR DECISION**

****** I choose to send the following message (please circle ONE): ******

Message C **Message D**

Experiment Instructions (continued)

Situation K (continued)

• Please answer the following questions for us:

1) Which option do you think your Receiver will choose based upon the message you sent? If you correctly predict your Receiver's choice, we will pay you an additional $1.

I believe my Receiver will choose (please CIRCLE ONE)

Option C Option D

Situation L

• In this situation, only you have to make a decision. Your decision will determine the payments to both you and your Receiver.

• Two possible monetary payments are available:

Option R: $5 to you and $6 to the other student (your Receiver).
Option S: $7 to you and $3 to the other student (your Receiver).

• Based on your decision, the computer will determine payments according to your chosen option with 80% probability, and according to the other option with 20% probability.

• You know what the Options are, but your Receiver **DOES NOT**. Your Receiver only knows that his/her (anonymous) Sender will make a choice and, based on that choice, the computer will determine payments to both of you.

• **YOUR DECISION:** **I choose (please CIRCLE ONE)**

OPTION R **OPTION S**

FINAL QUESTION:

2) Please can tell us your gender (please circle one): **Male** **Female**

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.2. First-Round Experiment: Delegated Treatment Instructions (Blue Dot Version)

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

NOW PLEASE PROCEED TO THE EXPERIMENTS ON THE NEXT PAGE…

Experiment 0

In another classroom, the following short experiment was conducted with **OTHER STUDENTS**. PLEASE READ THIS DESCRIPTION. YOU WILL NEED this information in Experiment 1.

There is a colored dot at the top of this page. What color is the dot?_____

There are two Messages that can be sent about the dot to someone who cannot see the dot and does not know its color:

Message GREEN: The dot is green.
Message BLUE: The dot is blue.

After seeing exactly the same dot as you, the **OTHER STUDENTS** could
• Choose **Message GREEN** and **RECEIVE $2**,
OR
• Choose **Message BLUE** and **RECEIVE $1**.

Based on their choices, we will call the **OTHER STUDENTS**
• **$2 Senders** if they earned $2 by choosing **Message GREEN.**

• **$1 Senders** if they earned $1 by choosing **Message BLUE**.

Roughly half of the **OTHER STUDENTS** were **$2 Senders**, and roughly half were **$1 Senders**.

Please answer the following:
What type of Senders earned more money by sending Message GREEN? (**PLEASE CIRCLE ONE**.)

$1 Senders **$2 Senders**

Experiment 1
• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Receiver." For purposes of this experiment, **YOU ARE A "SENDER."**

• Neither of you will ever know the identity of the other.

• You and your Receiver will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. Both of you will be paid for ONE of the two situations. The situation for which you will be paid is determined by a flip of a coin after all decisions have been made by all participants. YOU SHOULD THEREFORE MAKE YOUR DECISION IN EACH SITUATION AS IF IT IS THE ONE FOR WHICH YOU WILL BE PAID.

Experiment 1 Instructions (continued)
Situation K

• Two possible monetary payments are available to your Receiver and you (the SENDER) in this situation:

Option C: $7 to you and $3 to your Receiver.
Option D: $5 to you and $6 to your Receiver.

• You (and other Senders) know what the Options are, but the Receivers **DO NOT**.
• **ONE** of two possible messages will be sent to your Receiver:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

• Based **ONLY** on the Message, your Receiver will **CHOOSE ONE** of the two Options. Your Receiver's chosen Option will determine payments to both of you. Receivers will not know the true Options, or the sums to be paid under the different Options.

• In past sessions of this experiment, roughly **8 out of 10 Receivers chose the Option recommended in the Message (for example, Option C if Message C was sent)**.

• In another session of this experiment, **OTHER STUDENTS** had exactly the same options as above, and each chose whether to send Message C or D to his or her **OWN** Receiver.

• In **THIS** experiment, you will NOT CHOOSE a Message. Instead, **YOUR** Receiver will be sent the Message chosen by one of the **OTHER STUDENTS** for his or her **OWN** Receiver. For example, if the OTHER STUDENT sent Message C to his/her OWN Receiver, Message C will be sent to YOUR Receiver. **YOU** will make a **DECISION** that determines how this **OTHER STUDENT** is selected.

• There are **TWO TYPES OF SENDERS** to choose between, based upon decisions in Experiment 0:

$2 Senders earned $2 by choosing **Message GREEN**.
$1 Senders earned $1 by choosing **Message BLUE**.

• The Message to be sent to **YOUR** Receiver is the Message chosen by a randomly selected **STUDENT** from **ONE OF THESE TWO GROUPS**. **YOU WILL CHOOSE** whether this **OTHER STUDENT** is selected from the **$1 Senders OR** from the **$2 Senders**.

• **YOUR DECISION:**

****** I choose to have the "OTHER STUDENT" selected from (please CIRCLE ONE): ******

$1 Senders **$2 Senders**

Experiment 1 Instructions (continued)

Situation L

• In this situation, your Receiver does not make any decision.

• Two possible monetary payments are available:

Option R: $7 to you and $3 to the other student (your Receiver).
Option S: $5 to you and $6 to the other student (your Receiver).

• One of these two options will be CHOSEN. The computer will determine payments according to the CHOSEN option with 80% probability, and according to the other option with 20% probability.

• You know what the Options are, but your Receiver **DOES NOT**. Your Receiver only knows that someone else will make a choice and, based on that choice, the computer will determine payments to both of you.

• In **ANOTHER** session of this experiment, each SENDER **CHOSE** one of the two above OPTIONS (for him/her SELF and his/her OWN Receiver).

• In **THIS** experiment, we will implement the **CHOSEN OPTION of ANOTHER STUDENT** from the other session. For example, if the OTHER STUDENT chose Option R, we will implement Option R (with 80 percent probability) for you and your Receiver. Your "**OTHER STUDENT**" will be randomly selected from **ONE OF TWO GROUPS** (based on decisions in Experiment 0):

$2 Senders earned $2 by choosing **Message GREEN.**
$1 Senders earned $1 by choosing **Message BLUE.**

• **YOU WILL CHOOSE** whether your **"OTHER STUDENT"** is selected from the **$1 Senders OR** from the **$2 Senders**.

• **YOUR DECISION: **** I choose to have the "OTHER STUDENT" chosen from (please CIRCLE ONE): **

$1 Senders **$2 Senders**

Experiment 1 Instructions (continued)

Please answer the following questions. In each question, you are asked to make a prediction; you will be paid an ADDITIONAL $1 for **EACH** CORRECT PREDICTION.

1) In Situation K, do you predict that your Receiver will choose the Option recommended in the Message (for example, Option C if Message C was sent)? PLEASE CIRCLE ONE ANSWER:

YES	NO
(My Receiver will choose the recommended option)	(My Receiver will NOT choose the recommended option)

2) In Situation K, what percent of **$1 Senders** do you think will send the Untruthful Message C (that earns the Sender more money)? PLEASE CIRCLE ONE:

0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90 90-100

3) In Situation K, what percent of **$2 Senders** do you think will send the Untruthful Message C (that earns the Sender more money)? PLEASE CIRCLE ONE:

0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90 90-100

4) In Situation L, what percent of **$1 Senders** do you think will choose the Option R (that earns the Sender more money)? PLEASE CIRCLE ONE:

0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90 90-100

5) In Situation L, what percent of **$2 Senders** do you think will choose the Option R (that earns the Sender more money)? PLEASE CIRCLE ONE:

0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90 90-100

FINAL QUESTION:

6) Please can tell us your gender (please circle one):

Male **Female**

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.3. First-Round Experiment: Receiver Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Sender." For purposes of this experiment, you are a "Receiver."

• Neither of you will ever know the identity of the other.

• The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

• You and your Sender will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. Both of you will be paid for ONE of the two situations. The situation for which you will be paid is determined by a flip of a coin after all decisions have been made by all participants. YOU SHOULD THEREFORE MAKE YOUR DECISION IN EACH SITUATION AS IF IT IS THE ONE FOR WHICH YOU WILL BE PAID.

• You and your Sender will be paid for situation ***K*** if the coin toss comes up **Heads**.
• You and your Sender will be paid for situation ***L*** if the coin toss comes up **Tails**.

Situation K

• Two possible monetary payment Options (A and B) are available to you and your Sender in this situation.

• We showed the two payment options to your Sender.

• **YOU** will choose **ONE** of the two options, which will determine the payments to the two of you. The ONLY information you will have is a message sent to you on behalf of your Sender.

• Two possible messages could be sent to you:

Message A: "Option A will earn you more money than Option B."
Message B: "Option B will earn you more money than Option A."

• The following Message is sent to you on behalf of your Sender: ____

• We now ask you to choose ONE of the two options. **(PLEASE CIRCLE ONE)**

Option A **Option B**

Situation L

• In this situation, you are a passive participant, and your payment is determined by the computer based on your Sender's decision.

C.4. First Round Experiment: Prior Experiment Sender Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

Experiment 1

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Receiver." For purposes of this experiment, **YOU ARE A "SENDER."**

• Neither of you will ever know the identity of the other.

• You and your Receiver will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. Both of you will be paid for ONE of the two situations. The situation for which you will be paid is determined by a flip of a coin after all decisions have been made by all participants. YOU SHOULD THEREFORE MAKE YOUR DECISION IN EACH SITUATION AS IF IT IS THE ONE FOR WHICH YOU WILL BE PAID.

Experiment 1 Instructions (continued)

Situation K

• Two possible monetary payments are available to you and your Receiver in this situation:

Option C: $5 to you and $6 to the other student (your Receiver).
Option D: $7 to you and $3 to the other student (your Receiver).

• You know what the Options are, but your Receiver **DOES NOT**.

• You will choose between **ONE** of two possible messages, which will be sent to your Receiver:

Message C: "Option C will earn you (the Receiver) more money than Option D."
Message D: "Option D will earn you (the Receiver) more money than Option C."

• Based **ONLY** on your message, your Receiver will **CHOOSE ONE** of the two Options. Your Receiver's choice will determine the payments in the experiment. Receivers will not know the true Options, or the sums to be paid to you under the different Options.

• **Information from Past Sessions**:

• In past sessions of this experiment, roughly **8 out of 10 Receivers chose the Option recommended by their Senders**.

• NOW PLEASE MAKE YOUR DECISION:

• **YOUR DECISION**

****** I choose to send the following message (please circle ONE): ******

Message C **Message D**

Experiment 1 Instructions (continued)

Situation K (continued)

• Please answer the following questions for us:

1) Which option do you think your Receiver will choose based upon the message you sent? If you correctly predict your Receiver's choice, we will pay you an additional $1.

I believe my Receiver will choose (please CIRCLE ONE)

Option C Option D

Situation L

• In this situation, only you have to make a decision. Your decision will determine the payments to both you and your Receiver.

• Two possible monetary payments are available:

Option R: $5 to you and $6 to the other student (your Receiver).
Option S: $7 to you and $3 to the other student (your Receiver).

• Based on your decision, the computer will determine payments according to your chosen option with 80% probability, and according to the other option with 20% probability.

• You know what the Options are, but your Receiver **DOES NOT**. Your Receiver only knows that his/her (anonymous) Sender will make a choice and, based on that choice, the computer will determine payments to both of you.

• **YOUR DECISION:** ****I choose (please CIRCLE ONE)**

OPTION R **OPTION S**

Experiment 2

In this second short experiment, you will be paid for ONE of two SCENARIOS. The scenario for which you will be paid will be determined by a coin flip.

There is a colored dot at the top of **THIS** page. What color is the dot?_____
Here are two Messages that you can send about the dot to someone who cannot see the dot and does not know its color.
Message GREEN: The dot is green.
Message BLUE: The dot is blue.
For each of the following scenarios, please circle the Message you choose to send.

SCENARIO 1: You will get **$1 if you send Message GREEN** and **$1 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

SCENARIO 2: You will get **$1 if you send Message GREEN** and **$2 if you send Message BLUE.**
I choose to send (please CIRCLE ONE)

MESSSAGE GREEN MESSAGE BLUE

Please answer the following questions for us:
2) What proportion of students in this Session do you think will choose to send the Truthful Message in Scenario 1 (when they receive $1 regardless)? PLEASE CIRCLE ONE:

0-10% 10-25% 25-50% 50-75% 75-90 90-100%

3) What proportion of students in this Experiment 2 do you think will choose to send the Truthful Message in Scenario 2 (when they receive $1 for a Truthful Message and $2 for an Untruthful Message)? If your prediction is correct (within five percentage points of the actual percentage, plus or minus), you will receive an additional $1 payment. PLEASE CIRCLE ONE:

0-5 5-10 10-15 15-20 20-25 25-30 30-35 35-40 40-45 45-50

50-55 55-60 60-65 65-70 70-75 75-80 80-85 85-90 90-95 95-100

4) Please can tell us your gender (please circle one): **Male**
Female

THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A CLASS SESSION NEXT WEEK. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.5. Second-Round Experiment: Direct Deception Treatment Instructions (Black Lie Version)

Instructions **Your Registration Number:_____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENT ON THE NEXT PAGE…

Experiment 1

• You have been randomly matched with another student from a different session of this experiment, whom we will call your **RECEIVER**. In this experiment, **YOU ARE A SENDER.** Neither of you will ever know the identity of the other.

• The two of you will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. We will randomly select ONE of the two situations for payment; Situation K will be selected with (2/3) probability and Situation L with (1/3) probability.

• Both situations revolve around the colored dot at the top of this page. What color is the dot?_____

• Your Receiver will NOT KNOW the color of the dot. However, **ONE** of two possible messages will be sent to your Receiver on your behalf:

Message GREEN: I solemnly swear that the dot is GREEN.
Message BLUE: I solemnly swear that the dot is BLUE.

• Your Receiver will be told that the Message comes from a Sender who observes the true color of the dot. Based **ONLY** on the Message, your Receiver will be asked to **REPORT** the true color of the dot, whether GREEN or BLUE. Your Receiver's chosen Report will determine payments to both of you:

If he/she **REPORTS GREEN**, then payments will be: **$7 to YOU** and **$3 to HIM / HER**.

If he/she **REPORTS BLUE**, then payments will be: **$5 to YOU** and **$5 to HIM / HER**.

• Only you know what the Payment possibilities are. Your Receiver will have NO information about the sums to be paid in these different cases. He or she will only be told that their Report determines payments and that you have been informed of the payments connected to each Message.

• NOTE: In similar experiments, Receivers almost always Report according to the Message they receive.

SITUATION K

• In this Situation, you CHOOSE which Message to send, Message GREEN or Message BLUE:

• **YOUR DECISION:**

****** I choose to SEND THE FOLLOWING MESSAGE to my Receiver (**please CIRCLE ONE**): ******

Message GREEN	**Message BLU**E
("I solemnly swear that the dot is GREEN")	("I solemnly swear that the dot is BLUE")

SITUATION L

• In this Situation, no decision is required of you and Message BLUE will automatically be sent.

FINAL QUESTION:

Can you tell us your gender (please circle one): **Male** **Female**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.6. Second-Round Experiment: Delegated Deception Treatment (Black Lie Version)

Instructions **Your Registration Number:____ _____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENTS ON THE NEXT PAGE…

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your **RECEIVER**. In this experiment, **YOU ARE A SENDER** and there are also other students who we will call **AGENTS**. None of you will ever know the identities of the others.

• There is a colored dot at the top of this page. What color is the dot?_____

• Your Receiver will NOT KNOW the color of the dot. **ONE** of two possible messages will be sent to your Receiver from a student who observes exactly the same color dot as you:

Message GREEN: I solemnly swear that the dot is GREEN.
Message BLUE: I solemnly swear that the dot is BLUE.

• Your Receiver will be informed that the Message comes from a student who observes the true color of the dot. Based **ONLY** on the Message, your Receiver will be asked to **REPORT** the true color of the dot, whether GREEN or BLUE. Your Receiver's chosen Report will determine payments to both of you:

If he/she **REPORTS GREEN**, then payments will be:**$7 to YOU and $3 to HIM/ HER**.
If he/she **REPORTS BLUE**, then payments will be:**$5 to YOU and $5 to HIM / HER**.

• Your Receiver will have NO information about the sums to be paid in these different cases. He or she will only be told that their Report determines payments and that Senders and Agents are informed of the payments connected to each Message.

• NOTE: In similar experiments, Receivers almost always Report according to the Message they receive.

• In another session of this experiment, a group of other students – the AGENTS – saw exactly the same color dot as you and also saw the payments connected to each Message, as described above. EACH Agent is associated with 3 Messages. The 1st and 2nd are the Message CHOSEN by the Agent – *both* are *either* Message GREEN or Message BLUE. The 3rd is MESSAGE BLUE (assigned by us).

• In **THIS** experiment, **YOUR** Receiver will be sent ONE of the three Messages of ONE of the AGENTS. You will make a decision that determines which student is YOUR AGENT. The Agents are not paid based on the payments attached to the Messages, but rather according to how many Senders (like you) choose them as their own Agent.
• You have two Agents to choose between:
Agent 1 chose GREEN Messages (so 2 GREEN and 1 BLUE Message).
Agent 2 chose BLUE Messages (so 3 BLUE Messages).

• The Message to be sent to **YOUR** Receiver is ONE of **your** Agent's THREE Messages – each chosen with equal (one-third) probability. **You decide which Agent is YOURS**: Agent 1 OR Agent 2.

Experiment Instructions (continued)

• **YOUR DECISION:**
**I choose the following student to be MY AGENT (please CIRCLE ONE): **

AGENT 1 **AGENT 2**

ONE FINAL QUESTION:

Can tell us your gender (please circle one): **Male** **Female**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.7. Second-Round Experiment: Agent Instructions

Instructions **Your Registration Number:_____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENT ON THE NEXT PAGE…

Experiment

• In this experiment, you are an AGENT. You will be paid based on the number of students in other sessions of this experiment who choose YOU to be ***their*** AGENT. There are THREE different situations in which you make decisions. Based on your decisions, you will attract students who choose you as their Agent. We will give each student, in each Situation, a choice between TWO different Agents who have made different decisions on behalf of the students who select them. You will be paid $.50 (50 cents) for each student who selects you as their AGENT in each Situation. We expect payments to be about $10-$20, with the possibility for higher payments.

SITUATIONS 1-2: THE DOT

• Two of the 3 Situations revolve around the dot at the top of this page. What color is the dot?_____

• In each of these Situations, there are two types of players (other than AGENTS like you). There are SENDERS – the students who will be selecting AGENTS like you. And there are RECEIVERS. Each Sender is matched with a single Receiver. A subset of the Senders (those who can choose you as their Agent) have observed exactly the same color dot as you. Their Receivers will NOT KNOW the color of the dot. YOU will be choosing Messages that can be sent to each Sender's assigned Receiver. There are TWO possible Messages:

Message GREEN: I solemnly swear that the dot is GREEN.
Message BLUE: I solemnly swear that the dot is BLUE.

• Each Receiver will be sent one of these Messages and will be informed that the Message comes from a student who observes the true color of the dot. Based **ONLY** on the Message, the Receiver will be asked to **REPORT** the true color of the dot, whether GREEN or BLUE. The Receiver's chosen Report will determine payments to the Receiver and his/her matched Sender. Senders and Agents (like you) know the Payment possibilities. Receivers will have NO information about the sums to be paid in these different cases. They will only be told that their Report determines payments and that Senders and Agents are informed of the payments connected to each Message.

• NOTE: In similar experiments, Receivers almost always Report according to the Message they receive.

• Payments vary between the TWO Situations:

Situation 1:

If the Receiver **REPORTS GREEN**, then payments will be: **$6 to the SENDER and $6 to the RECEIVER**.
If the Receiver **REPORTS BLUE**, then payments will be: **$5 to the SENDER and $5 to the RECEIVER**.

Situation 2:

If the Receiver **REPORTS GREEN**, then payments will be: **$7 to the SENDER and $3 to the RECEIVER**.
If the Receiver **REPORTS BLUE**, then payments will be: **$5 to the SENDER and $5 to the RECEIVER**.

Experiment (cont.)

• For EACH Situation, you will be associated with THREE Messages. You choose the first two, but BOTH must be the SAME: either both are the GREEN Message OR both are the BLUE Message. Message BLUE will automatically be assigned as your 3rd Message.

• Senders will each be given a choice between TWO Agents who make different decisions – that is, between Agents who choose Message GREEN vs. Message BLUE. Once a Sender chooses an Agent, a Message is randomly chosen for that Sender from among the chosen Agent's 3 Messages, and that Message will be sent to the corresponding Receiver.

• You will be matched with Senders based on your choices and those of other Agents. For example, if there are 3 Agents who choose the GREEN Message, and 30 student Senders, then each of the 3 Agents will be one of the two alternatives offered to 10 students (30 Senders, divided by 3 Agents). Similarly, if 2 agents choose the BLUE Message, then each of the 2 agents will be the "BLUE Message Agent" offered to 15 students. It is therefore advantageous to make choices **different from other Agents** (in order to be on offer more often), but equally advantageous to make choices that **Senders are likely to prefer** (so that you are more likely to be selected out of the Agents offered to different Senders).

• We now ask you to make your Message choice for each of the three Situations. The Table below summarizes the payments to Sender and Receiver in each of the Situations.

SITUATION 1: For my 1st and 2nd Message, I choose the following to be sent on behalf of Senders who choose me as their Agent: **PLEASE CIRCLE ONE**

Message GREEN **Message BLU**E

SITUATION 2: For my 1st and 2nd Message, I choose the following Message to be sent on behalf of Senders who choose me as their Agent: **PLEASE CIRCLE ONE**

Message GREEN **Message BLU**E

Table: Payments in Situations 1-2
Payments to (Sender, Receiver)

Receiver Report →	GREEN	BLUE
Situation 1	($6,$6)	($5,$5)
Situation 2	($7,$3)	($5,$5)

Experiment (cont.)

Situation 3

• In this Situation, there are two possible Payment combinations for Sender and Receiver:

Option A: $7 to the Sender and $3 to the Receiver.
Option B: $5 to the Sender and $5 to the Receiver.

• Senders and Agents know what the Options are, but Receivers **DO NOT**. Receivers only know that someone else will make a choice and, based on that choice, the computer will determine payments to Sender and Receiver.

• As before, you (and other Agents) are each associated with 3 decisions. The 1st and 2nd are YOUR OWN OPTION CHOICE (one selection), either Option A or Option B. The 3rd is determined automatically (by us) as Option B. Senders are given a choice between TWO Agents who make different decisions – that is, between Agents who choose Option A vs. Option B. Once a Sender chooses an Agent, ONE Option is randomly chosen for that Sender from among the chosen Agent's 3 Option determinations.

• You will be matched with Senders based on your choices and those of other Agents in the same way as in Situations 1-2.

SITUATION 3: For my 1st and 2nd Option, I choose **PLEASE CIRCLE ONE**

Option A **Option B**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.8. Second-Round Experiment: Direct Deception Treatment Receiver Instructions

Instructions **Your Registration Number:_____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

Experiment

• In this short experiment, you are randomly matched with another participant in a different session of this experiment, whom we will call your SENDER. You are a RECEIVER. Neither of you will ever know the identity of the other.

• We have shown your Sender a COLORED DOT that is either BLUE or GREEN.

• After seeing the color, your Sender chose one of two possible Messages:

Message GREEN: I solemnly swear that the dot is **GREEN**.
Message BLUE: I solemnly swear that the dot is **BLUE**.

• The Message sent to you on behalf of your Sender is as follows:

I solemnly swear that the dot is GREEN BLUE

• Based only on this Message, we ask you to REPORT the COLOR of the dot, whether it is GREEN or BLUE. Your choice will determine the payments to you and your Sender according to two different options (option A and option B). If you report the actual color shown to your Sender, then we will implement payment option A; if you report the other color, then we will implement payment option B. Your Sender was told the monetary values connected to each Report.

• **YOUR DECISION:**

** I choose to REPORT that the shape is a (PLEASE CIRCLE ONE)**

GREEN **BLUE**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.9. Second-Round Experiment: Delegated Deception Treatment Receiver Instructions

Instructions **Your Registration Number:____ _____**

Introduction: Your experiment manager will introduce you to the experiment.

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

PLEASE ANSWER THE QUESTIONS IN SEQUENCE AS YOU READ THEM.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

You will earn some MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENT ON THE NEXT PAGE…

Experiment 1

• In this short experiment, you are randomly matched with another participant in a different session of this experiment, whom we will call your SENDER. You are a RECEIVER. Neither of you will ever know the identity of the other.

• We have shown your Sender a COLORED DOT that is either BLUE or GREEN. A few other students were shown a dot with exactly the SAME COLOR as shown to your Sender.

• After seeing the color, each of the other students chose between two possible Messages:

Message GREEN: I solemnly swear that the dot is **GREEN**.
Message BLUE: I solemnly swear that the dot is **BLUE**.

• ONE of these Message choices has been randomly selected to be sent to you on behalf of your Sender. It is as follows:

I solemnly swear that the dot is GREEN BLUE

• Based only on this Message, we ask you to REPORT the COLOR of the dot, whether it is GREEN or BLUE. Your choice will determine the payments to you and your Sender according to two different options (option A and option B). If you report the actual color shown to your Sender, then we will implement payment option A; if you report the other color, then we will implement payment option B. Your Sender and the other students were told the monetary values connected to each Report.

• **YOUR DECISION:**

** I choose to REPORT that the shape is a (PLEASE CIRCLE ONE)**

GREEN **BLUE**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! <u>PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER</u>. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.10. Second-Round Experiment: Direct Dictator Treatment Instructions

Instructions **Your Registration Number:____ _____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENTS ON THE NEXT PAGE…

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Co-Participant." Neither of you will ever know the identity of the other.

• The two of you will participate in two different decision-making situations, which we identify by ***K*** and ***L*** below. We will randomly select ONE of the two situations for payment; Situation ***K*** will be selected with (2/3) probability and Situation ***L*** with (1/3) probability.

• In both situations, there are two possible monetary payments:

Option A: $7 to you and $3 to the other student (your Co-Participant).
Option B: $5 to you and $5 to the other student (your Co-Participant).

• One of these options will be CHOSEN. The computer will determine payments according to the CHOSEN OPTION with 80% probability, and according to the other option with 20% probability.

• You know what the Options are, but your Co-Participant **DOES NOT**. Your Co-Participant only knows that someone else will make a choice that determines payments to both of you.

Situation K

• In this Situation, YOU CHOOSE one of the Options.

• **YOUR DECISION: **** I choose (please CIRCLE ONE): ******

Option A **Option B**

Situation L

• In this Situation, no decision is required of you and Option B will be chosen automatically.

ONE FINAL QUESTION:

Can tell us your gender (please circle one): **Male**
Female

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.11. Second-Round Experiment Delegated Dictator Treatment Instructions

Instructions **Your Registration Number:____ _____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING!

Experiment

• You have been randomly matched with another student from a different session of this experiment, whom we will call your "Co-Participant." There are also other students who we will call **AGENTS**. None of you will ever know the identities of the others.

• In this experiment, there are two possible monetary payments:

Option A: \$7 to you and \$3 to the other student (your Co-Participant).
Option B: \$5 to you and \$5 to the other student (your Co-Participant).

• One of these options will be CHOSEN. The computer will determine payments according to the CHOSEN OPTION with 80% probability, and according to the other option with 20% probability.

• You know what the Options are, but your Co-Participant **DOES NOT**. Your Co-Participant only knows that someone else will make a choice that determines payments to both of you.

• In another session of this experiment, a group of other students – the AGENTS – saw exactly the same payment options. EACH Agent is associated with 3 Option selections. The 1st and 2nd are the Agent's CHOICE – *both* are *either* Option A or Option B. The 3rd is Option B (assigned by us).

• In **THIS** experiment, we will randomly choose ONE of the option choices made by ONE of the Agents; this option will determine payments for you and your Co-Participant (with 80% probability). You will make a decision that determines which student is YOUR AGENT. The Agents are not paid based on the payment options, but rather according to how many Senders (like you) choose them as their own Agent.

• You have two Agents to choose between:

Agent 1 chose Option A (so 2 Option A and 1 Option B choice).
Agent 2 chose Option B (so 3 Option B choices).

• The **CHOSEN Option** for you and your Co-Participant will be ONE of the 3 selections of YOUR AGENT – each chosen with equal (one-third) probability. **You decide which Agent is YOURS**: Agent 1 OR Agent 2.

• **YOUR DECISION:**
**I choose the following student to be MY AGENT (please CIRCLE ONE): **

AGENT 1 **AGENT 2**

ONE FINAL QUESTION:

Can you tell us your gender (please circle one):

Male **Female**

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

C.12. Plus-One Experiment Instructions

Instructions **Your Registration Number:____ _____**

Introduction: Welcome to our short experiment!

IF YOU HAVE QUESTIONS AT ANY TIME, PLEASE RAISE YOUR HAND AND ONE OF US WILL HELP YOU.

PLEASE DO NOT COMMUNICATE WITH OTHER PARTICIPANTS UNTIL THE END OF THE SESSION.

PLEASE READ ALL INSTRUCTIONS CAREFULLY AND ANSWER ALL QUESTIONS.

The decisions that you make in this experiment will be COMPLETELY ANONYMOUS.

YOU WILL EARN SOME MONEY in the experiments to follow. The money that you earn will be paid to you at a future time to be announced by your session manager. The payment will be made in cash and in confidence.

Your registration number is given on the attached tag and will be used to identify you for payment. PLEASE TEAR OFF AND KEEP THE ATTACHED TAG, AND BRING IT WITH YOU TO THE FUTURE PAYMENT SESSION NEXT WEEK to be announced by your experiment manager.

THANK YOU FOR PARTICIPATING! NOW PLEASE PROCEED TO THE EXPERIMENT ON THE NEXT PAGE.

Experiment

In this experiment, you (player A) will be matched with another student in a different classroom (player C). **Each of you obtains a base payment of $3** for this experiment. Based on a decision that you make, this initial set of payments can be changed in the following way:

Payment Change: You each can earn $1 more, for total payments of **$4 to YOU and $4 to PLAYER C**.

You will also be matched with two other students, players B and D, who are in a similar situation as you and player C. Both players B and D also get a base payment of $3 each, and player B decides whether to make the Payment Change, so that B and D each earn $1 more (total of $4 each).

Once this experiment is over, we will calculate the percentage of *all* "player B's" who chose to make the Payment Change. We will call this percentage Q. For example, if 9 in 10 player B's make the Payment Change, then Q will be 90 percent.

YOU have three options:

Option 1: Implement PLAYER B's decision.
In this case, if your matched player B chooses to make the Payment Change for himself/herself and his/her player D, then the Payment Change will be made for you and player C as well. Likewise, if your player B chooses not to make the change, then the Payment Change will not be implemented for you and player C.

Option 2: YOU choose to MAKE the Payment Change directly ($1 more for you and C).
In this case, the Payment Change will be implemented for you and player C with probability Q.

Option 3: DO NOT make the Payment Change (so that you and C each obtain the initial $3).

YOUR DECISION: I choose ***PLEASE CIRCLE ONE***

Option 1	**Option 2**	**Option 3**
(Implement PLAYER B's decision)	(MAKE the Payment Change directly)	(Do NOT make the Payment Change)

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

Experiment 1 (Reverse)

In this experiment, you (player A) will be matched with another student in a different classroom (player C). **Each of you obtains a base payment of $3** for this experiment. Based on a decision that you make, this initial set of payments can be changed in the following way:

Payment Change: You each can earn $1 more, for total payments of **$4 to YOU and $4 to PLAYER C**.

You will also be matched with two other students, players B and D, who are in a similar situation as you and player C. Both players B and D also get a base payment of $3 each, and player B decides whether to make the Payment Change, so that B and D each earn $1 more (total of $4 each).

Once the experiment is over, we will calculate the percentage of player B's who chose to make the Payment Change. We will call this percentage Q. For example, if 9 in 10 player B's make the Payment Change, then Q will be 90 percent.

YOU have three options:

Option 1: YOU choose to MAKE the Payment Change directly ($1 more for you and C). *In this case, the Payment Change will be implemented for you and player C with probability Q.*

Option 2: DO NOT make the Payment Change (so that you and C each obtain the initial $3).

Option 3: Implement PLAYER B's decision.
In this case, if your matched player B chooses to make the Payment Change for himself/herself and his/her player D, then the Payment Change will be made for you and player C as well. Likewise, if your player B chooses not to make the change, then the Payment Change will not be implemented for you and player C.

YOUR DECISION: I choose ***PLEASE CIRCLE ONE***

Option 1	**Option 2**	**Option 3**
(MAKE the Payment Change directly)	(Do NOT make the Payment Change)	(Implement PLAYER B's decision)

THAT'S IT! THANK YOU FOR YOUR PARTICIPATION!! PLEASE TURN IN YOUR QUESTIONNAIRE TO THE EXPERIMENT MANAGER. PAYMENTS WILL BE MADE FOR THESE EXPERIMENTS AT A FUTURE CLASS SESSION. PLEASE BRING YOUR TAG TO THE CLASS FOR PAYMENT.

Appendix D: The Hurkens-Kartik Statistic

Let p_k = mean of Bernoulli distribution k, q_k = sample mean from distribution k, and n_k = number of observations for sample k. We have four distributions and four corresponding independent random samples, k=1 (direct deception), k=2 (delegated deception), k=3 (direct dictator), and k=4 (delegated dictator). In each case, the Bernoulli observation takes a value of one if the choice is untruthful (selfish), and zero if the choice is truthful (generous). The null hypothesis is that the ratio of untruthful to selfish for delegated decisions equals the ratio of untruthful to selfish for direct decisions:

H_0: $(p_1/p_3)-(p_2/p_4)=0 \leftrightarrow (p_1p_4)-(p_2p_3)=0$
H_1: $(p_1p_4)-(p_2p_3) \neq 0$

A few preliminary observations aid derivation of the test statistic for this null.
(i) $E(q_kq_m) = p_k\, p_m$ and $E(q_k^2\, q_m^2) = E(q_k^2)E(q_m^2)$ for $k \neq m$ by independence of observations;
(ii) $E(q_k^2) = E\{ [\sum_i x_i^2 + \sum_i \sum_{i \neq j} x_i x_j] / n_k^2 \} = (1/n_k)[p_k + (n_k-1)p_k^2]$ where x_i and x_j are (Bernoulli) observations from sample k and the second equality follows from independence of observations in the random sample ($E(x_ix_j)= p_k^2$) and $E(x_i^2)=p_k$ for observation i in sample k;
(iii) for $k \neq m$, and using properties (i) and (ii),

$$V(q_kq_m) = E\{ (q_kq_m - E(q_kq_m))^2 \} = E\{(q_kq_m)^2 \} - E\{ q_kq_m\}^2$$
$$= E(q_k^2)E(q_m^2) - p_k^2p_m^2$$
$$= p_k\, p_m \{1 + (n_k-1)\, p_k + (n_m-1)\, p_m - (n_k+n_m - 1)\, p_k\, p_m \}/ (n_k\, n_m)$$

(iv) by independence of the random samples, the variance of the difference in sample mean products is

$$V(q_1q_4 - q_2q_3) = V(q_1q_4) + V(q_2q_3)$$

By these properties, the Central Limit Theorem and the Law of Large Numbers, the following statistic is approximately distributed as a standard normal under the null:

$$t_0 = \{(q_1q_4)-(q_2q_3)\} / \{[W(q_1q_4)+ W(q_2q_3)]^{1/2}\} \overset{a}{\sim} N(0,1)$$ under H_0,
where W = estimated $V = q_k\, q_m \{1 + (n_k-1)\, q_k + (n_m-1)\, q_m - (n_k+n_m-1)\, q_k\, q_m \}/ (n_k\, n_m)$ for (k=1,m=4) and (k=2,m=3).