White lies

Erat, S.; Gneezy, U.

DOI
10.1287/mnsc.1110.1449

Publication date
2012

Document Version
Final published version

Published in
Management Science

MANAGEMENT SCIENCE
Vol. 58, No. 4, April 2012, pp. 723–733
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1110.1449

White Lies

Sanjiv Erat, Uri Gneezy
Rady School of Management, University of California, San Diego, La Jolla, California 92093
{serat@ucsd.edu, ugneezy@ucsd.edu}

In this paper we distinguish between two types of white lies: those that help others at the expense of the person telling the lie, which we term *altruistic white lie*s, and those that help both others and the liar, which we term *Pareto white lies*. We find that a large fraction of participants are reluctant to tell even a Pareto white lie, demonstrating a pure lie aversion independent of any social preferences for outcomes. In contrast, a non-negligible fraction of participants are willing to tell an altruistic white lie that hurts them a bit but significantly helps others. Comparing white lies to those where lying increases the liar's payoff at the expense of another reveals important insights into the interaction of incentives, lying aversion, and preferences for payoff distributions. Finally, in line with previous findings, women are less likely to lie when it is costly to the other side. Interestingly though, we find that women are more likely to tell an altruistic lie.

Key words: deception; lies; cheap talk; experiments
History: Received June 14, 2010; accepted July 1, 2011, by Teck Ho, decision analysis. Published online in *Articles in Advance* November 4, 2011.

1. Introduction

When people communicate, they sometimes lie. Because communication is indispensable in almost any economic and social interaction, understanding when and why people choose to lie is important. A growing body of evidence suggests the decision to lie is sensitive to incentives. People care about their own gain from lying as well as the harm the lie may cause another.[1]

An example displaying the importance of even "small" dishonest behavior is illegal downloads versus purchases in the music industry. Evidence from this industry suggests that millions of people are willing to download music without paying, even when the value is very small—in many cases under $1. One person who downloads music without paying engages in dishonest behavior with very little effect on the industry. Millions of people who do this change the industry profoundly. Such markets often show the presence of both dishonest and honest behavior.

To date, the economics literature has concentrated on contexts like the one above, which we call "selfish black lies," involving acts that help the liar at the expense of another. But sometimes lies may benefit the other person. If a person believes others are likely to misunderstand the truth or that knowledge of the truth is likely to result in disutility to others, then lying may help those other people. Such lies may have a paternalistic flavor in that the liar may believe, to quote Jack Nicholson from the film *All the Good Men*, that others "can't handle the truth."

Consider, for example, a supervisor giving performance feedback to an employee. Should/would the supervisor give truthful feedback to a poorly performing employee, even when such truthful feedback has the potential to reduce the employee's confidence and future performance?

In medicine, a physician may give a placebo to a patient, even if the physician believes the substance has no specific pharmacological effect on the condition being treated. Some moral philosophers argue against such practices even when it helps the patient at "no cost." Saint Augustine (421; 1996, p. 21) wrote, "To me, however, it seems certain that every lie is a sin …." Later philosophers like Kant (1785) took this extreme approach in arguing against all types of lies, regardless of their consequences.

Still, a recent study (Sherman and Hickner 2008) found 45% of doctors had given placebos in clinical practice, but only 4% told patients of the placebos. On one hand, it seems as if a large fraction of doctors may be violating the moral approach cited above. On the other hand, if a doctor believes the placebo has therapeutic value (as 96% of the physicians in the study

[1] For a discussion of the role of incentives in the decision to lie, see Gneezy (2005). See also Boles et al. (2000), Crawford (2003), Brandts and Charness (2003), Croson et al. (2003), Ellingsen and Johnnesson (2004), Cai and Wang (2006), Charness and Dufwenberg (2006), Sánchez-Pagés and Vorsatz (2007), Mazar and Ariely (2006), Mazar et al. (2008), Dreber and Johannesson (2008), Sutter (2009), Lundquist et al. (2009), Ellingsen et al. (2009), Gino and Pierce (2009), Gino and Ayal (2011), and Gintis (2012). For a discussion of the social psychology literature on lies, see DePaulo et al. (1996).

believed), then why not use a placebo if it helps the patient without causing harm?[2]

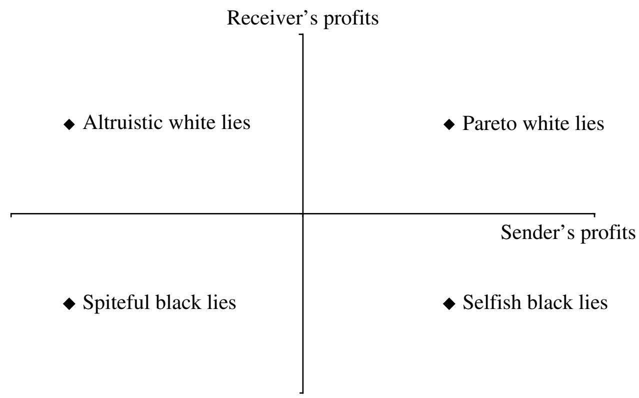
The utilitarian approach (e.g., Bentham 1789) argues that one *should* lie in such situations. When considering whether to lie, a utilitarian would argue, a person should weigh benefits against harm and happiness against unhappiness. The act of lying in itself carries no bad consequences. To quote Martin Luther, "What harm would it do, if a man told a good strong lie for the sake of the good and for the Christian church . . . a lie out of necessity, a useful lie, a helpful lie? Such lies would not be against God; He would accept them" (Bok 1978, p. 47).

These "useful lies" are the focus of the current paper. We wish to expand the discussion in the economics literature to "white lies," or lies that help others. In some cases, a lie can harm the liar but help the other person; we call these "altruistic white lies." People may choose to tell such lies because they care about the other person's payoffs. For example, such lies may create efficiency of the type Charness and Rabin (2002) discuss, if the lie results in a smaller loss to the liar relative to the gain obtained by the other person. Although such a lie is not Pareto improving (in monetary terms), if someone cares enough about the other person and about increasing the total pie, she may choose to lie. The second type of white lies are those that constitute a Pareto improvement, i.e., when both sides earn more as a result of the lie. We call such lies "Pareto white lies." Absent a cost of lying, one would expect people to always tell such lies.

Figure 1 presents our taxonomy of lies based on consequences. The dimensions in the figure are the change in payoffs resulting from a lie. Understanding when people choose to tell white lies is crucial to our understanding of deception. First, people who are reluctant to tell Pareto white lies demonstrate lie aversion independent of social preferences for outcomes. Such people refrain from lying not (merely) because of the consequences, but because they simply view lying as a bad act in itself. This provides the best test of a pure cost of lying in line with a moral stand.

Second, the consequences of black lies where the liar benefits at the other's expense is likely to have very different distributional concerns compared to consequences of white lies where both parties benefit (Fehr and Schmidt 1999). Thus, contrasting white lies with black lies helps explain the interaction between distributional concerns and lying aversion. It also helps identify different types of people. For example, one may expect the motivation to tell a selfish black lie arises from putting more weight on the liar's payoffs. On the other hand, the motivation to tell an altruistic white lie may rise from placing a higher weight on another's payoffs in the liar's utility function.

Figure 1 **Taxonomy of Lies Based on Change in Payoffs**



Notes. The origin in the figure represents the payoffs resulting from telling the truth. We define "white lies" as lies that increase the other side's payoffs, i.e., above the zero line in the horizontal dimension in the figure. "Black lies" are those that decrease the other side's payoffs, i.e., below the horizontal zero axes. Costly lies are defined as below zero on the vertical dimension.

Third, as Dreber and Johannesson (2008) demonstrated, there are gender differences in the tendency to lie. In particular, men are more likely to tell a selfish black lie. Contrasting men's and women's behavior in different domains of lying allows us to test gender differences in the interaction of distributional concerns and lying aversion. For the domain of selfish black lies, our data is consistent with the results of Dreber and Johannesson (2008): Men are more likely in our experiment to tell a selfish black lie. Moreover, men are significantly more likely to tell a Pareto white lie. Interestingly, however, women are significantly more likely than men to tell an altruistic white lie.

This result shows a valuable interaction between lie aversion and social preferences. It appears as if women have a higher cost of lying, but at the same time are more sensitive to another person's payoffs (for surveys of gender differences in social preferences, see Croson and Gneezy 2009, Eckel and Grossman 2008).

2. Experimental Design and Procedure

2.1. The Deception Game

Two players acted sequentially in the role of sender and receiver, respectively. We rolled a six-sided die before the start of the game and communicated the outcome only to the sender. The sender was then asked to send a message to the receiver from a pool of six possible messages. The six possible messages

[2] We abstract in the example from strategic issues such as reputation that are associated with the use of placebo, e.g., "In the clinical setting, the use of a placebo without the patient's knowledge may undermine trust, compromise the patient-physician relationship, and result in medical harm to the patient" (Council on Ethical and Judicial Affairs 2006).

are "the outcome of the roll of die was i," where $i \in \{1, 2, 3, 4, 5, 6\}$. The sender was told the payment in the experiment would depend on a choice made by the receiver. She was also told that the only information the receiver would have regarding the actual outcome of the roll of the die was her message. There were two payment options, A and B. The sender knew the payoffs (to both players) associated with each option, and she was told the receiver would not know the payoffs. Finally, the sender was told that if the receiver chose the real outcome of the die roll, payment option A would be implemented. Otherwise, both would be paid according to option B (the payoffs associated with each option are described below).

The message from the sender was the only information given to the receiver regarding the roll of the die. That is, the receiver did not know the actual number that came up but only the message received from the sender. After observing this message, the receiver was asked to choose a number from the set $\{1, 2, 3, 4, 5, 6\}$. This choice determined which of two possible payoff options, A or B, were implemented. Importantly, the receiver was not told what the actual payoffs associated with option A and option B were. She was told, just as the sender was, that if she chose the actual outcome of the roll of the die, payoff option A would be implemented, and for any other choice, payoff option B would be implemented.

2.2. Procedure

Table 1 summarizes the payoffs in the five different treatments used in this study (all in a between-subject design). As Table 1 illustrates, the treatments differ in the relative gains/losses to the sender or receiver when the receiver chooses a number other than the actual outcome of the roll of the die. Note that in all treatments, if the receiver chooses the actual outcome of the roll of the die, each player gets \$20.

The experiment was conducted in a classroom setting. The 517 subjects who participated were undergraduate students taking introductory courses in management. Full instructions are reported in Appendix A. As we describe in the experimental instructions, one of 20 senders was paired with a receiver, and this pair was paid according to the instructions.

Table 1 Payoffs (in \$) in the Different Treatments

	N			
Treatment	Male	Female	A	B
$T[-1, 10]$	62	39	(20, 20)	(19, 30)
$T[1, 10]$	63	38	(20, 20)	(21, 30)
$T[10, 10]$	55	47	(20, 20)	(30, 30)
$T[1, -5]$	59	45	(20, 20)	(21, 15)
$T[10, 0]$	64	45	(20, 20)	(30, 20)

Note. *N* is the number of senders per treatment.

3. Results

Before discussing the actions of the senders, on which this paper focuses, we discuss how the richer message space we employ in our design can address the issue of sophisticated deception (Sutter 2009). Consider a sender who prefers option B over option A (i.e., $u_B > u_A$). Suppose that the sender believes that with probability p the receiver will not believe her message and try to invert her message by choosing a number randomly from the remaining five messages. Then, her (expected) value from sending a truthful message is $V_T = (1-p)u_A + pu_B$, and her value from sending a deceptive message (assuming no lying costs) is $V_D = (1-p)u_B + p(4/5u_B + 1/5u_A)$. Thus, a sender will engage in sophisticated deception (i.e., $V_T > V_D$) only if $p > 83\%$. Although it is still feasible in our game to have some senders who believe that more than 83% of the receivers will not follow the message, the expanded message space makes it less likely that such sophisticated deception will play a major role. Hence, the rest of the discussion assumes that senders expect the receivers to follow the message they send.

The fraction of people who chose to lie, per treatment, is presented in Table 2 and graphically in Figure 2. The origin in the figure represents the payoffs of \$20 to each player and is obtained when the receiver chooses the actual outcome. Hence, the different treatments represent the deviations in payoffs resulting from the receiver choosing a number different than the actual outcome of the die roll. The fractions represent the actual fraction of senders who chose to lie.

We now discuss the main empirical results. The first result regards altruistic white lies, in which the sender loses money and receiver gains money if option B is implemented.

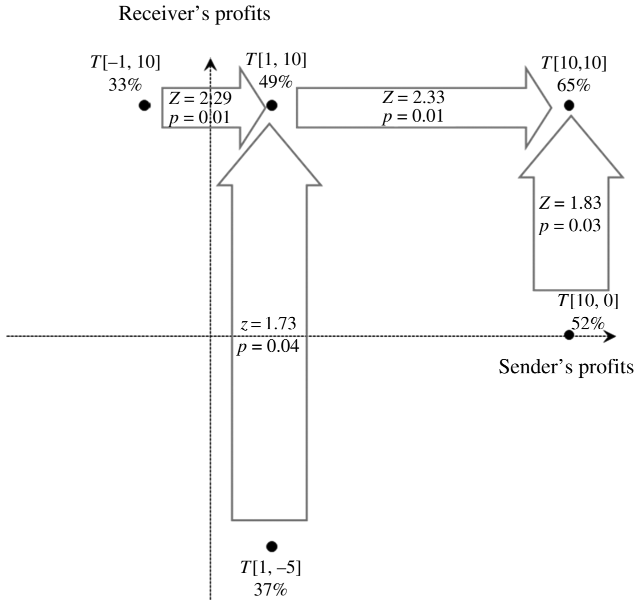
Result 1. A significant fraction of senders lie when it costs them a little but helps the receiver a lot.

A substantial fraction of the senders in the altruistic white lies treatment ($T[-1, 10]$) chose to lie. This finding lends strong support for the role of social preferences in determining deception. Findings indicate that 33% of the senders were willing to lose \$1 in order for the receiver to gain \$10. Although the payoff resulting from lying is not a Pareto improvement over telling the truth, lying does increase the total surplus.

Table 2 Proportion of Senders Who Lied in Each of the Five Treatments

Treatment	Fraction of lies
$T[-1, 10]$	33/101 (33%)
$T[1, 10]$	49/101 (49%)
$T[10, 10]$	66/102 (65%)
$T[1, -5]$	38/104 (37%)
$T[10, 0]$	57/109 (52%)

Figure 2 Sensitivity of Proportion of Lies to Incentives



Notes. The arrows represent the sensitivity of proportion of lies to incentives. The numbers given within the arrows represent the significance level.

The second result is on Pareto white lies and lie aversion.

Result 2. A significant fraction of senders do not lie even when lying results in a Pareto improvement.

Even in $T[10, 10]$, when each participant earns an extra \$10 from a lie, only 65% of the senders lie. This finding offers strong support to the hypothesis that lying incurs a cost for some people. This result represents clean evidence of lie aversion, without any confounding of distributional preferences. Thus, the aversion to lying cannot be (fully) explained by the harm it causes others, as the aversion is present even when everyone gains from the lie.

It would be interesting in future research to see how other factors influence this cost of lying. For example, in the domain of selfish black lies, Lundquist et al. (2009) study lying aversion along two dimensions not studied by Gneezy (2005). First, they show that the aversion to lying grows stronger the further one deviates from the truth. Second, using a richer set of messages, they allow for promises (as in Ellingsen and Johannesson 2004, Charness and Dufwenberg 2006) and find that aversion to lying depends on the strength of the promise.

Karthik (2009), using the classic strategic communication setting of Crawford and Sobel (1982) with the additional assumption that some messages entail exogenous or direct lying costs for the sender, shows that lying costs or aversion are important in modeling deception (see also Matsushima 2008 and Ottaviani, Squintani 2006).

Table 3 Logistic Regression of All of the Five Cases

Variable	*df*	Parameter estimate	Error	Wald χ^2	Pr > χ^2
Intercept	1	−0.62	0.15	18.15	0.01
Own benefit	1	0.09	0.02	20.73	0.01
Others benefit	1	0.03	0.01	3.27	0.07

Note. df, degree of freedom.

The third result we discuss is the role of consequences on deception.

Result 3. Senders are sensitive to own and receiver's cost/benefit associated with a lie.

Not surprisingly, more people were willing to lie in $T[10, 10]$ (65%) than in $T[1, 10]$ (49%) ($Z = 2.33$, $p = 0.01$).[3] That is, the cost of lying was high enough for some senders that they avoided sending a deceptive message when their benefit was only \$1, but not when their benefit was \$10. This finding provides support for the hypothesis that the decision to lie depends on the incentives involved. More senders are likely to tell a Pareto lie when the incentives are higher.

Similarly, more people were willing to lie in $T[1, 10]$ (49%) compared to $T[-1, 10]$ (33%) ($Z = 2.29$, $p = 0.01$). That is, Pareto lies seem easier to tell compared to altruistic lies.

The difference between the fraction of lies in $T[1, 10]$ (49%) and $T[1, -5]$ (37%) is significant ($Z = 1.73$, $p = 0.04$). This result mirrors other findings in the literature that for their own given benefit, senders are less likely to lie when the cost to the receiver increases (Gneezy 2005). Similarly, the fraction of lies in $T[10, 10]$ (65%) is significantly greater than the fraction of lies in $T[10, 0]$ (52%) ($Z = 1.83$, $p = 0.03$). This finding extends past findings of the role of receiver costs to show that senders are less likely to lie even when the benefit to the receiver is smaller.

More generally, Table 3 shows results from a regression testing how the likelihood of lying is affected by the liar's own and others' gains/losses. Overall, the likelihood of lying is significantly greater when the liar's own or another's benefit is greater. Figure 2 illustrates this important trend. More senders lie as we move right, i.e., increase the sender's payoff. At the same time, moving up (i.e., decreasing the receiver's cost or increasing the receiver's benefit) results in an increase in the fraction of senders who lie. Our results generalize findings regarding the importance of incentives in the decision to lie to

[3] The p-values are calculated from a one-tailed test of the equality of proportions, using normal approximation to the binomial distribution.

the case of white lies, and show senders are sensitive to their payoffs (benefits or costs) as well as the receivers' payoffs (benefits or costs).

This aggregate result is likely to hide individual differences. For example, it seems plausible that an "altruistic" sender may be willing to tell an altruistic white lie, but not a selfish black lie. A "selfish" sender may have just the opposite preferences.

An important aspect of this individual contrast is the case of gender differences. Do women and men have different propensity to lie, and if so, how does this depend on incentives? The next section addresses this question.

3.1. Gender Differences in the Propensity to Lie

The literature documents systematic differences between women and men in decision making (for surveys, see Eckel and Grossman 2008, Croson and Gneezy 2009). Is there such a difference in the propensity to lie?

Dreber and Johannesson (2008) use Gneezy's (2005) design and find, in the domain of selfish black lies, that men are significantly more likely to lie to secure a monetary benefit. Our data shows the same pattern: In the $T[1, -5]$ treatment, the fraction of men who lie (39%) is more than the fraction of women who lie (33%) ($Z = 0.59$, $p = 0.28$). Can we conclude from this that men are always more likely to lie? The following result suggests that we cannot.

Table 4 **Logistic Regression for Pareto White Lies**

Variable	*df*	Parameter estimate	Error	Wald χ^2	$\Pr > \chi^2$
Intercept	1	−0.48	0.28	3.08	0.08
Treatment ($T[10, 10] = 1$)	1	0.74	0.29	6.36	0.01
Gender (*Male* = 1)	1	0.68	0.30	5.20	0.02

Notes. Treatments $T[1, 10]$ and $T[10, 10]$. *df*, degree of freedom.

Result 4. Women are more likely to tell an altruistic white lie than men when the lie hurts oneself a little but helps the other a lot.

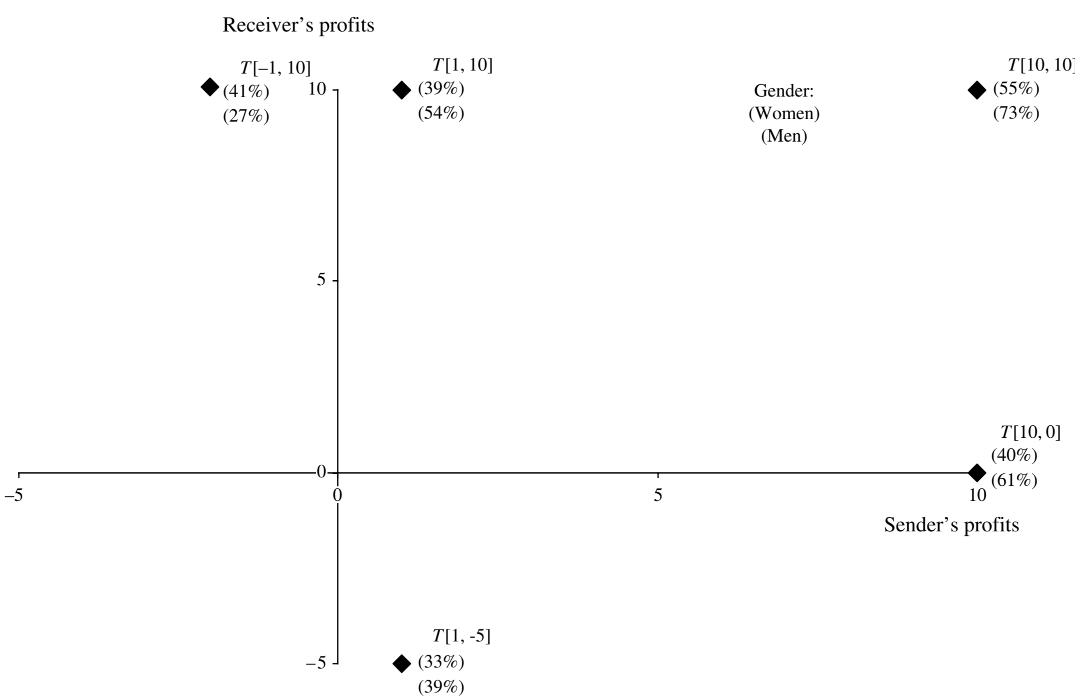
The fraction of men and women senders who chose to lie, per treatment, is presented in Figure 3. In $T[-1, 10]$, more women lied than men ($Z = 1.42$, $p = 0.08$). Unlike in the case of selfish black lies, women senders were significantly more willing than men to lie and endure a small cost to themselves, as long as it helps the receivers a lot.

The final result we document is about gender differences in the domain of Pareto white lies.

Result 5. Women are less likely to tell a Pareto white lie than men.

A logistic regression of propensity to lie in the Pareto lie treatments $T[1, 10]$ and $T[10, 10]$, shown in Table 4, reveals a significant main effect of gender (Wald $\chi^2 = 5.20$, $p = 0.02$); for more women than men, the cost of lying seems higher than the benefit of the Pareto white lies ($Z = 1.41$, $p < 0.08$ for $T[1, 10]$, and

Figure 3 **Proportion of Women and Men Senders Who Lied in Each of the Five Treatments**



Note. The number on the top in each treatment is the fraction of women who lied, and the number on the bottom is the fraction of men who lied.

$Z = 1.83$, $p = 0.03$ for $T[10, 10]$). For example, only 55% of women, compared to 73% of men, were willing to lie when the lie increased both participants' payoffs by $10.

Furthermore, comparing the fraction of men and women who lie in treatments $T[-1, 10]$ and $T[1, 10]$ shows an effect of how gender influences sensitivity to incentives: When a liar's own payoffs change from a $1 cost to a $1 benefit, the fraction of men who lie increases significantly (from 27% to 54%, $Z = 3.02$, $p < 0.01$); however, the fraction of women who lie does not change significantly (from 41% to 39%).

4. A Within-Subject Design

The evidence above is based on a between-subjects design, in which each participant is faced with one possible set of payments. As a result, the statistical analysis is based on the assumption of random assignment, and hence reveals the aggregate differences between treatments. An alternative approach that we report in this section is a within-subject design, in which each participant is faced with more than one set of payoffs and makes multiple decisions. Such a design, unlike the between-subject one reported earlier, allows the analysis to be done on an individual basis, and thus to compare the behavior of a given participant in different payment schemes. The disadvantage is the risk of an "experimenter demand," by which the participant feels that she is expected to change her behavior in response to the changes in payoffs. We believe that a combination of these two approaches can give us important converging evidence with regard to the response of participants to changes in payoffs.

4.1. Design and Procedure

The same sender-receiver game as before was employed with a within-subject design, in which the senders are asked to make choices for four different payment possibilities. For each of the four different payment possibilities, the senders are asked to send a message and are also informed that one of these four payment possibilities would be chosen at random at the end of the experiment for payment.

For each payment possibility i ($i = 1, 2, 3, 4$), the senders are asked to send a message. If the receiver, after seeing this message, chooses the actual outcome of the die roll, payment option Ai is implemented, and if the receiver's choice differs from the actual outcome of the die roll, then payment option Bi is implemented. The payments Ai and Bi (for $i = 1, 2, 3, 4$) are given in Table 5.

Table 5 Payoffs (in $) in the Within-Subject Design

Payment possibility	A	B
1. ($T[-1, 10]$)	(20, 20)	(19, 30)
2. ($T[1, 10]$)	(20, 20)	(21, 30)
3. ($T[10, 10]$)	(20, 20)	(30, 30)
4. ($T[1, -5]$)	(20, 20)	(21, 15)

Table 6 Fraction of Lies in the Different Treatments

Payment possibility	Fraction of lies
1. ($T[-1, 10]$)	25/58 (43%)
2. ($T[1, 10]$)	38/58 (66%)
3. ($T[10, 10]$)	44/58 (76%)
4. ($T[1, -5]$)	30/58 (52%)

Table 7 Contingency Tables for the Within-Subject Design

	$T[-1, 10]$	
$T[1, 10]$	Lie	Truth
Lie	19	19
Truth	6	14
	$T[1, 10]$	
$T[10, 10]$	Lie	Truth
Lie	34	10
Truth	4	10
	$T[1, -5]$	
$T[1, 10]$	Lie	Truth
Lie	27	11
Truth	3	17

All the payment possibilities were shown on a single instructions page, and the actual ordering of these possibilities was randomized. The experiment was conducted in a classroom setting. The subjects were undergraduate students taking an introductory course in accounting. Full instructions are reported in Appendix B.

4.2. Results

A total of 58 participants, 28 male and 30 female, completed the experiment. The percentage of lies for each of the payment options is presented in Table 6. As in our earlier between-subject design, the following may be observed:

(a) *Result* 1: a significant fraction of participants are willing to tell an altruistic lie (43% in $T[-1, 10]$).

(b) *Result* 2: a significant fraction of subjects do not tell even a Pareto white lie (24% in $T[10, 10]$).

Table 7 shows the 2×2 contingency tables used for testing for increase in propensity to lie with incentives.[4] As with Result 3 identified earlier, the

[4] The p-values are given for one-sided McNemar's test.

Table 8 Logistic Regression for the Within-Subject Design

Variable	*df*	Estimate	Error	Wald χ^2	Pr > χ^2
Intercept	1	0.06	0.21	0.09	0.76
Treatment ($T[1,-5]$)	1	−1.06	0.53	4.02	0.04
Treatment ($T[-1,10]$)	1	−1.66	0.54	9.56	0.01
Treatment ($T[10,10]$)	1	1.12	0.64	3.09	0.08
Participant fixed effects	—				

Note. df, degree of freedom.

Table 9 Percentage of Men and Women Who Lie

Payment possibility	Fraction of lies among men	Fraction of lies among women
1. ($T[-1,10]$)	9/28 (32%)	16/30 (53%)
2. ($T[1,10]$)	20/28 (71%)	18/30 (60%)
3. ($T[10,10]$)	22/28 (79%)	22/30 (73%)
4. ($T[1,-5]$)	18/28 (64%)	12/30 (40%)

propensity to lie is (i) significantly higher in $T[1,10]$ compared to $T[-1,10]$ ($\chi^2 = 5.75$, $p < 0.01$); (ii) significantly higher in $T[10,10]$ compared to $T[1,10]$ ($\chi^2 = 1.79$, $p = 0.09$); and (iii) significantly higher in $T[1,10]$ compared to $T[1,-5]$ ($\chi^2 = 3.5$, $p = 0.03$). Thus, the propensity to lie is indeed affected both by own and the other persons payoffs.

Identical results, shown in Table 8, are obtained with a logistic regression of propensity to lie in the four payment options. Specifically, there is lower likelihood of lying in $T[-1,10]$ compared to $T[1,10]$ ($Z = -2.00$, $p = 0.04$); lower likelihood of lying in $T[1,-5]$ compared to $T[1,10]$ ($Z = -3.09$, $p = 0.01$); higher likelihood of lying in $T[1,10]$ compared to $T[10,10]$ ($Z = 1.76$, $p = 0.08$).

The within-subject design also allows a somewhat stronger test. Specifically, if a person switched from truth to lying when going from $T[-1,10]$ to $T[1,10]$, then this person would continue to lie in $T[10,10]$. Indeed, our data indicates that every one of the 19 subjects who switched from telling the truth to lying when going from $T[-1,10]$ to $T[1,10]$ also continue to lie in $T[10,10]$.

Finally, the results shown in Table 9 indicate that the systematic differences between men and women to lie and its dependence on the type of lies are also present in this new experiment. Results from logistic regressions (shown in Table C.1 in Appendix C) indicate that women are more likely to lie than men in the altruistic lies region (Result 4) whereas men are more likely to lie than women in the Pareto white lies region (Result 5).

5. Conclusion

Why should economists care about small lies with seemingly little economic consequences? In contrast with psychologists, an important challenge to behavioral economists is to show that our findings have some impact on important economic behavior.

The importance of welfare-improving dishonest behavior is illustrated in monitoring and control. The performance feedback that a manager gives has the potential to change an employee's confidence and, consequently, the employee's performance in future tasks (see, for instance, Compte and Postlewaite 2004 for a model of a single decision maker whose performance depends on her confidence). In such a setting of strategic performance feedbacks, the alignment of incentives (as in the Pareto white lies domain) combined with lying aversion, results in managers giving employees false praise to build up confidence. Furthermore, such false praise is potentially welfare improving.[5]

The examples we discussed in the introduction further illustrate the economic importance of understanding dishonest behavior, both when dishonesty helps and when it hurts others. So, why do some people avoid lying? One explanation (put forward by Dufwenberg and Gneezy 2000, Charness and Dufwenberg 2006, and Dufwenberg and Battigalli 2009) is that people experience belief-based guilt. Specifically, people experience a disutility when they let down others. Furthermore, these models suggest the size of this disutility is related to the difference between the consequence of the action and the other person's expectation of the consequence of her own action. The aversion to lying in our experiment in the domain of Pareto white lies cannot be explained by this theory of guilt based on monetary consequences, because the consequence of lying in each case is a Pareto improvement. In particular, the receiver of the lie also earns more money as a result of the lie. This suggests that at least part of the reason people do not lie may be connected with their endogenous lying cost, not with guilt over the consequences. Indeed, our results do not preclude the possibility that people feel guilty over the act of lying itself. However, our results do demonstrate aversion to lying cannot be solely explained by the negative consequences of the lie.

A different take is proposed by Ellingsen and Johannesson (2004), who model a commitment-based guilt where the act of breaking a promise incurs an endogenous cost. In a recent study, Vanberg (2008) contrasted these two theories of guilt (expectation-based and commitment-based) and offered evidence

[5] Note that lying aversion appears to be crucial for this result. For instance, if no one exhibited lying aversion, then irrespective of the employees past performance, the manager should give the feedback that yields the greatest confidence. Consequently, a rational employee should disregard any feedback as it is uninformative, and thus the manager ex post has no reason to give false praise.

suggesting people's second-order beliefs (i.e., beliefs about others' expectations) do not significantly affect whether or not they break a promise. From this, he concludes that people have a preference for promise keeping, per se. Still, because the study considers settings where breaking a promise results in benefit to oneself and cost to another, it leaves open the question of whether the consequences affect this preference. Specifically, do people have a preference for keeping their promise because it is a "promise," or does the preference for truthfulness arise because of the negative consequences of breaking the promise? In line with Ellingsen and Johannesson's (2004) model, our study suggests such preferences, although influenced by the consequences, are present even when the consequences are positive (Pareto improving).

We do not claim expectations are irrelevant to whether a person lies or not. Someone who does not "bluff" in poker is a bad player. The expectation in such a context would be that everyone lies. A related alternative is that people feel guilty when they lie because they violate a social norm. The amount of guilt they feel depends on the descriptive norm, that is, their beliefs about adherence to the norm in their peer group. Specifically, people would feel greater guilt about lying when they expect their peer group to be more honest (see Lundquist et al. 2009 for experimental evidence).[6]

For demonstration, consider the simple model with outcomes A and B. Suppose that a person's utility from outcomes A and B are U_A and 0, respectively. Outcome B is obtained when the person is honest and outcome A when the person lies (and violates the norm). A person experiences disutility c when she believes no one in her peer group lies and she lies. When she believes that a fraction, p, of her peers lie, she experiences $(1-p)c$ disutility from lying. Hence, her total utility from lying is $U_A - (1-p)c$ and from being honest is zero.

This explanation is consistent with Vanberg's (2008) experimental finding that second-order beliefs may not impact lying aversion. It is also consistent with Charness and Dufwenberg's (2006) results if we admit the possibility of false consensus, i.e., correlation between p (a person's belief about prevalence of lying) and her second-order expectation about others' beliefs about lying. This correlation could explain why, in their experiment, people who had higher second-order beliefs in the experiment were also more likely to lie.

Our simple explanation predicts that a person's belief about the descriptive norm in a given context affects lying aversion and offers a rationale for why people may not feel much guilt about lying in poker, as people expect no one to tell the truth in such contexts. Note that in contrast to a second-order beliefs explanation (people expect that in poker others do not expect them to tell the truth), our explanation relies on much simpler first-order beliefs about descriptive norms (people expect that in poker no one tells the truth). An important avenue for future research includes investigation into this conjectured relationship between descriptive norms and actual deception behavior, and equally importantly the source of beliefs about the descriptive norm of a person's own peer group.

Our results also help in testing a recent explanation of Gneezy's (2005) result offered by Hurkens and Kartik (2009). They offer an explanation that is a special case of the Gneezy (2005) one. In their explanation, 50% of the people are ethical and never lie, and the remaining 50% will choose to lie whenever the lie will yield their preferable outcome. More generally, they claim that Gneezy's (2005) data or their own data cannot reject the hypothesis that people have either zero or infinite cost of lying.[7]

To test their hypothesis, one needs to find a design in which social preferences work in the same way, and hence only the cost of lying is relevant. As noted, Gneezy's (2005) or Hurkens and Kartik's (2009) design cannot test this; our design can. In particular, inequity models, or any other model of social preferences we are aware of, would predict that a person with no lying costs in treatment $T[10, 10]$ strictly prefers (30, 30) over (20, 20), and one in $T[10, 0]$ strictly prefers (30, 20) over (20, 20).[8] Hence, Hurkens and Kartik (2009) predict that the same fraction of people will choose to lie in both cases. Because the difference between the fraction of participants who lie in $T[10, 10]$ (65%) and $T[10, 0]$ (52%) is large and significant ($Z = 1.83$, $p = 0.03$), we can reject their hypothesis. Thus, our results reject the unrealistic assumption that a person's propensity to lie is independent of incentives.

Finally, the different reaction of men and women to lies provides a deeper understanding of the influences on the decision to lie. Female senders are more

[6] An alternate explanation can focus on a person's perception (belief) about the strength of the injunctive "do not lie" social norm, i.e., a person's beliefs about what fraction of people find lying acceptable. Still, these two beliefs (the injunctive norms and the descriptive norms) are likely to be highly correlated (unless the person believes that his or her peer group members are hypocrites). So for the purposes of our explanation, it seems adequate to focus only on one.

[7] Given the implausible and atheoretical nature of the hypothesis, not rejecting the hypothesis arguably speaks more of the power of the previous empirical tests rather than the truth of the hypothesis.

[8] For instance, in Fehr and Schmidt (1999) where $U(x, y) = x - \alpha \max\{x-y, 0\} - \beta \max\{y-x, 0\}$, because $\alpha < 1$, the outcome (30, 30) or (30, 20) are always preferable to (20, 20).

likely than male senders to lie when the lie hurts the sender a little but helps the receiver a lot (upper left quadrant of Figure 1). In addition, female senders are less likely than male senders to lie when the lie helps the sender a little but hurts the receiver equally or even a little more (lower right quadrant). And female senders are also less likely to lie when doing so will help both the sender and receiver (upper right quadrant). This could be viewed as stronger altruistic preferences among women compared to men. Because in one case women are more likely and in the other case less likely than men to lie, the results are not consistent with an explanation that includes only a simple cost of lying. If lying aversion was the only reason to avoid lying, we would expect the tendency to lie in these two situations to be similar. Although people avoiding lying in the Pareto white lies case implies aversion to lying is important, gender differences suggest social preferences are important, too. The value of these results lies in the interaction of these two forces.

Appendix A

Instructions for Player 1 (Sender)
Welcome to our short experiment. Please read these instructions carefully. You may earn a considerable sum of money, depending on the decisions you make in the experiment. We will randomly choose 1 participant out of 20 and pay this person next week. The rest of the instructions describe what the procedure will be should you be chosen.

You will be matched randomly with another participant in this experiment. Neither of you will know the identity of the other.

Before starting this experiment, we have rolled a 6-sided die, and obtained the outcome 2. The other participant will not be informed that the outcome of the die roll was 2. However, he or she will be told that you have been informed the outcome of the die roll.

We now ask you to send a message to the other participant. The message will correspond to a number from 1 to 6. There are six possible messages:

Message 1: "The outcome from the roll of the 6-sided die is 1."

Message 2: "The outcome from the roll of the 6-sided die is 2."

Message 3: "The outcome from the roll of the 6-sided die is 3."

Message 4: "The outcome from the roll of the 6-sided die is 4."

Message 5: "The outcome from the roll of the 6-sided die is 5."

Message 6: "The outcome from the roll of the 6-sided die is 6."

Your message will be shown to the other participant, and then he or she will be asked to choose a number between 1 and 6.

The choice of the number by the other participant will determine the payments in the experiment.

- If he or she chooses the actual outcome of the roll of the dice, 2, then you will receive $20 and he or she will receive $20.
- If he or she chooses a number different than the actual outcome, you will receive a $X and he or she will receive $Y.

Finally, only you will be informed of the particular monetary value connected to each message. The other participant will not be informed of these monetary values. However, he or she will be told that you have been informed of the monetary value connected to each message.

The message I will send is "The outcome from the roll of the 6-sided die is ______."

Treatment	X	Y
$T[-1, 10]$	19	30
$T[1, 10]$	21	30
$T[10, 10]$	30	30
$T[1, -5]$	21	15
$T[10, 0]$	30	20

Instructions for Player 2 (Receiver)
Welcome to our short experiment. Please read these instructions carefully. You may earn a considerable sum of money, depending on the decisions you make in the experiment.

You will be matched randomly with another participant in this experiment. Neither of you will know the identity of the other.

Before starting this experiment, we have rolled a 6-sided die, and told the outcome of it to the other participant, but we are not going to tell it to you.

After being informed of the roll of the die, the other participant has sent a message to you. The message corresponds to a number from 1 to 6. There are six possible messages:

Message 1: "The outcome from the roll of the die is 1."
Message 2: "The outcome from the roll of the die is 2."
Message 3: "The outcome from the roll of the die is 3."
Message 4: "The outcome from the roll of the die is 4."
Message 5: "The outcome from the roll of the die is 5."
Message 6: "The outcome from the roll of the die is 6."

The message the other participant sent is as follows:

Message: "The outcome from the roll of the 6-sided die is ___."

Now we ask you to choose a number between 1 and 6. The message you received is the only information you will have regarding the roll of the die. Your choice of a number will determine the payments in the experiment according to two different options (option A and option B), known only to the other participant.

If you will choose the same number as the number that came up in the roll of the die, both of you will be paid according to option A. If you will choose a number different than the actual number, you will both be paid according to option B.

Do you have any questions?

The number I choose is _____.

Appendix B

Instructions for Player 1 (Sender)
Welcome to our short experiment. Please read these instructions carefully. You may earn a considerable sum of money, depending on the decisions you make in the experiment.

We will randomly choose 1 participant out of 20 and pay this person next week. The rest of the instructions describe what the procedure will be should you be chosen.

You will be randomly matched with another participant in this experiment. Neither of you will know the identity of the other.

Before starting this experiment, we have rolled a 6-sided die, and obtained the outcome 2. The other participant will not be informed that the outcome of the die roll was 2. However, he or she will be told that you have been informed about the outcome of the die roll.

There are four different payment possibilities. Only you will be informed of the particular monetary values in each payment option. The other participant will not be informed of these monetary values. However, he or she will be told that you have been informed of these monetary values.

At the end of the experiment, we will choose one of these payment possibilities at random for actual payment.

For each of the payment possibilities, we now ask you to send a message to the other participant. The message will correspond to a number from 1 to 6. There are six possible messages:

Message 1: "The outcome from the roll of the 6-sided die is 1."

Message 2: "The outcome from the roll of the 6-sided die is 2."

Message 3: "The outcome from the roll of the 6-sided die is 3."

Message 4: "The outcome from the roll of the 6-sided die is 4."

Message 5: "The outcome from the roll of the 6-sided die is 5."

Message 6: "The outcome from the roll of the 6-sided die is 6."

Your message will be shown to the other participant, and then he or she will be asked to choose a number between 1 and 6.

The choice of the number by the other participant will determine the payments in the experiment.

Payment possibility #1

• If he or she chooses the actual outcome of the roll of the dice, 2, then you will receive $20 and he or she will receive $20.

• If he or she chooses a number different than the actual outcome, you will receive $19 and he or she will receive $30.

The message I will send in this payment option is

"The outcome from the roll of the 6-sided die is ______."

(*Please fill in a number from* 1 *to* 6.)

Payment possibility #2

• If he or she chooses the actual outcome of the roll of the dice, 2, then you will receive $20 and he or she will receive $20.

• If he or she chooses a number different than the actual outcome, you will receive $30 and he or she will receive $30

The message I will send in this payment option is

"The outcome from the roll of the 6-sided die is ______."

(*Please fill in a number from* 1 *to* 6.)

Payment possibility #3

• If he or she chooses the actual outcome of the roll of the dice, two, then you will receive $20 and he or she will receive $20.

• If he or she chooses a number different than the actual outcome, you will receive $21 and he or she will receive $30.

The message I will send in this payment option is

"The outcome from the roll of the 6-sided die is ______."

(*Please fill in a number from* 1 *to* 6.)

Payment possibility #4

• If he or she chooses the actual outcome of the roll of the dice, 2, then you will receive $20 and he or she will receive $20.

• If he or she chooses a number different than the actual outcome, you will receive a $21 and he or she will receive $15.

The message I will send in this payment option is

"The outcome from the roll of the 6-sided die is ______."

(*Please fill in a number from* 1 *to* 6.)

Appendix C

Unlike in the regression, which examined treatment effects (Table 8 in the main text), in this regression, which examines gender differences, we cannot simply put a fixed effect for each participant. Hence, we estimated a logistic model with a participant random effect. That is, the model is

$$\begin{aligned}\text{LIE} = {} & \text{SUBJECT\_RANDOM\_EFFECT} \\ & + \text{TREATMENT\_DUMMY} \\ & + \text{GENDER} + \text{TREATMENT} : \text{GENDER}\end{aligned}$$

The results, shown in Table C.1, indicated that the gender dummy ($1\{Gender = Male\}$) is positive and significant; 1.37 (0.72), $p = 0.03$. Furthermore, the interaction term corresponding to the altruistic lies, $1(Treatment = [-1, 10]) \times 1(Gender = \text{M})$, is negative and significant; -2.68 (0.87), $z = -3.098$, $p < 0.01$.

Thus, the results with respect to gender differences are broadly consistent with our previous findings, and indicate that women are more likely to speak an altruistic lie (Result 4), but that men are more likely to speak a Pareto lie (Result 5).

Table C.1 Random Effects Model for Propensity to Lie

Variable	*df*	Estimate	Error	Wald χ^2	Pr > χ^2
Intercept	1	0.56	0.49	1.30	0.25
Treatment ($T[1, -5]$)	1	−1.13	0.59	3.75	0.03
Treatment ($T[-1, 10]$)	1	−0.38	0.58	0.43	0.25
Treatment ($T[10, 10]$)	1	0.83	0.62	1.77	0.09
Subject random effects					
Gender (*Male*)	1	1.38	0.72	3.65	0.03
($T[1, -5]$): *Male*	1	−0.67	0.89	0.56	0.22
($T[-1, 10]$): *Male*	1	−2.68	0.87	9.60	0.001
($T[10, 10]$): *Male*	1	−0.95	0.95	1.01	0.16

Note. df, degree of freedom.

References

Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*. Reprint, Clarendon Press, Oxford, UK, 1907.

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

Boles, T. L., R. T. A. Croson, J. K. Murnighan. 2000. Deception and retribution in repeated ultimatum bargaining. *Organ. Behav. Human Decision Processes* **83**(2) 235–259.

Brandts, J., G. Charness. 2003. Truth or consequences: An experiment. *Management Sci.* **49**(1) 116–130.

Cai, H., J. Wang. 2006. Overcommunication in strategic information transmission games. *Games Econom. Behav.* **56**(1) 7–36.

Charness, G., M. Dufwenberg. 2006. Promises and partnership. *Econometrica* **74**(6) 1579–1601.

Charness, G., M. Rabin. 2002. Understanding social preferences with simple tests. *Quart. J. Econom.* **117**(3) 817–869.

Compte, O., A. Postlewaite. 2004. Confidence-enhanced performance. *Amer. Econom. Rev.* **94**(5) 1536–1557.

Council on Ethical and Judicial Affairs. 2006. Placebo use in clinical practice. CEJA Report 2-I-06. Accessed October 22, 2011, http://www.ama-assn.org/ama1/pub/upload/mm/369/ceja_recs_2i06.pdf.

Crawford, V. 2003. Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *Amer. Econom. Rev.* **93**(1) 133–149.

Crawford, V., J. Sobel. 1982. Strategic information transmission. *Econometrica* **50**(6) 1431–1451.

Croson, R., U. Gneezy. 2009. Gender differences in preferences. *J. Econom. Literature* **47**(2) 448–474.

Croson, R., T. Boles, J. K. Murnighan. 2003. Cheap talk in bargaining experiments: Lying and threats in ultimatum games. *J. Econom. Behav. Organ.* **51**(2) 143–159.

DePaulo, B. M., D. A. Kashy, S. E. Kirkendol, M. M. Wyer, J. A. Epstein. 1996. Lying in everyday life. *J. Personality Soc. Psychol.* **70**(5) 979–995.

Dreber, A., M. Johannesson. 2008. Gender differences in deception. *Econom. Lett.* **99**(1) 197–199.

Dufwenberg, M., P. Battigalli. 2009. Dynamic psychological games. *J. Econom. Theory* **144**(1) 1–35.

Dufwenberg, M., U. Gneezy. 2000. Measuring beliefs in an experimental lost wallet game. *Games Econom. Behav.* **30**(2) 163–182.

Eckel, C. C., P. J. Grossman. 2008. Differences in the economic decisions of men and women: Experimental evidence. C. Plott, V. Smith, eds. *Handbook of Experimental Economics Results*. Elsevier, New York, 509–519.

Ellingsen, T., M. Johannesson. 2004. Promises, threats, and fairness. *Econom. J.* **114**(495) 397–420.

Ellingsen, T., M. Johannesson, J. Lilja, H. Zetterqvist. 2009. Trust and truth. *Econom. J.* **119**(534) 252–276.

Fehr, E., K. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quart. J. Econom.* **114**(3) 817–868.

Gino, F., S. Ayal. 2011. Honest rationales for dishonest behavior. M. Mikulincer, P. R. Shaver, eds. *The Social Psychology of Morality: Exploring the Causes of Good and Evil*. American Psychological Association, Washington, DC.

Gino, F., L. Pierce. 2009. Dishonesty in the name of equity. *Psych. Sci.* **20**(9) 1153–1160.

Gintis, H. 2012. Behavioral ethics. E. Slingerland, M. Collard, eds. *Integrating Science and the Humanities: Interdisciplinary Approaches*. Forthcoming.

Gneezy, U. 2005. Deception: The role of consequences. *Amer. Econom. Rev.* **95**(1) 384–394.

Hurkens, S., N. Kartik. 2009. Would I lie to you? On social preferences and lying aversion. *Experiment. Econom.* **12**(2) 180–192.

Kant, I. 1785. *Fundamental Principles of the Metaphysics of Morals*. [Translated by T. K. Abbott. Pearson, London, 1949.]

Kartik, N. 2009. Strategic communication with lying costs. *Rev. Econom. Stud.* **76**(4) 1359–1395.

Lundquist, T., T. Ellingsen, E. Gribbe, M. Johannesson. 2009. The aversion to lying. *J. Econom. Behav. Organ.* **70**(1–2) 81–92.

Matsushima, H. 2008. Role of honesty in full implementation. *J. Econom. Theory* **139**(1) 353–359.

Mazar, N., D. Ariely. 2006. Dishonesty in everyday life and its policy implications. *J. Public Policy Marketing* **25**(1) 117–126.

Mazar, N., O. Amir, D. Ariely. 2008. The dishonesty of honest people: A theory of self-concept maintenance. *J. Marketing Res.* **45**(6) 633–644.

Ottaviani, M., F. Squintani. 2006. Naive audience and communication bias. *Internat. J. Game Theory* **35**(1) 129–150.

Saint Augustine. 421. *Enchiridion on Faith, Hope, and Love*. [Translated and published by Regnery Publishing, Washington, DC, 1996.]

Sánchez-Pagés, S., M. Vorsatz. 2007. An experimental study of truth-telling in a sender–receiver game. *Games Econom. Behav.* **61**(1) 86–112.

Sherman, R., J. Hickner. 2008. Academic physicians use placebos in clinical practice and believe in the mind–body connection. *J. General Internal Medicine* **23**(1) 7–10.

Sutter, M. 2009. Deception through telling the truth? Experimental evidence from individuals and teams. *Econom. J.* **119**(534) 47–60.

Vanberg, C. 2008. Why do people keep their promise? An experimental test of two explanations. *Econometrica* **76**(6) 1467–1480.