SCHOOL OF ECONOMICS
AND MANAGEMENT
Lund University

Department of Economics

Magister Thesis

June 2007

Information Transmission and its Distortions: A Study of Cheap-Talk Games

Supervisor: Professor Håkan Jerker Holm

Author: Jakob Jeanrond

Abstract

Lies as distorters of information transmission are examined in this paper. A survey of research conducted on games featuring costless information transmission leads to a characterization of several types of lies. The relevant type of lie depends on the shape of players' strategy sets. Individuals' perception of different types of lies varies. Active misrepresentation is viewed as the most serious violation and thus individuals experience substantial lie aversion when confronted with such an alternative. For active misrepresentation to succeed additional characteristics such as differing player knowledge or non detectable lies are often required. When multiple players are present on either the sender or receiver side lying becomes more difficult in general, although there are situations where the opposite is true. The characterization of distortions in information transmission can help make policy decisions aimed at increasing truthful revelation more effective since different lies are prevented by dissimilar strategies.

Table of Contents

1. Introduction

Communication is becoming increasingly refined, not least due to the rapid development of the internet. Improvements in information transmission have included substantial reductions in the costs of sending information. In fact, the expense of sending a message today is often so low that economists disregard it when trying to model its expected content. If a message can be sent virtually for free, what are the incentives for individuals to be truthful? The role of lies as information distorters in environments where information transmission is costless is the subject of this paper. Knowledge regarding in which situations the truth can be expected or when and to what extent it might be distorted can assist decision-makers in their choice of communicational forum.

A suitable starting point for a discussion of lying in free message environments is the seminal work by Crawford and Sobel (1982). This paper shows the inherent dilemma present in costless message games and thus can deliver the necessary understanding for the following extensions and variations of the theory. The models following Crawford and Sobel (1982) employ varied concepts of lying. Such definitional variations allow a wider spectrum of communicative distortion to be encompassed in the discussion. The extensions also involve situations in which lying can be avoided. A discussion of the nature and relevance of policies aimed at increasing truthful disclosure is derived from the models. Naturally there is a limit to the amount of research that can be covered in detail, therefore additional results from related research are presented to complement and underscore the results obtained.

The paper is organized as follows. Section 2 presents relevant concepts and terminology. The model of Crawford and Sobel follows in section 3, which also presents some relevant related work. The extensions are covered in sections 4 and 5. Section 6 contains the analysis and section 7 concludes.

2. Terminology and Background

What is a lie? When faced with two alternatives, such as x and y, over which an individual has certain preferences, say $x \succ y$, a misrepresenting statement of these preferences $(y \succ x)$ composes a lie. Similarly, if an individual has observed the realization of one state (a) out of several possible states $\{(a),(b),(c),(d)\}$ a statement that asserts the realization of either $\{(b),(c),(d)\}$ would also commonly be considered a lie. Now consider a modification of the previous example: An individual observes outcome (a) out of the possible states $\{(a),(b),(c),(d)\}$. In his report to an uninformed decision maker the individual reports that either (a) or (b) occurred as the outcome. In one sense, this is the truth, since the report includes no misrepresentation of what happened. On the other hand however, the individual has withheld some information that could possibly have altered the decision making process of the counterpart. In this paper, such withholding will also be regarded as a type of lie. Lying constitutes *a message consciously sent to misrepresent the truth in order to achieve personal gain*.[1] Here, misrepresention is interpreted in a wider sense to encompass all representations that convey something other than *the whole unaltered truth*. Although is not difficult to come up with examples of lies that are told without any sort of direct personal gain in mind (see below chapter 5.4), the gain may be indirect in some sense. For economic decision-making the relevant classes of lies are incorporated by the definition above.

The theory of information transmission is founded on the assumption of imperfectly distributed information. Decision makers are assumed to be unaware of a certain state of the world that is important to a decision they intend to make. An informed expert possesses the relevant information and transmits a message (a signal) to the decision-maker prior to the decision. There exists a class of games where the signal sent has direct costs for the sender. Let the decision-maker be a company that is seeking to hire. The expert is the prospective employee who wants to convince the company to hire him. The company's decision can be influenced by sending documentation of education and acquired skills. Costs of sending the signal are comprised of those associated with the

---

[1] Cf. the definition of a lie in *dtv-Lexikon*, vol. 11, 149. – For a discussion of lying in Christian moral tradition see Preston (1986). Interestingly, the Swedish National Encyclopedia does not have an entry for either "lie" or "lying".

acquisition of the relevant qualification. These costs can include college tuition, internships etc. A key feature of expensive signals is that they are expected to be informative.[2] When sending a costly signal it is natural to include some sort of information that is expected to create gains that outweigh the costs.

A second class of information transmission games includes signals which are costless. An example could involve the formation of a study group where an individual is questioned by another whether his ability in the subject at hand is high or low. The signal, here represented by the two possible answers, is costless in the sense that the individual can simply assert his qualifications. There is a different cost associated with the signal though. This cost is associated with the reaction taken by the counterpart in response to receiving the signal. In the example above this cost could be represented by missing out on participation in the study group. Games which include costless signaling but where the signal is determined by the payoff associated with the action taken in response are called cheap-talk games.

Cheap-talk games are the focus of this paper since they provide an environment in which lying can be expected to take place frequently. Research has been conducted regarding when, or if, any information is actually transferred in cheap talk games. A discussion of cheap talk is presented in Farell and Rabin (1996), in which the authors conclude that communication works when correlation between the sender's true type and his preferences over receiver's beliefs is perfect. On the other hand, if correlation between type and beliefs fails completely no information is transmitted. Farell (1993) coined the term *babbling equilibria* for all those equilibria where information is transferred but is not informative due to large enough differences in preferences between sender and receiver. In general, cheap talk cannot guarantee an efficient outcome in games but it can in some cases enable coordination and help avoid misunderstandings (Farell and Rabin (1996)). As will be shown below, there are cases when it can. How changes in preferences can affect the messages sent, and the outcomes observed, in cheap-talk games is a concept explored by Crawford and Sobel (1982). Their model is presented in the next section.

[2] See Spence (1973)

3. A Model of Strategic Information Transmission

3.1. Introducing the Idea

The following model, originally presented by Crawford and Sobel (1982) will act as a benchmark for the extensions and discussion that follow. Since all the models that will later be presented draw upon this framework in one way or another, it is imperative to first understand certain concepts developed in Crawford and Sobel's paper. Since the paper is rather technical, only the main thoughts and results will be presented in an overview.

Attention is turned towards the informational activity in a game that features two players: a so called sender (*S*), and a receiver (*R*). *S* has private information and faces a decision on how much of it he wishes to share with *R*, or to be more precise in this case, how explicitly he wants to convey the details of his private knowledge to *R*. In the model this private information is modeled as a randomly distributed, single dimensional variable *m* that is only visible to *S*. Upon receiving a message from *S*, *R* takes an action that will determine the utility of both players. *S* thus faces the complication of trying to decide on how specific to be in his message to *R*. Crawford and Sobel show that under their assumptions *S* decision will depend on how well aligned preferences are between *S* and *R*. In the model this is represented by a preference similarity parameter denoted by *b*. The object of investigation is how the amount of information transmitted depends on variations in preference similarity. A somewhat more detailed presentation of assumptions and outcomes follows below.

3.2. The Model

The game is played between *S* and *R*, where *S* observes the random variable *m* which has a differentiable probability distribution function, $F(m)$ with density $f(m)$ and is supported on $[0,1]$. *S* and *R* both have twice differentiable von Neumann-Morgenstern utility functions which are denoted by

$$U^S(y,m,b)$$

for *S* and

$$U^R(y,m)$$

for R. Here y, a real number, represents the action taken by R in response to the message sent by S. As noted above b is a scalar parameter that represents the similarity of interests. Every aspect of this game except m is common knowledge. It is assumed that U^i, $i = R, S$, has a maximum in y for each pair of (m, b).

The game proceeds through three stages. At stage one (time zero), S observes his private information (his type). Then in stage two S sends a signal, a possibly noisy estimate of m to R. Finally in stage three, R observes and processes the signal, whereupon an action (taken by R) determines the payoffs of both players. The solution concept employed in the game is Harsanyi's Bayesian Nash Equilibrium. This is simply a Nash Equilibrium in the decision rules that agents utilize to relate their actions to their respective information and the situations in which they find themselves. In other words, each agent maximizes expected utility, responding optimally to his opponent's strategy choice, taking into account its implications given his probabilistic beliefs.

Crawford and Sobel characterize the equilibrium as a family of signaling rules for S and an action rule for R such that S maximizes expected utility for each possible state of information he could observe, taken R's action rule as given. R in turn responds optimally to each possible type of signal. Note that the signal sent by S does not involve any explicit costs. However, there are endogenous costs involved since the signal will affect the action taken by R and thus the outcome.

All equilibria take the form of so called *partition equilibria*. These types of equilibria are characterized by the fact that S introduces noise into his message by not discriminating as finely as possible among the information he can distinguish. In addition, there are a finite number of equilibria, i.e. there is an upper bound on the size of possible partitions. There must exist one equilibrium of each partition size from one through $N(b)$ which denotes the largest partition possible (and depends on the preference similarity parameter!). Under the assumptions present in the model Crawford and Sobel prove that ex ante, agents can reasonably be expected to coordinate on the equilibrium of size $N(b)$.

3.3. A Very Simple Illustration

This very simple illustration will hopefully provide some further intuition to the abstract setting. Here, S observes one of four possible states denoted (A, B, C, D). In his signal to

R, S might obscure his observation depending on how aligned his preferences (payoffs) are with R's. Given a set of different b:s the following partitions could be possible:

{(A), (B), (C), (D)}, {(A, B), (C), (D)}, {(A, B, C), (D)}, {(A, B), (C, D)}, {(A, B, C, D)}

Note that this list is not all exhaustive in the sense that e.g. {(A, B, D), (C)} is not represented.[3] The states should be viewed as non definite, that is they have not been independently identified. Therefore, ex ante {(A, B, D), (C)} = {(A, B, C), (D)}. The message chosen by the sender is drawn from his preferred partition.
The most informative partition

{(A), (B), (C), (D)}

can be interpreted as that regardless of which state S observes, he will reveal the state in his signal to R. In the most uninformative partition

{(A, B, C, D)}

no information regarding the true state observed is revealed in the signal sent (only the possible outcomes). A possible middle ground is provided by the partition

{(A, B), (C), (D)}

in which information will subsequently be perfectly transmitted in the cases of C and D but not in A or B. This could be due to the fact that payoffs vary for different outcomes associated with R's action rule and thus it is not in S's best interest to convey the whole truth in case of A or B. As the preference similarity parameter approaches zero, the probability that

{(A), (B), (C), (D)}

is possible in equilibrium increases. Thus the number of partitions $N(b)$ decrease and the partition sizes increase as b increases.

3.4. Remarks

Crawford and Sobel (1982) deliver a setting that is convenient in the sense that more similar preferences imply more precise information transmitted. On the other hand, as acknowledged by the authors in their conclusion, outcomes related to lying are not fully

[3] An all exhaustive listing would simply have used too much space.

satisfactory. This game features no form of active misrepresentation, only intentional vagueness. Even though the definition of lying as presented above included intentional vagueness, this is but one of several possible ways to lie.

3.5. Associated Research

The beginning of the eighties saw several important contributions to the area of strategic information transmission. In addition to Crawford and Sobel (1982), an influential paper was presented by Green and Stokey (1981).[4] In the latter the authors concentrate on changes in the information structure in situations where agent's preferences are held constant. Green and Stokey present an argument in which improved information structures will not generally be beneficial to both agents' welfares. Improved information can however be welfare improving under the condition that it is "success enhancing". This implies that the probability of receiving an uninformative observation is decreased and the probability of receiving other informative messages is increased.

The concepts of Crawford and Sobel are empirically tested by Dickhaut et al. (1995). To a large extent, the results confirm the theory presented in Crawford and Sobel (1982), i.e. messages become less informative when preferences diverge and optimal actions are less frequently matched to the actual state. Dickhaut et al. (1995) mention that it might be of interest to empirically investigate how multiple receivers can affect a sender's information disclosure. For instance, a firm that wants to attain benefits by revealing information to the capital market may have to balance these against the dangers associated with revealing too much to competitors. A model taking into account this precise example is presented in Farell and Gibbons (1989). In their one sender versus two receivers setting, they distinguish between five cases that can occur. Apart from the two more obvious possibilities, no revelation and full revelation, there are three more interesting situations. In one, the communication in public (to both receivers) is disciplined by one receiver, so called *one sided discipline*, i.e. there exists a private equilibrium with one receiver and not with the other, but there is a public equilibrium. This can be thought of as a case where a firm which will not communicate anything to a

[4] An updated version was presented by Green and Stokey in 2003, this is the version referred to in this paper.

potential competitor in private decides to make information public so as to receive benefits from the capital market. Here the expected benefits from the capital market discipline the firms action and create a public equilibrium where information is also transferred to potential competitors.[5] A second case is *mutual discipline*. There exists no private equilibrium but there exists an equilibrium in public. This type of state is often associated with lying in both directions. A firm could have incentives to lie to two interest groups in private but is forced to reveal the truth on average in public.[6] The last case is called *subversion* where there exists a private equilibrium with one receiver but no public equilibrium. This can be viewed as the opposite to one sided discipline, where a firm decides not to communicate anything in public to the capital market due to the potential threat of entry. The opposite case where there exists more than one sender but only one receiver is discussed in Battaglini (2002) and will be presented in more detail in section 5.

An interesting more recent contribution is presented by Ottaviani and Squintani (2006) who provide an extension of Crawford and Sobel (1982) to cases where receivers are naïve and blindly believe senders messages. Under such assumptions there exist equilibria where communication is characterized by an *inflated language* and equilibrium outcomes are biased. However, information transmission is more precise than in Crawford and Sobel. In their comparative statics section, Ottaviani and Squintani show that educating receivers on the behavior of senders has a perverse, decreasing effect on information transmission. Rather, it is through reduction in the sender's bias that the amount of information transmission is increased. Inflation of language as presented in Ottaviani and Squintani (2006) constitutes another form of distortion and therefore a lie.

In games where a lie is to be successful, the liar must believe that his lie will be successful in a sense that may differ depending on the relevant situation. Different expectations driven by varying contexts are covered in the extensions that follow below.

[5] These can be thought of as possible entrants.

[6] Farell and Gibbons (1989) present an example in which a firm may lie to both shareholders and unions in private but were telling the truth is optimal in public. This is due to the opposite interests of the receivers. The truth becomes a kind of average.

4. Extensions 1: Modeling Successful Lies

Two models that include lying as active misrepresentation are presented. The first model, by Pitchik and Schotter (1987), describes a simple setting in which dishonesty not only exists but in which it may actually increase as agents interests converge. The second model by Crawford (2003) introduces different player types.

4.1. Case 1: Purchasing a Car repair

A consumer needs to purchase one of two possible repairs, one inexpensive (*I*) and one expensive (*E*), from an expert who can observe the true state of the world, i.e. which repair is required. The probability that an expensive repair is required is exogenously given as *r*, where *r* is assumed to be known by all consumers. The profitability of selling *E* is assumed to be greater than from selling *I*, i.e. $\pi(E \mid I) > \pi(I \mid I)$. Here π denotes profit and $|$ denotes conditional on needing. The profit functions are increasing in prices.[7] For consumers it holds that $u(E \mid E) > u(I \mid E)$ and $u(I \mid I) > u(E \mid I)$ where u is the consumer's payoff. This assumption implies that consumers prefer the appropriate type of repair required. Consumer's payoffs $u(I \mid \cdot)$ and $u(E \mid \cdot)$ are decreasing in repair prices.[8] Furthermore it is assumed that consumers are only able to detect whether the defect still exists after the repair has taken place. Purchasing *E* will remove any defect that could have been repaired with *I*. There are therefore no incentives to perform *I* when *E* is required since the fraud will always be discovered by the consumer. If the defect is gone, the consumer cannot judge whether the repair type purchased was the appropriate one.

The game proceeds as follows. First the expert offers the consumer one type of repair. Then the consumer chooses whether to accept or reject the offer. Here rejecting an offer implies obtaining a remedy somewhere else. The game can be reduced to the set of strategies illustrated in the table below, which should be read as follows: The first character in the consumer's strategy describes his response to an offer of *E*. *A* stands for accept and *R* for reject.

[7] This relationship is suppressed here in order to simplify the illustration.
[8] Note that $u(I \mid E)$ is decreasing in the price of E since the consumer ultimately buys the expensive repair in this case.

Strategy / Payoff Table

		Expert	
		TT	*TN*
Consumer	*AA*	(C_1, E_1)	(C_2, E_2)
	RA	(C_3, E_3)	(C_4, E_4)

The second character denotes the response to an offer of *I*. Similarly, the first character in the expert's strategy set tells whether he is willing to tell the truth *(T)* if the required repair is *E*. The second character corresponds to telling the truth or not *(N)* when the repair required is *I*.[9] In accordance with earlier assumptions on payoffs: $E_2 > E_1, E_3 > E_4, C_1 > C_3$ and $C_4 > C_2$ to prevent the expert from lying all the time.[10] The payoffs are calculated as follows for the customer:

$$C_1 = ru(E \mid E) + (1-r)u(I \mid I), C_2 = ru(E \mid E) + (1-r)u(E \mid I), C_3 = C_4 = ru(I \mid E) + (1-r)u(I \mid I)$$

And for the expert:

$$E_1 = r\pi(E \mid E) + (1-r)\pi(I \mid I), E_2 = r\pi(E \mid E) + (1-r)\pi(E \mid I), E_3 = (1-r)\pi(I \mid I), E_4 = 0$$ [11]

The equilibrium in this game is mixed. The expert will always tell the truth if he observes that *E* is required and he will tell the truth with probability $0 < 1 - \hat{p} < 1$ if *I* is required.

(1.1) $$1 - \hat{p} = \frac{r\left[u(E \mid E) - u(I \mid E)\right]}{(1-r)\left[u(I \mid I) - u(E \mid I)\right]}$$ [12]

[9] The reduced form presented adheres from the fact that *AA* and *RA* dominate *AR* and *RR,* and *TT* and *TN* dominate *NT* and *NN*, since an offer of *I* always is truthful in this model as discussed above.

[10] The assumption $C_4 > C_2$ is equal to imposing an upper bound on *r*.

[11] Note that E_3 and E_4 are lower since they include one and two rejected offers respectively.

[12] The expert chooses $1 - \hat{p}$ so as to make the consumer indifferent between accepting and rejecting the expensive offer. The consumer is indifferent

The consumer will always accept an offer of *I* and will accept an offer of *E* with probability $0 < \hat{q} < 1$, where

$$(1.2) \quad \hat{q} = \frac{\pi(I \mid I)}{\pi(E \mid I)} \text{ [13]}$$

The probabilities $\hat{p}$ and $\hat{q}$ assure indifference of the counterpart as is usual in mixed equilibria. The equilibrium payoffs are $\hat{q}\left[r\pi(E \mid E) + (1-r)\pi(E \mid I)\right]$ for the expert and $ru(I \mid E) + (1-r)u(I \mid I)$ for the consumer.[14]

A Simple Numerical Example

The following numerical values are assumed:

$u(E \mid E) = 200, u(I \mid I) = 100, u(E \mid I) = 50, u(I \mid E) = 0, \pi(E \mid E) = 800, \pi(I \mid I) = 400,$
$\pi = (E \mid I) = 1000$

and finally $r = 0,1$.

According to the definitions of C_i and E_i above, the following payoff matrix is obtained:

Expert

Consumer		*TT*	*TN*
	AA	(110, 440)	(65, 980)
	RA	(90, 360)	(90, 0)

when $pC_1 + (1-p)C_2 = pC_3 + (1-p)C_4 \rightarrow p = \dfrac{C_4 - C_2}{C_1 - C_2 - C_3 + C_4}$ but $C_3 = C_4 \rightarrow p = \dfrac{C_4 - C_2}{C_1 - C_2}$. It now follows that $1 - p = \dfrac{C_1 - C_2}{C_1 - C_2} - \dfrac{C_4 - C_2}{C_1 - C_2} = \dfrac{C_1 - C_4}{C_1 - C_2}$

13 The consumer chooses $\hat{q}$ so as to make the expert indifferent between lying and not. The expert is indifferent when $qE_1 + (1-q)E_3 = qE_2 + (1-q)E_4 \rightarrow q = \dfrac{-E_3}{E_1 - E_2 - E_3}$

14 Note that the consumers equilibrium payoff is independent of $\hat{p}$ since $C_3 = C_4$

It is instantly clear that there exists no pure strategy equilibrium in this game. The mixed strategy equilibrium, calculated according to the formulas presented above, involves the following probabilities:

$$\hat{p} = \frac{90-65}{110-65} = \frac{25}{45} = \frac{5}{9} \rightarrow 1-\hat{p} = \frac{4}{9} \text{ , } \hat{q} = \frac{-360}{440-980-360} = \frac{-360}{-900} = \frac{2}{5}$$

The equilibrium payoffs become:

$\frac{2}{5}\left[0{,}1 \cdot 800 + 0{,}9 \cdot 1000\right] = 392$ for the expert and $0{,}1 \cdot 0 + 0{,}9 \cdot 100 = 90$ for the consumer.

4.2. Comparative Statics of the Model

In their comparative static's section, Pitchik and Schotter (1987) discuss the dynamics of price changes on the respective repairs. If, for example, the price of the expensive repair is decreased, then $\pi(E \mid I)$ decreases, while the difference $\pi(E \mid E) - \pi(E \mid I)$ remains the same. Furthermore, $u(E \mid E)$, $u(E \mid I)$ and $u(I \mid E)$ increase, while the difference $u(E \mid E) - u(I \mid E)$ remains the same. From (1.1) and (1.2) it follows that $\hat{p}$ and $\hat{q}$ both increase. Thus, a decrease in the price of E implies more dishonesty from the expert and a higher probability of acceptance from the consumer. The intuition is that by decreasing the price of the more expensive repair the expected cost of following the expert's advice is decreased. The consumer can be less particular about choosing an expensive repair. As Pitchik and Scotter phrase it, the expert can exploit the consumer's loss of vigilance. These results are confirmed in a continuation of the numerical example.

Numerical Example Continued

Assume that the decrease in price has the following effects:

$u(E \mid E) = 250, u(I \mid I) = 100, u(E \mid I) = 50, u(I \mid E) = 50, \pi(E \mid E) = 600, \pi(I \mid I) = 400,$
$\pi = (E \mid I) = 800$

The new payoff matrix becomes:

		Expert	
		TT	*TN*
Consumer	*AA*	$(115, 420)$	$(69, 780)$
	RA	$(95, 360)$	$(95, 0)$

It follows that $\hat{p} = \frac{13}{23} > \frac{5}{9}$ and $\hat{q} = \frac{1}{2} > \frac{2}{5}$ which confirms the theory above.[15]

4.3. Comparison with Crawford and Sobel

The model distinguishes itself from Crawford and Sobel (1982) through its different comparative statics. In Crawford and Sobel's continuous setting, as *b* approached zero, i.e. agents preferences grew more similar, information revelation was increased through a finer partitioning equilibrium. In Pitchik and Schotter (1987) as the agents become more similar, honesty of the expert does not increase. While there exists no explicit preference similarity parameter such as *b*, Pitchik and Scotter assume that agents' preferences can be made more similar by letting their payoffs converge through price changes. The different similarity assumptions between the models cause different outcomes. In Pitchik and Scotter's discrete setting, closer payoffs may not imply changes in payoff optimizing actions. As a result $\hat{p}$ does not always decrease when agents become more similar and neither are they both better of which is in contrast to Crawford and Sobel (1982).

[15] Note that $u(E \mid I) = 50$ is the same as before. In reality one might have assumed a small rise but for simplicity in calculations this payoff has been left unchanged.

4.4. Case 2: Attacking the "Wrong" Target

A second model that illustrates the possibility of lying for strategic advantage, as successful misrepresentation of preferences, is presented in Crawford (2003). In his model Crawford focuses on the Allies attack on Normandy as a case of successful misrepresentation.[16] The key feature of this model is the existence of multiple player types. Different types have different levels of rationality (intelligence in a sense). Here the case is generalized and simplified to facilitate the explanation of the key features.

Decision Matrix

		Receiver	
		Defend Left	*Defend Right*
Sender	*Attack Right*	-a a	0 0
	Attack Left	0 0	-1 1

The Game

The players and their strategies are illustrated in the matrix above. The game, which is based on the class of two player, zero sum, perturbed matching pennies games, features a sender (who can be thought of as an attacker) and a receiver (defender). The relationship **a**>1 is assumed to hold.[17] Before the game is played the sender sends a nonbinding, costless, noiseless message in which he states whether he will play *Attack Right* or *Attack Left*. Thereafter the players choose actions simultaneously. The game setting is common knowledge.

In a "standard" analysis, this game would feature an uninformative message and a mixed equilibrium depending on the size of (**a**), thus showing some similarity to Pitchik and Schotter (1987). Communication is ineffective and misrepresentation is unsuccessful.

---

[16] The Germans expected an attack at Calais which was more probable ex ante.
[17] This can be thought of as **a** being the easier object to attack.

To make things interesting it is now assumed that both players can assume two different types. The player types consist of a rationally bounded, so called *mortal*, player and a fully strategically rational, so called *sophisticated*, player. Although both players are not aware of the other player's type, they are both familiar with the overall distribution of both types. Mortal players are rationally bounded in the sense that although they maximize payoffs, their beliefs that rationalize their behavior generally differ from equilibrium beliefs. The mortal players use a step-by-step procedure that determines unique, pure strategies but avoids any determination of simultaneous decisions used to define equilibrium.[18] This independent determination of Mortal's strategies with regard to each other as well as with regard to sophisticated players allows for an exogenous treatment of this player type. Therefore it suffices to focus on sophisticated players of each role in a reduced game. However, the results of this reduced game are altered through the existence of mortal players.

In the reduced game between sophisticated players, these have to weigh the responses of sophisticated opponents against those of mortal opponents. This implies that the game is no longer necessarily zero-sum since mortal opponents payoffs may differ from those of sophisticated opponents. Information in this reduced game is incomplete. When a sender sends a message regarding his intentions the message reveals his type to a sophisticated receiver.

4.5. Explaining Mortals Behavior

The nature of the sequential equilibrium in the reduced game depends on the prior probability distribution of mortal players' behavior.[19] Mortal senders and receivers can assume two different subtypes. Mortal senders can either be liars or truthtellers whereas mortal receivers are represented by either believers or inverters.[20] The mortal senders expect their attempt to misrepresent their preferences to succeed and therefore respond to the payoff advantage in **a**>1. They thus try to induce the receiver to play *Defend Left* where after they play *Attack Right* on the equilibrium path. The underlying assumption

[18] According to Crawford (2003) this assumption is empirically plausible in communication games.
[19] A sequential equilibrium specifies not only a strategy but also a belief for each of the players.
[20] Inverters are characterized by their skeptical approach towards the signal received. They therefore invert the signal received (i.e. play the opposite from what is recommended by the message)..

here is that not only senders, but also receivers are convinced that their decision rule will be successful in outmaneuvering the counterpart. Crawford lists a set of decision rules which are appointed different "levels" of rationality, starting at level 0. Believers and truthtellers are assumed to represent level 0, 2, 4 etc. Inverters and liars represent level 1, 3, 5 etc. The table below illustrates the idea according to which mortal types act.[21]

Table: Rationality levels of boundedly Rational Players

Sender rule	Behavior	Receiver rule	Behavior
Credible $\equiv W0$	tells the truth	Credulous $\equiv S0$	believes (b.r. to $W0$)
$W1$ (Wily)	lies (b.r. to $S0$)	$S1$ (Skeptical)	inverts (b.r. to $W1$)
$W2$	tells the truth (b.r. to $S1$)	$S2$	believes (b.r. to $W2$)
$W3$	lies (b.r. to $S2$)	$S3$	inverts (b.r. to $W3$)

b.r. $\equiv$ best response

Each decision rule, apart from the level $W0$ is a best response to the associated rule one level below. Every mortal player believes that he is one level above his counterpart.

4.6. Equilibrium with Successful Misrepresentation

Crawford (2003) shows that when probabilities of sophisticated senders and receivers are relatively high, the equilibrium will be of a mixed type very much similar to the mixed strategy equilibrium in a standard (no different player types) game. However, when probabilities of sophisticated sender and receiver types are relatively low, there exists a case where one sophisticated sender fools a sophisticated receiver in equilibrium. If the probability of a single mortal sender type is high enough, a sophisticated sender may pool with the mortal senders message. Therefore, if the probability of a sophisticated sender is low enough and the probability of a believer type mortal receiver is above the intermediate range, there exists a sequential equilibrium in which a sophisticated sender can fool a sophisticated receiver. An example could look as follows: Assuming that the

[21] The notation of levels of rationality as $W0$ and $S0$ etc. is used by Goffman (1969) and also in Crawford (2003).

mortal sender type predominantly consists of truth tellers who send *Attack Right*, a sophisticated sender will pool and send the message *Attack Right* but will play *Attack Left* whereas both the sophisticated receiver and the believer play *Defend Right*. Since mortal senders will play *Attack Right* (in response to the expected payoff advantage) the sophisticated receiver will play *Defend Right* in any case. It follows that there can be no equilibrium in which a sophisticated receiver plays *Defend Left*, but a sophisticated sender plays *Attack Right*. This game structure is founded on a setting in which, ex ante, sophisticated receivers are prepared and perhaps more willing to be fooled by sophisticated senders than being fooled by mortals given the ex ante probability distributions of all player types.

Players' payoffs differ depending on the distribution of mortal versus sophisticated types. When the probabilities for sophisticated player types are high, mortals and sophisticated players have identical expected payoffs. The existence of sophisticated players on both sides thus protects the mortal players from exploitation. When probabilities of mortals are high however, sophisticated player types have higher payoffs than their mortal counterparts since they can avoid being fooled except when they choose to (choose the lesser evil).

To summarize, the existence of different player types can, under certain distributive circumstances, allow successful misrepresentation of preferences, i.e. a sophisticated opponent can be fooled. An interesting aspect of this model is that it represents a class of situations where no information is transmitted in Crawford and Sobel (1982) due to the large difference in b. Here, a certain amount of information is indeed transmitted in equilibrium. This is due to the existence of the different player types where the messages of the mortal player types are deciphered by their sophisticated counterparts.

5. Extensions 2: Inducing Truthful Revelation and Lie-aversion

There are contexts within which lying is disadvantageous or simply not pursued. The next model shows situations in which the policy is of a multidimensional character. The final case is based on empirical experiments. As a consequence of the observed results the term lie-aversion is introduced.

5.1. Case 3: Multidimensional Cheap Talk

This setting originally presented in Marco Battaglini (2002) differs from previous representations by introducing a multidimensional decision variable. In the multidimensional case Battaglini shows that full revelation of information is generically possible regardless of how large the conflict of interest between players is. Multidimensionality of a problem is prominent in, e.g. political decision-making, where issues generally have several different areas of impact.[22]

The game has three players, one policy-maker and two experts. The set of alternatives for the policy-maker is denoted as $Y \equiv \Re^d$.[23] For any policy $y \in Y$, the outcome is $x = y + \theta$ where θ is a $d-$dimensional vector in $\vartheta \equiv \Re^d$, implying that there is a distinction between the policy space and the outcome space $x \in X$. Nature chooses θ according to a continuous distribution function $F(\theta)$ with density $f(\theta)$, supported on ϑ and with a zero expected value. In choosing y the policy-maker has no knowledge of θ. The realization of nature is only observed by the experts. Each of the three players has a von Neumann Morgenstern utility function $u_i : X \to \Re$. All agents are thus only interested in the policy outcome x rather than the actual policy chosen. The utility functions are assumed to be continuous, quasiconcave and differentiable.[24] The preferred points of the two experts are denoted as x_i where $i = 1,2$.[25] The policy-maker's preferred

---

[22] As an example consider the issue of the environment, where several goals are combined, e.g. reducing pollution and sustaining economic growth simultaneously.

[23] Where *d* represents the dimension of the policy and therefore also the maximum number of experts that have to be consulted.

[24] A function f defined over a convex set $S \subseteq \Re^n$ is quasiconcave if the upper level set $P_a = \{\mathbf{x} \in S : f(x) \geq a\}$ is convex for each number a (FMEA p.68).

[25] The set of experts is denoted E.

point is denoted x_p and is normalized to the origin. Quadratic utilities are assumed for simplification: $u_i(x) = -\sum_{j=1}^{d} (x_i^j - x^j)^2$ where x_i^j and x^j denote the j th coordinate of, i 's preferred point and the outcome x. Utility functions and preferred points are common knowledge.

The game proceeds as follows:

I. Nature chooses θ, which is observed by each expert, according to $F(\theta)$.

II. The experts simultaneously (privately) report θ to the policy-maker.

III. The policy-maker decides y and $x = y + \theta$ is realized.

A strategy for the policy-maker can be described as a function $y : \vartheta \times \vartheta \to Y$ where each couple of signals is associated with a policy in Y. The policy-maker's belief function is denoted $\mu : \vartheta \times \vartheta \to P(\vartheta)$. The belief function assesses a posterior belief P over ϑ for each pair of signals. Both experts' strategies are represented by a function $s_i : \vartheta \to \vartheta$. Thus, for each realization of nature the experts report a state of the world in ϑ. The equilibrium concept is the Perfect Bayesian Equilibrium.

The quintessential assumption of this model is that in a multidimensional policy space, agents with different preferred points share preferences over lower dimensional subsets. This implies that a sender and a receiver may have very different preferred points on a plane, but if they are forced to choose on a line in this plane, their preferred points are the same. Restricting a sender's influence to such a subset would leave no incentives for the sender to lie. The main result of the model is the generic existence of an equilibrium in which each sender only influences the dimension of common interest with the receiver and these dimensions are sufficient to identify the state of the world θ.

Battaglini (2002) distinguishes between a so called fully revealing equilibrium and a truthful fully revealing equilibrium. A fully revealing equilibrium does not necessary imply the experts telling the truth, it merely requires that the policy-maker "understand" the signal received. However Battaglini (2002) presents in a lemma the existence of a truthful fully revealing equilibrium whenever a fully revealing equilibrium exists.[26] In the case of a one dimensional state variable Battaglini (2002) shows that it is

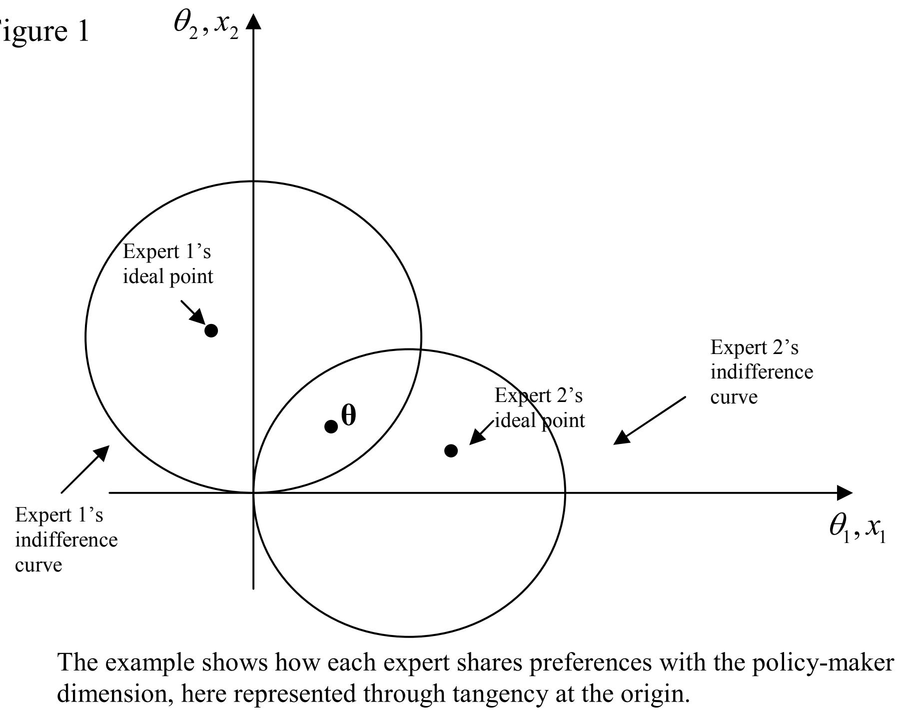
[26] See Battaglini (1999) for a proof.

possible to construct fully revealing equilibria but that these are implausible because they to a large extent rely on ad hoc constructions of out of equilibrium beliefs.

5.2. A Simple Example

Battaglini (2002) presents an example in which the policy-maker chooses a two dimensional policy $y = (y_1, y_2)$. There are two experts who may each send a signal to the policy-maker. The experts each observe the state $\theta \in \Re^2$ where θ is unknown to the policy-maker. The policy outcome is represented by the vector $x = y + \theta$. It is assumed that all agents have quasi-concave utilities over outcomes and the policy-maker's preferred point is normalized to the origin. Thus, the policy-maker wishes to minimize the distance between the outcome of the policy (x) and his preferred point. For example if he believes the optimal state to be μ, then his optimal policy would be given by $y = -\mu$. The policy-maker may listen to the experts' advice but bases his actual decision on his posterior beliefs. Figure 1 below presents a possible case.

Figure 1



The example shows how each expert shares preferences with the policy-maker in one dimension, here represented through tangency at the origin.

Each expert is only allowed influence in the dimension where his preferences coincide with the policy-maker's. Let $\theta = (2,2)$ and the policy-maker has the beliefs presented above, then the following system of equalities holds:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \theta_1 - s_1(\theta) \\ \theta_2 - s_2(\theta) \end{pmatrix} = \begin{pmatrix} 2-2 \\ 2-2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Any deviation from $s_1(\theta)$ for expert 1 is strictly negative since this would imply that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \theta_1 - \hat{s}_1 \\ \theta_2 - s_2(\theta) \end{pmatrix} = \begin{pmatrix} 2-\hat{s}_1 \\ 2-2 \end{pmatrix} = \begin{pmatrix} 2-\hat{s}_1 \\ 0 \end{pmatrix}$$

Assume $\hat{s}_1 = 3 \rightarrow x = (-1, 0)$. The above result implies that since expert 2 plays the optimum strategy, any movement achieved through deviation by expert 1 will be along the horizontal axis. Since tangency of expert 1's indifference curve on the horizontal axis occurs at $x_1 = 0$ no deviation is optimal for expert 1. If expert 1 is only allowed influence in x_1 and expert 2 in x_2 the optimal outcome given the other players strategy is thus the origin.

The example presented in figure 1 is of course a special case in which the indifference curves are tangent to the coordinate axes at the policy-maker's preferred point. The same reasoning however holds in general for quasiconcave utilities. This is shown in the generalization that follows.

5.3. Existence of a Fully Revealing Equilibrium in Two Dimensions

General existence of a fully revealing equilibrium in two dimensions is shown through a proposition and the corresponding proof presented in Battaglini (2002). This case can be further generalized to cases with more than two dimensions and more general utility functions but for the purpose of this paper, the explanation below should suffice.[27]

Proposition: *If $d = 2$, then for any x_1 and x_2 such that $x_1 \neq \alpha x_2 \ \forall \alpha \in \Re$, there exists a fully revealing, robust equilibrium.*

Proof: Some definitions are in order to begin with. For any $\alpha \in \Re$, and for any $i = 1, 2$, define

$$l_i(a) \equiv \{z \in \Re^2, \nabla u_i(0,0) \cdot z = a\}$$

[27] Battaglini (2002) proves the theory under more general assumptions following his main discussion.

Figure 2 presents an illustration. Here, $l_i(0)$ can be interpreted as the tangent of the indifference curve of the *i*th agent at the preferred point of the policy-maker. For any $a \in \Re$, $l_i(a)$ identifies one and only one line parallel to $l_i(0)$. Since there exists no $\alpha \in \Re$ such that $x_1 = \alpha x_2$, $l_1(0)$ and $l_2(0)$ are linearly independent vector spaces. Linear independence of the preferred points implies linear independence of the gradients of utility at the origin and therefore it implies linear independence of any $\alpha \in l_1(0)$, $\beta \in l_2(0)$ since they are orthogonal to $\nabla u_1(0,0)$ and $\nabla u_2(0,0)$ respectively. It follows that $\forall\, \theta \in \Re^2$, there exists a unique vector $(a_1, a_2) \in \Re^2$ such that $\theta = l_1(a_1) \cap l_2(a_2)$. Define the function $a(\theta): \Re^2 \to \Re^2$ that, for each θ, associates the couple $a_1(\theta), a_2(\theta)$ uniquely defined by the previous equality. In addition $\forall (a,b,c,d) \in \Re$:

$$l_i(a) \cap l_j(b) + l_i(c) \cap l_j(d) = l_i(a+c) \cap l_j(b+d)$$

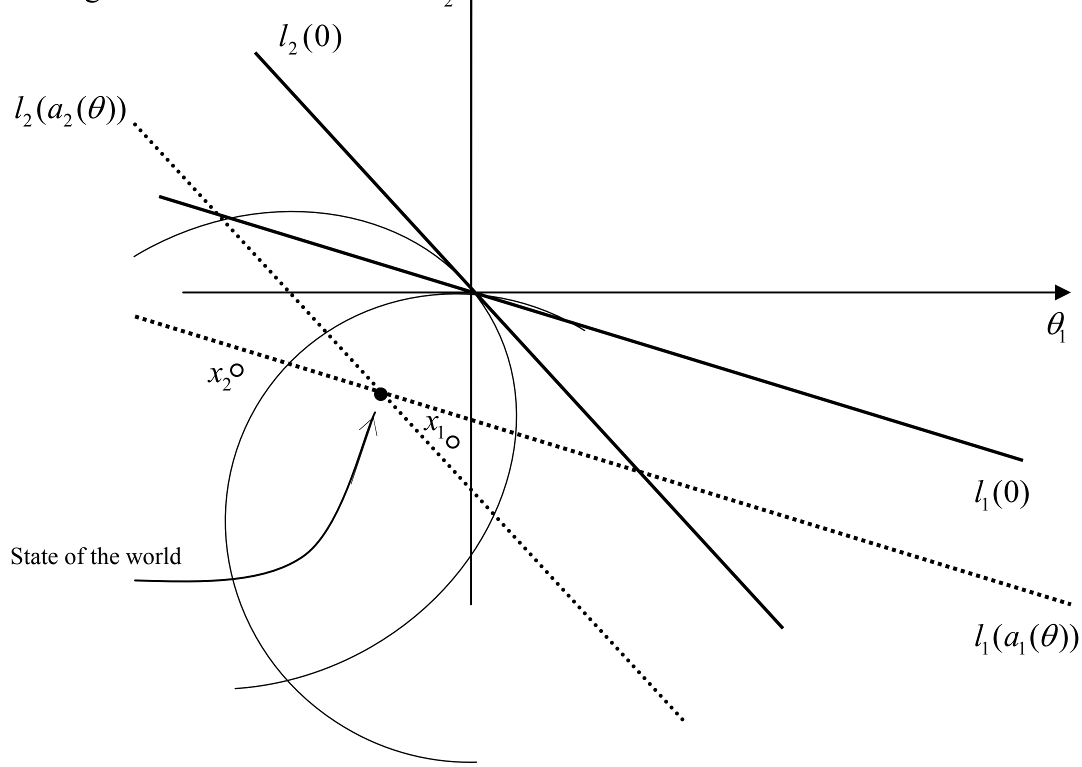
The proof of the proposition now follows. Each expert is required to report a number, s_i. Consider the following strategies and beliefs:

(1.3) $\quad s_i(\theta) = a_j(\theta) \quad \forall i, j = 1,2, \qquad i \neq j$

(1.4) $\quad \mu(s_1, s_2) = l_1(s_2) \cap l_2(s_1),$

(1.5) $\quad y(s_1, s_2) = -\mu(s_1, s_2).$

Figure 2



It is possible two construct a new coordinate system to exploit experts' conflict of interest

The claim is that the above strategies and beliefs constitute a robust equilibrium. Given the other players' strategies, player *I*, choosing $ś_i$, may induce a point:

$$\theta - \mu(s_j(\theta), ś_i) = \theta - l_i(s_j(\theta)) \cap l_j(ś_i) \qquad \text{by (1.4),}$$

$$= \theta - l_i(a_i(\theta)) \cap l_j(ś_i) \qquad \text{by (1.3),}$$

$$= l_i(a_i(\theta)) \cap l_j(a_j(\theta)) - l_i(a_i(\theta)) \cap l_j(ś_i) \text{ by definition of } a(\theta),$$

$$= l_i(0) \cap l_j(a_j(\theta) - ś_i) \text{ by the claim above.}$$

$ś_i$ is any number in $\Re$, so agent *i* can choose any value for $(a_j(\theta) - ś_i)$ implying any point in $l_i(0)$. But since u_i is tangent to $l_i(0)$ only at the origin (the preferred point of the policy-maker) this is the optimal outcome that *i* can induce. The optimal strategy is therefore to set $ś_i = a_j(\theta)$, as claimed above, since there is no profitable deviation for agent *i*, $\forall i = 1, 2$. This equilibrium contains consistent beliefs and an optimal policy choice given these beliefs. Robustness is guaranteed since there are no out-of-equilibrium message pairs. Thus beliefs are always well defined. *Q.E.D.*

In the general case the same result is achieved as in the introductory example by creating a new coordinate system which requires only that x_1 and x_2 are linearly independent. Assuming quadratic utilities, the tangents at $(0,0)$ of the respective utilities are linearly independent so they span. When agent *i* is required to choose an outcome in $l_i(0)$, he chooses $(0,0)$, the policy-maker's preferred point. This happens since agent *j* will be honest on the l_j dimension which forces *i* to choose in $l_i(0)$.

To summarize the above, the policy-maker makes each sender influential only in the dimension of common interest. The dimensions are chosen so that they span the entire outcome space. The combination of signals will make it possible for the policy-maker to extract all available information. In this model it is not the distance between two preferred points that is relevant, but rather it is their linear independence. This implies that information is transferred regardless of preference dissimilarity.

5.4. Case 4: Lie-aversion

The discussion presented in Gneezy (2005) serves to round up the series of extensions presented in this paper. Gneezy introduces results based on an empirical experiment in which individuals are faced with no explicit penalty when they lie. However, lying will have so called *consequences* that take the shape of differing states of wealth for individuals depending on whether the truth or a lie is told.

Gneezy distinguishes between four different types of lies. There are lies that don't hurt anyone, or are beneficial to the liar and the counterpart, e.g. lying with compliments. Second there are lies that help the counterpart and hurt the liar, e.g. cases of altruism. The third category involves lies that hurt all or at least the counterpart, e.g. acts of nemesis in response to unfair behavior. The forth category is the most relevant for economic theory in general and it is even more prominent in cheap-talk games. These lies include all those that make the liar better of at the expense of the counterpart. Note that all models discussed in dept in this paper are built on an assumption of, or close to, this last type.

An overview of the cheap talk experiment conducted by Gneezy (2005) is presented below so as to make clear the intuition behind his results. A sample of students is divided into senders and receivers. Both player types are informed about two options (A and B) that yield differing payoffs to themselves and to the counterpart. The private information on the side of the sender consists of the knowledge regarding the different payoffs of A and B. The sender first sends one of two possible messages to the receiver:

1. "Option A gives you a higher payoff than option B".
2. "Option B gives you a higher payoff than option A".

Upon receiving either message, the receiver chooses which option is to be implemented.

The usual equilibrium outcome in this type of game where preferences are standard and objectives are in conflict (which all players are aware of) includes a message with no information whatsoever. In this experiment however, the receiver is not aware of the alignment of incentives.[28]

[28] Note that a technical presentation would have to include the receivers beliefs on the alignment of incentives!

The element of primary interest is of course the message sent by the sender. The content of this message is affected by the beliefs of the sender regarding the receiver's reaction. Participants in the experiment where questioned regarding their beliefs concerning the receiver's response. Upon questioning, approximately 80% of the respondents assumed the receiver would follow the message sent to him. After conducting the experiment, it turned out that almost 80% of the receivers had indeed followed the sender's message. These findings imply that one could generally expect senders to believe that their advice will be followed. If senders in turn are selfish, they will always choose the option yielding them the highest payoff.

Now consider the following three pairs of alternatives in the table below.

Table with game types and respective payoffs

		Payoff to	
Game Type	**Alternative**	**Player 1**	**Player 2**
1	**A**	**5**	**6**
	B	**6**	**5**
2	**A**	**5**	**15**
	B	**6**	**5**
3	**A**	**5**	**15**
	B	**15**	**5**

In the above game types message **B** represents the lie as it gives player 2 (the receiver) a lower payoff than message **A**. For game type 1, the result displayed a lying ratio of approximately 1/3. In game type 2, where the loss for player 2 is significantly increased, the amount of liars dropped to less than 20%. Finally, in game type 3 about half the senders (player 1) decided to lie.

In order to identify the possible existence of lie-aversion Gneezy performed an alternate so called dictator treatment in which the senders were simply allowed to choose which of the two allocations they preferred in each game type. In this setting receivers had no say but player 1's choice was implemented only 80% of the time so as to resemble the situation in the deception game, thereby making the treatments equivalent.[29] The idea is that if the sender now chooses the materially advantageous allocation more frequently this would count as evidence in favor of lie-aversion. The results from the dictator setting, although displaying the same pattern as in the deception game above, included a far higher frequency of self interested choice; e.g. about 90% of the participants decided to lie in the case that resembled game type 3.[30]

The above results give reason to believe that lie-aversion exists in addition to care about others in general. Selections between payoffs are treated differently depending on if they involve a preceding lie or a simple "dictated" choice, as Gneezy describes it: "…people are not indifferent to the process leading up to the outcome." The fact that individuals seem to care not only about their own gain but also about the potential losses for the counterpart is also intuitively appealing. Gneezy presents the example of a simple car purchase. Individuals can generally trust a seller's assurance on functional brakes more than on the standard of air conditioning.

[29] In 20% of the cases the alternative set of payoffs is implemented.
[30] In the cases that resembled game type 1 and 2 lying was up to about 40% and 2/3 respectively.

6. Analysis

This section characterizes how and when individuals can be expected to lie. To facilitate the illustration, a summary of the results obtained in the different models introduced so far is presented in the liars table below.

Liars Table

Gametype	Sender Lies	Set of Strategy Choices (Static/Continuous)	Nr. of Receivers	Nr. Of Senders	Aligned Preferences over Possible Outcomes Required	Receiver has Knowledge of Game Structure
C&S (1982)	No	Continuous	1	1	Yes	Yes
C&S (1982)	Yes	Continuous	1	1	No (but can differ more or less)	Yes
Farrell/ Gibbons	Yes (Subversion)	Discrete	>2	1	No	Yes
Farrell/ Gibbons	No (Possibly One Sided Discipline or Mutual Discipline)	Discrete	>2	1	No	Yes
Battaglini	No	Continuous	1	>1	No (but senders preferred states are linearly independent)	Yes
Pitchik/ Schotter	Yes	Discrete	1	1	Not completely (determined as proximity of expected payoffs)	Yes
Pitchik/ Schotter	No	Discrete	1	1	Yes (determined as proximity of payoffs)	Yes
Crawford (2003)	Yes	Discrete	1	1	No	Yes, but actual receiver/sender types are given a certain probability
Gneezy	In general No	Discrete	1	1	No, but big loss for Receiver and not too big gains for Sender associated with lie	No
Gneezy	In general Yes	Discrete	1	1	No, and big gains associated with lying for Sender	No

A discussion of how individuals lie presupposes that there exist different types of lies. Intuitively such a conclusion is not difficult to establish. As mentioned above, lying can take place in a several ways, e.g. withholding information, inflating information, deflating information and active misrepresentation. However, not all of these opportunities to be untruthful are always available. The setup of the game, or more precisely the available strategy sets for sending messages and choosing reactions, determine which lies are possible.

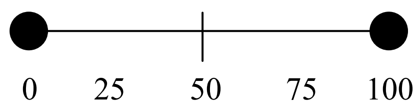
6.1. Lying in Discrete and Continuous Strategy Type Situations

The column denoted "set of strategy choices" in the liars table above shows that there are models where the strategy set is discrete and where it is continuous. An example of a discrete choice is illustrated below.

Figure 1: Discrete Strategy Choice

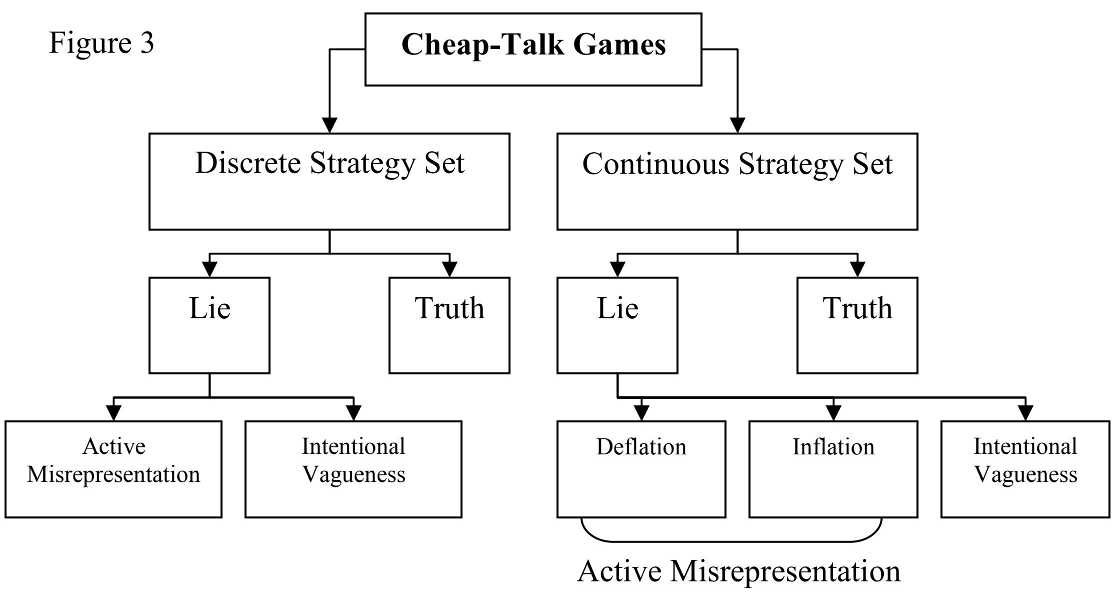
Above, there are only two choices, two extremes one might say. If the sender observes one and reports the other, he lies through active misrepresentation. If he decides to report both outcomes as possible, he is intentionally vague as in Crawford and Sobel. There are no other possibilities to lie in this game. A continuous version is presented in figure 2.

Figure 2: Continuous Strategy Choice



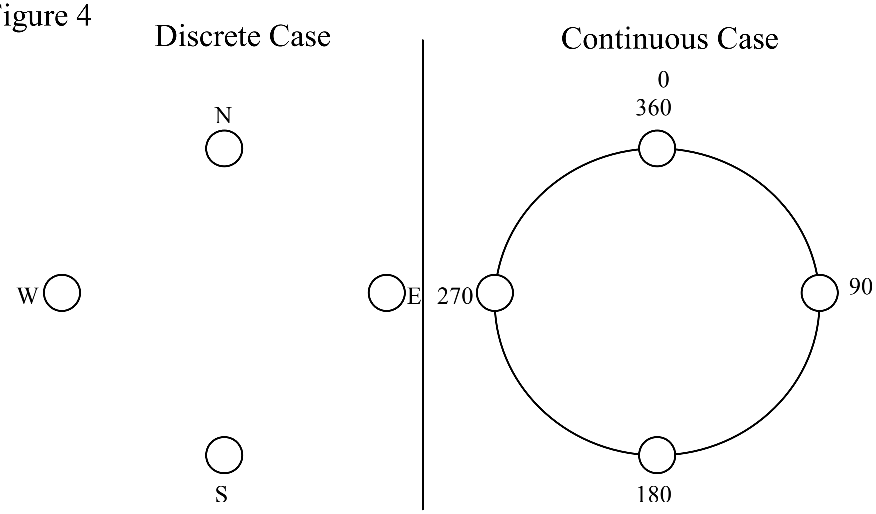
In the continuous strategy setting, the observed outcome and the strategy choices may take on any value between 0 and 100. Untruthful behavior can now take on more shapes in addition to those already mentioned in the static case. If the true state of nature were 70 for instance, a sender could inflate or deflate language by reporting 80 or 60 respectively.

Figure 3



As the overview in figure 3 reveals, deflation and inflation of language become equal to the discrete case active misrepresentation in their extremes. Note that games with discrete strategy sets are not restricted to two possible outcomes. Consider the following example with four strategies presented in figure 4.

Figure 4



The key difference between the discrete and the continuous case is not the amount of strategies per se, but whether choices exist that are situated in between the opposite poles (just as *maybe* lies between *yes* and *no*). Distinguishing amongst cases in this type of manner does have the drawback that there may be "close calls" but they will have to be regarded as a necessary evil.

6.2. Levels of Sin

When and how often lying occurs is determined by how individuals perceive lying as an alternative. Crawford and Sobel (1982) and Battaglini (2002) both have treatments in which a lie occurs simply because of a difference in preferences between the sender and the receiver. Gneezy (2005) claims that a decision to lie is not that simple since lie-aversion can be observed amongst senders. These different treatments may however both be correct. Lying in continuous strategy environments is easier due to the extended scope of possible lies. Furthermore, lying through inflation, deflation or intentional vagueness is not viewed as equally evil by the counterpart. Often such distortions are not even considered lies. There are surely several philosophical factors involved in determining an individual's treatment of an informational distortion and these will not be examined here.

There may however be a simple explanation that can account for much of the difference in perception of lies. Lying in continuous strategy set environments will in general still transmit some information to the receiver. Saying *100%* when the truth is *0%* transmits no information. Reporting *80%* when *70%* is true is not as serious since it allows a conclusion that the answer is at least not *0%*.[31] It is important to add that that the environment in which lying occurs can be equally important. Lying on a test is cheating regardless of to which extent. Lying with facts and figures in a political campaign can be viewed differently depending on the type of lie used. According to Gneezy (2005), senders care not only about their own payoffs but also about receivers'. Lie aversion in continuous environments may be reduced since senders know that their lie still transmits some information to the receiver, allowing a decision that will benefit him to some extent.

Active misrepresentation of outcomes or preferences in a discrete strategy environment is the most evil of lies since it transmits no useful information to the receiver. Moral barriers and honor code should not be underestimated as preventions for this type of lie, but still it does occur. Explanations may be many and there will be no attempt to even try and create an exhausting list. However, some factors that may affect lie aversion in discrete strategy set situations are of economic interest.

In Gneezy's experiment the receiver is unaware of the structure of the game. The sender is aware of this fact. In a game where the rules are not common knowledge fooling an opponent could be considered more evil and thus associated with higher lie aversion. On the other hand, if everyone knows or even expects that others cheat then lie aversion is lower.

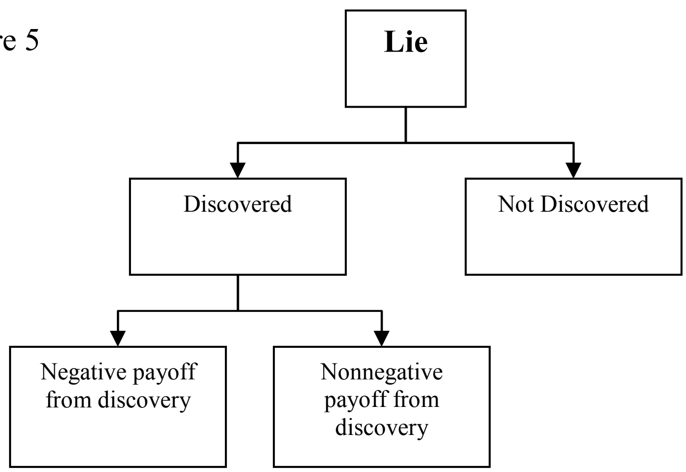
There is no information regarding player's total wellbeing in Gneezy (2005). In reality, decisions may often be affected by knowledge regarding the overall situation of the counterpart. If a decision to lie by the sender will completely ruin the receiver one might expect the sender to be somewhat less willing to lie. On the other hand, if the sender's lie is very costly to the receiver, but the receiver is rich and will be so after the consequences of the lie have been worn off, there might be less of a disincentive to lie (a Robin Hood effect!). The key point is that the actual change in welfare for the counterpart

[31] When the game structure is common knowledge the receiver is aware of the available strategy-set for the counterpart and for himself and therefore knows what type of lie he can expect.

does not so much create lie aversion as the relation of this change to the counterpart's welfare reference level. Gneezy (2005) points out that deceiving corporations is perceived as less evil than lying to individuals.

For a sender to consider a lie, he must be confident in its success. Discovery of a lie is associated with a negative payoff unless the sender does not care about the receiver at all. This is not in accordance with Gneezy (2005) but natural in Crawford (2003). A lie concerning an attack location will of course be discovered, but since the decision to lie is taken in order to increase harm to the receiver this is not a deterrent. In many cases however, possible discovery of a lie is a large deterrent. Senders might only lie if they believe receivers will not be able to notice ex post. This is the case in Pitchik and Schotter (1987). The possibilities of active misrepresentation are illustrated in figure 5.

Figure 5



Both Crawford (2003) and Pitchik and Schotter (1987) assume situations where players have different levels of competence. In an extension Pitchik and Schotter expand their model to include cases where the seller (sender) can vary in competence, which is a step in the direction of Crawford (2003). The distinction of player types in Crawford's model is an appealing explanation of the presence of successful lies. One might think of the different player types as players who have spent differing amounts of resources on something, giving them a diverse level of understanding of the actual situation. The important difference between Crawford (2003) and Crawford and Sobel (1982) is, as mentioned above, that in the prior information is transmitted under conflict of interest. Fooling an opponent is possible even when information is transmitted due to the diverse level of understanding. An interesting difference between Pitchik and Schotter (1987) and

Crawford and Sobel (1982) is the dissimilarity regarding the interpretation of convergence of interest. The preference similarity parameter *b* in Crawford and Sobel seems more appealing than the payoff similarity point of view in Pitchik and Schotter. Crawford and Sobel's presentation allows more successful modeling of additional possible situations. These include classes of lying, as described in Gneezy (2005), where the sender may be altruistic.

6.3. Effects of One Sided Multiplayer Presence

The introduction of additional senders or receivers significantly alters and often diminishes the possibilities for successful misrepresentation. That lying to many is more difficult than lying to few seems self evident, but that one additional individual can completely change the state of information transmission may seem less intuitive. As shown in Battaglini (2002), and Farell and Gibbons (1989), certain preference assumptions are required to hold for a revealing equilibrium to be optimal. The idea is simpler in the case with several receivers, where numerous examples illustrate the validity of the theory.

A theory of limited influence as presented in Battaglini (2002) is attractive for policymakers who are faced with competing lobby groups. Unfortunately it cannot be excluded that the experts collude. Battaglini recognizes that Collusion proofness is not generally the case of the truthful equilibrium outcome. If experts can collude they will often find another equilibrium point that will not conform to the policy-maker's preferred choice. Another issue is that for the state of the world to be completely and truthfully revealed the policy-maker needs to share preferences with the experts in at least one sub-dimension. This assumption is not especially restrictive if one allows for compromises on the side of the receiver, i.e. the policy-maker can choose to share the preference of a sender even if this does not induce his preferred point as long as the outcome is better than what could have been expected from the beginning.

A key feature in Battaglini's model is that individuals care only about the outcome and not about the policy that induces it. This assumption stands in contrast with Gneezy (2005) who points out that there seems to be support for the assumption that individuals do care about the actual process leading to the outcome. One must keep in mind however, that Gneezy makes this observation after comparing situations where on the one hand a lie leads to an outcome and on the other a dictator-setting enables a free choice of outcomes for the sender. The equilibrium created in Battaglini (2002) is one where lying is not optimal so it will not be pursued.

What might be the implications of introducing boundedly rational players into multiplayer models? On the sender side one could perhaps visualize this as a case where lobby groups spend differing amounts of money on their campaign. Some lobby groups may therefore not be adequately informed about reaction functions of the policy-makers (receivers). Given certain probability distributions of player-types there could be situations where the true state of the world is not completely revealed since sophisticated groups utilize the less informed, boundedly rational ones to misrepresent the truth. In other words, one might expect the results of Crawford (2003) to hold in multiplayer games as well. Such a difference would only be effective if there existed no sophisticated player with interests in a given dimension and such an assumption may not find large support in reality.

In general, additional players on either the sender or receiver side reduce the attractiveness of lying. If multiplayer presence is on the sender side then the receiver can under circumstances effectively limit the dimension of senders influence and thereby induce truthful behavior as optimal. Multiple receivers can induce truthful behavior by enforcing public communication.

6.4. Optimal Policies

The results derived above yield that lying occurs predominantly in cases where:

1. A continuous strategy set exists.

Or in discrete strategy set environments when:

2. The players are heterogeneous in their level of rationality (intelligence).
3. Lies cannot be detected ex post or the payoff from discovery is nonnegative for the sender.

In deciding how to develop efficient policies that support truthful information revelation it would be of advantage to have data that gave some sort of indication as to when lying costs the most. Even without such data some results can be derived. Regardless of the costs to society associated with the act of lying itself, policies aimed at raising the subjective costs of lying can help reduce active misrepresentation in discrete games, but may be less successful in games with continuous strategy sets. In a continuous strategy environment lying takes a less explicit form and liars face less lie-aversion. To reduce lying in continuous strategy set situations one can try to design policies that necessarily create multiplayer situations which allow induction of preferred outcomes through limited influence or one sided/mutual disciplining. If the existence of boundedly rational players makes lying feasible, policies aimed at reducing liars bias may be more efficient than educating the less intelligent (naïve) players (Ottaviani and Squintani (2006)).

Economic reasoning may not always justify active measures aimed at increasing truthful revelation. Lying may have merely redistributive or even welfare enhancing consequences for society. In deciding when lying can have non-negative effects on overall welfare a relevant factor will be how high the inherent value of truthful revelation is. The value of the truth as determined by society is of course a crucial factor in explaining the existence and size of lie-aversion. How to measure the value of truthful revelation and how such value may vary over different lies are interesting questions that require more than game theoretic exposure to be answered.

7. Conclusion

Lying is not simply a homogeneous activity but can take several forms such as inflation, deflation or withholding of information. Distortions in information transmission vary in type depending on the available strategy set for senders and receivers. When continuous strategy sets are available, lie-aversion is lower since some information is still transferred even when transmission itself is distorted. In a discrete strategy type setting, lie-aversion can be overcome when the lie is expected to succeed. This may happen because the lie cannot be discovered ex post or because players are heterogeneous in knowledge regarding the subject. Differences in knowledge between senders and receivers allow some information to be transferred under conflicting preferences. Lying can often be avoided in situations where multiple senders or receivers are involved since the multiplicity of players allows placement of restrictions that can bind senders to the truth.

A lie is necessarily an action of self interest but individuals seem to care about more than themselves. They can be expected to take into account the change in the welfare reference level of the counterpart when deciding to lie. Lie-aversion seems to be stronger in situations where the counterpart is not aware of the structure of the game and therefore may not be able to successfully anticipate the occurrence of lies.

Knowledge regarding the relevant type of lie is important if one is to be successful in implementing policies aimed at increasing truthful revelation. How detrimental lying really is depends on whether society's welfare is decreased, increased or simply redistributed through lying. An important aspect is the possibility of an inherent value of truthful revelation.

8. List of References

Battaglini, M. "Multiple Referrals and Multidimensional Cheap Talk." The Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 1295 (1999).

Battaglini, M. "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica*, vol. 70 (2002), pp. 1379-1401.

Crawford, V.P. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *The American Economic Review*, vol. 93 (2003), pp. 133-149.

Crawford, V.P., Sobel, J. "Strategic Information Transmission." *Econometrica*, vol. 50 (1982), pp. 1431-1451.

Dickhaut, J.W., McCabe, K.A., Mukherji, A. "An Experimental Study of Strategic Information Transmission." *Economic Theory*, vol. 6 (1995), pp. 389-403.

dtv-Lexikon, vol.11, Munich: DTV, 1990, p. 149.

Farell, J. "Meaning and Credibility in Cheap-Talk Games." *Games and Economic Behavior*, vol. 5 (1993), pp. 514-531.

Farell, J., Gibbons, R. "Cheap Talk with Two Audiences." *The American Economic Review*, vol. 79 (1989), pp. 1214-1223.

Farell, J., Rabin, M. "Cheap Talk." *The Journal of Economic Perspectives*, vol. 10 (1996), pp. 103-118

Goffman, E. *Strategic Interaction*. Philadelphia: University of Pennsylvania Press, 1969.

Gneezy, U. “Deception: The Role of Consequences.” *The American Economic Review*, vol. 95 (2005), pp. 384-394.

Green, J., Stokey, N. “A Two-Person Game of Information Transmission.” The Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 418 (May 2003). Internet Source: http://ideas.repec.org/p/nwu/cmsems/418.html#provider (May 2007)

Ottaviani, M., Squintani, F., “Naïve Audience and Communication Bias.” *The International Journal of Game Theory*, vol. 35 (2006), pp. 129-150.

Pitchik, C., Schotter, A. “Honesty in a Model of Information Transmission.” *The American Economic Review*, vol. 77 (1987), pp. 1032-1036. See also “Erratum,” vol. 78, (1988), p. 1164.

Preston, R., “Lying”, in Macquarrie, J., and Childress, J., eds., *A New Dictionary of Christian Ethics*. London: SCM, 1986, p. 363.

Spence, M. “Job Market Signaling.” *The Quarterly Journal of Economics*, vol. 87 (1973), pp. 355-374.

Sydsaeter, K., Hammond, P., Seierstad, A., Stroem, A. *Further Mathematics for Economic Analysis*. Harlow, Essex: Pearson Education, 2005.