CARF Working Paper

CARF-F-174

Implementation and Mind Control

Hitoshi Matsushima
The University of Tokyo

September 2009

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Citigroup, Dai-ichi Mutual Life Insurance Company, Meiji Yasuda Life Insurance Company, Nippon Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

Implementation and Mind Control[+]

Hitoshi Matsushima[*]

Faculty of Economics, University of Tokyo

First Version: December 8, 2008

This Version: July 8, 2009

[+] The title of the earlier version is “Implementation and Social Influence,” which is available as a discussion paper (Matsushima [2008]). The earlier version investigated additional issues related to incomplete information, mixed strategies, and iterative dominance, which have been omitted from the present version. The research for this paper was supported by a Grant-In-Aid for Scientific Research (KAKENHI 21330043) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese government.

[*]Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hitoshi at mark e.u-tokyo.ac.jp

Abstract

This paper incorporates social psychology into implementation theory, where an uninformed principal manipulates a dynamic decision-making process without employing any tailored contractual device. We demonstrate the principal's mind-control method through which he can effectively utilize social psychology tactics to incentivize informed agents to announce their information in keeping with his wishes. We show that with incentive compatibility, the principal can implement any alternative that he wishes as the unique Nash equilibrium outcome, even if the psychological cost of each agent from disobeying the principal's wishes is small as compared to his total material benefits.

Keywords: Implementation, Unique Nash Equilibrium, Dynamic Decision-Making, Social Psychology, Mind Control Methods, Expectation-Based Obedience, Tail-Chasing Competition.

Journal of Economic Literature Classification Numbers: C72, D78, D81, D86

1. Introduction

This paper investigates a decision problem that involves a principal's attempt to select the alternative that is most compatible with his wishes despite being unaware of which alternative would be the most desirable. Besides the principal, there are many agents who possess private signals about such possible alternatives. Hence, the principal requires these agents to disclose these private signals in the form of an announcement. In order to obtain the information from the agents, the principal has to devise various ways with which to incentivize each agent to reveal the information that would help the principal determine the desired alternative. However, in this case, it is insufficient for their obedient announcement to satisfy incentive compatibility, because there may exist other self-enforcing announcements by them that are not true to the principal's wishes, thereby preventing the principal from arriving at his desirable alternative. Hence, in addition to incentive compatibility, the principal has to utilize additional incentive devices that eliminate unwanted equilibria; in other words, the principal needs to obtain their obedient announcements as the *unique* Nash equilibrium.

The issue of uniqueness has been studied intensively in the standard theory of implementation; it was generally assumed that agents care only about their material benefits as shortcuts and enjoy full autonomy in making their announcements. Following this assumption, the authors in the literature pertaining to this field have generally confined their attention to inventing material-based contractual devices, such as the modulo mechanisms (Maskin [1977/1990]); and the Abreu-Matsushima mechanisms (Abreu and Matsushima [1992]) that implement, at least in the virtual sense, any value of the fixed social choice function as the unique Nash equilibrium outcome in compensation for artificial tailoring.[1]

In contrast to this standard theory, any real person cares about not only material benefit but also any psychological factor of social influence; the person experiences feelings of guilt for disobeying the authority's wishes, and this feeling is intensified when he expects his reference group to obey this authority's wishes. In this regard, several studies in social psychology such as Ash (1955), Milgram (1974), Zimbardo et al. (1977), and Hofling et al. (1966) have commonly reported that subjects in laboratories and fields tended to be obedient to the authorities[2] and tended to seek conformity to their reference groups' behavioral modes.[3]

On the basis of the above arguments, we demonstrate a new concept for the implementation of the principal's desirable alternative as follows. From the vast store of knowledge pertaining to social psychology and daily life, it is natural to infer that the aforementioned principal pragmatically considers how to utilize social psychology tactics to influence the agents with respect to their choice of announcements.[4] With the continuous time horizon, given the incentive-compatible decision function, the principal will manipulate the decision-making process in the following ways.

---

[1] For the surveys on the standard theory of implementation, see Moore (1992), Palfrey (1992), Osborne and Rubinstein (1994, Chapter 10), and Maskin and Sjöström (2002).

[2] Many of these experiments reported that the subjects are obedient to the authority, even if the authority attempts to disturb social order.

[3] For issues on social influence in general, see Cialdini (2001).

[4] Attempts to incorporate social psychology into economics are not new but are attracting growing interest. See, for instance, Akerlof and Dickens (1982), Geanakoplos, Pearce, and Stacchetti (1989), Bernheim (1994), Gneezy (2005), Charness and Dufwenberg (2006), and Bébabou (2007).

(i) The agents are required to make their announcements at the initial time. Each agent can make a different announcement at any later time and as many times as he wants.

(ii) This process is randomly terminated at a constant hazard rate. According to the decision function, the principal selects the alternative that corresponds to the announcements that are effective at the terminal time.

(iii) During this process, the agents are prohibited from communicating with or monitoring each other.

In addition to this process manipulation, we shall take into account a concept of social psychology that we refer to as *expectation-based obedience*. Expectation-based obedience implies that the degree to which each agent experiences feelings of guilt with regard to disobeying the principal's wishes depends to a great extent on his expectations about the other agents' behavioral modes; in other words, as a rule, an agent will experience greater feelings of guilt about disobeying the principal's wishes if he expects that no agent has been disobedient before. Thus, if he expects that someone has already disobeyed the principal's wishes before, he does not necessarily experience guilt.

Since the process manipulation prohibits monitoring and communication, there is no room for an agent to influence the other agents by being disobedient himself; there is no means for him to incite disobedience and eliminate feelings of guilt experienced by members of his reference group. In other words, manipulating the dynamic decision-making process in the above manner is assumed to allow the principal to successfully defend himself against any possibility of civil disobedience from the agents, a real-life example of which would be the Montgomery Bus Boycott by Rosa Parks that eventually led to the modern civil rights movement in the United States.

The result of this paper is quite permissive from the principal's viewpoint; even if an agent's psychological cost of disobeying is small as compared to his total material benefits, the principal can incentivize the agents to make announcements obediently at all times, that is, he can implement any alternative that he wishes as the unique Nash equilibrium outcome.

Since the decision-making process is randomly terminated in a continuous time horizon, any point-wise change of announcement hardly influences the alternative choice. Moreover, according to expectation-based obedience, each agent can slightly reduce his psychological cost by waiting for someone else to disobey. This tiny psychological cost reduction, along with random termination, is sufficient to trigger a *tail-chasing competition* among the agents, eliminating all possibilities of them beginning to disobey the principal's wishes in due order.

Our tail-chasing competition stems from the basic concept of Abreu-Matsushima mechanisms (Abreu and Matsushima [1992]) explored in the standard theory of implementation; the standard theory was generally devoted to inventing material-based contractual devices. In contrast, the present paper shows that if there is a little room for the principal to infringe on the agents' autonomy, the principal can apply the same logic as the standard theory to the invention of *mind-control methods*, rather than contractual devices.

The earlier works of Matsushima (2008a, 2008b) took into account the psychological aspects in the implementation literature.[5] These works, however, did not introduce expectation-based obedience, and therefore, still needed to consider tailored contractual devices à la Abreu-Matsushima to incentivize the agents.

This paper is organized as follows. Section 2 illustrates the model. Section 3 introduces expectation-based obedience. Section 4 presents the main theorem that represents the unique Nash equilibrium wherein agents make announcements obediently at all times.

[5] There are a few other works in the implementation theory that are relevant to psychological aspects, such as Eliaz (2002) and Glazer and Rubinstein (1998).

2. The Model

Let $N \equiv \{1, 2, ..., n\}$ denote the set of *agents*, where $n \geq 2$. Let A denote the set of *alternatives*. Let us consider a decision problem with the continuous time horizon $[0, \infty)$, in which, a *principal* makes an alternative choice according to the following process given by $\Gamma \equiv (M, g, r)$. Let M_i denote the set of *messages* for each agent $i \in N$. Let $M \equiv \underset{i \in N}{\times} M_i$ denote the set of message profiles. At the initial time 0, the principal requires each agent i to announce any message, $m_i \in M_i$. Further, at any time after the initial time 0, an agent can change his message as frequently and whenever he wants.

We assume that at any time, any agent cannot monitor the other agents' announcements, and therefore, changes in choice by an agent are not contingent on the other agents' past announcements. This assumption is crucial in the present study since it takes away any means of civil disobedience available to the agents. On the basis of this assumption, we define a *strategy* for agent i as a function $s_i : [0, \infty) \to M_i$, where $s_i(t) \in M_i$ denotes the message that agent i announces at time t, that is, the message that stands at time t when the process terminates. We assume that s_i is right-continuous, that is, for every $t > 0$, either

$$s_i(\tilde{t}) = s_i(t) \quad \text{for all} \quad \tilde{t} \geq t,$$

or there exists $t' > t$ such that

$$s_i(t') \neq s_i(t), \text{ and } \; s_i(\tilde{t}) = s_i(t) \quad \text{for all} \quad \tilde{t} \in [t, t').$$

Let S_i denote the set of strategies for agent i. Let $S \equiv \times_{i \in N} S_i$ denote the set of strategy profiles. Let $S_{-i} \equiv \times_{j \in N \setminus \{i\}} S_j$ for each $i \in N$.

The principal randomly terminates the dynamic decision-making process at a constant *hazard rate* $r \in (0, \infty)$. For every $t \in [0, \infty)$, the probability that this process terminates at or after any time t is given by

$$\exp(-rt).$$

When the process terminates at time t, the principal makes an alternative choice on the basis of the message profile $s(t) = (s_i(t))_{i \in N} \in M$ that has been announced at terminal time t; he selects the alternative $g(s(t)) \in A$ according to the *decision function* given by $g : M \to A$, along with the message profile $s(t)$ announced at the terminal time t.

An additional account for this process is given as follows. Before the initial time 0, the principal explains his *wishes* to each agent, in words such as "I wish to aid the poorest persons."[6] The principal then requests each agent to provide a message containing any relevant information that is unknown to the principal, asking questions such as "Where do the poorest persons live?" Given that the agents have announced a message profile $m \in M$ at the randomly determined terminal time, the principal will regard the corresponding alternative $g(m) \in A$ to be the desirable one in light of his wishes.

For each $i \in N$, let us set a message $m_i^* \in M_i$ as the *truthful message* for agent i, which indicates the obediently announced message by agent i, i.e., the best answer by

[6] The arguments in this paper are irrespective of whether the principal's wishes are prosocial, antisocial, or neither.

agent i in line with the principal's wishes. Let $m^* = (m_i^*)_{i \in N} \in M$ denote the *truthful message profile*. We define the *truthful strategy* $s_i^* \in S_i$ for agent i by

$$s_i^*(t) = m_i^* \text{ for all } t \geq 0.$$

According to s_i^*, agent i obediently announces information at all times. Let $s^* = (s_i^*)_{i \in N} \in S$ denote the *truthful strategy profile*.

Let us denote the *payoff* function for agent i by $U_i : S \to R$, where $U_i(s)$ implies the payoff for agent i when he follows the strategy $s_i \in S_i$ and he expects the other agents to follow the profile of strategies $s_{-i} \in S_{-i}$. We define a *game* as a combination of the dynamic decision-making process and the profile of the payoff functions, given by $(\Gamma, (U_i)_{i \in N})$. A strategy profile $s \in S$ is said to be a *Nash equilibrium* in the game of $(\Gamma, (U_i)_{i \in N})$ if for every $i \in N$,

$$U_i(s) \geq U_i(s_i', s_{-i}) \text{ for all } s_i' \in S_i.$$

We assume that the payoff $U_i(s)$ for agent i is separated into two parts:

$$U_i(s) = V_i(s) - W_i(s).$$

The first part $V_i(s)$ is called the *material benefit*, and the second part $W_i(s)$ the *psychological cost*. The material benefit $V_i(s)$ implies the expected value of the intrinsic utility given by $v_i(a) \in R$, which is derived directly from the alternative choice, that is,

$$V_i(s) \equiv \int_{t=0}^{\infty} v_i(g(s(t)))d[1 - \exp(-rt)].$$

Let us assume *incentive compatibility* in terms of the intrinsic utility in that each agent can maximize his intrinsic utility by obediently making announcements, provided he expects the other agents to do the same; hence, for every $i \in N$,

(1) $$v_i(g(m^*)) \geq v_i(g(m^*/m_i)) \quad \text{for all} \quad m_i \in M_i .$$

From this assumption, it is clear that for every $i \in N$,

(2) $$V_i(s^*) \geq V_i(s^*/s_i) \quad \text{for all} \quad s_i \in S_i ,$$

which implies that each agent can maximize his material benefit by obediently announcing information at all times, provided he expects the other agents to do so at all times, too.

Any agent $i \in N$ cares more or less about any psychological factor of social influences that determine his psychological cost $W_i(s)$. We assume that any agent feels guilty about being disobedient at any time, if he expects the other agents to obediently announce information at all times; thus, for every $i \in N$,

(3) $$W_i(s^*) < W_i(s^*/s_i) \quad \text{for all} \quad s_i \in S_i \setminus \{s_i^*\} .$$

It is implicit in this assumption that the degree to which any agent i can reduce his psychological cost is very limited.

3. Expectation-Based Obedience

Let us introduce another assumption on $W_i(s)$ that we refer to as *expectation-based obedience*. For every $s_i \in S_i / \{s_i^*\}$, we define $t_i(s_i) \in [0,\infty)$ by

$$s_i(t_i(s_i)) \neq m_i^*, \text{ and } s_i(\tilde{t}) = m_i^* \text{ for all } \tilde{t} < t_i(s_i),$$

which indicates the first time that agent i makes a disobedient announcement. Let us denote $t_i(s_i^*) = \infty$. For every $t > 0$ and every strategy $s_i \in S_i / \{s_i^*\}$ for agent i, we define another strategy $s_{i,t} \in S_i$ for agent i as follows:

$$s_{i,t}(\tilde{t}) = m_i^* \text{ for all } \tilde{t} \in [0,t), \text{ and } s_{i,t}(\tilde{t}) = s_i(\tilde{t}) \text{ for all } \tilde{t} \geq t.$$

According to $s_{i,t}$, agent i continues to obediently announce information until time t, while he follows s_i at or after time t. Let us define

$$L_i \equiv \max_{(a,a')\in A^2} |v_i(a) - v_i(a')|,$$

which implies the upper bound of differences in intrinsic utility for agent i.

Expectation-Based Obedience: For every $i \in N$, every $j \in N / \{i\}$, and every $s \in S$, if

$$t_i(s_i) \leq t_j(s_j) \leq t_h(s_h) \text{ for all } h \in N / \{i\},$$

then

(4) $$\lim_{\varepsilon\downarrow 0}\frac{W_i(s)-W_i(s_{i,t_j(s_j)+\varepsilon},s_{-i})}{\varepsilon} > L_i r\exp(-rt)\,.$$

The degree to which each agent feels guilty about disobeying the principal's wishes depends on his expectation about the other agent's behavioral modes; his feelings of guilt about being disobedient would increase if he expects no agent to have been disobedient in the past, rather than otherwise. In other words, if he expects someone to have already been disobedient, he does not necessarily experience guilt. Hence, he can relieve his feelings of guilt by postponing his first act of disobedience to after another agent is disobedient, that is, by avoiding having to be the first person to behave disobediently.

To be more precise about expectation-based obedience, let us suppose that agent i is the first person to be disobedient, whereas agent $j\in N/\{i\}$ is the first person except for agent i to be disobedient. Then, by postponing his first act of disobedience from time $t_i(s_i)$ to time $t_j(s_j)+\varepsilon$, agent i can avoid being the first person to behave disobediently; thus, he can save his psychological cost by a positive amount of $W_i(s)-W_i(s_{i,t_j(s_j)+\varepsilon},s_{-i})>0$. Given that ε is positive but close to zero, the inequality (4) in expectation-based obedience implies that the reduction in costs that happen in this manner is greater than

$$\varepsilon L_i r\exp(-rt_j(s_j))\,.$$

Expectation-based obedience carries an implicit assumption that the degree to which each agent i can save his psychological cost is very limited.

Example: Let us denote the psychological cost of each agent i by

$$W_i(s) \equiv \int_{t=0}^{\infty} w_i(s;t)d[1-\exp(-rt)]\,.$$

In this equation, we defined $w_i(s,t)$ as

$$w_i(s,t) = \lambda \frac{\int_{\tau=0}^{t} \iota_i(s_i(\tau))dt}{t} + \eta \qquad \text{if } t_i(s_i) \le t \text{ and } t_i(s_i) \le t_h(s_h) \text{ for all } h \in N\,,$$

and

$$w_i(s,t) = \lambda \frac{\int_{\tau=0}^{t} \iota_i(s_i(\tau))dt}{t} \qquad \text{otherwise,}$$

where $\lambda > 0$, $\eta > 0$, and the function $\iota_i : M_i \to \{0,1\}$ is defined by

$$\iota_i(m_i^*) = 0\text{, and } \iota_i(m_i) = 1 \text{ for all } m_i \in M_i \setminus \{m_i^*\}\,.$$

Note that $\frac{\int_{\tau=0}^{t} \iota_i(s_i(\tau))dt}{t}$ implies the proportion of the time that agent i is disobedient when the decision-making process terminates at time t. Hence, the greater this proportion is, the greater is his psychological cost.

More importantly, $\eta > 0$ implies the additional increase in the psychological cost when the agent becomes the first person to be disobedient in his expectation. Since the presence of this positive value $\eta > 0$ renders the left-hand side of (4) equal to infinity, our example automatically satisfies the expectation-based obedience irrespective of the specification of the hazard rate r.

Our example also satisfies

$$\max_{(s,s')\in S^2} \left| W_i(s) - W(s') \right| \leq \lambda + \eta \; .$$

Hence, by letting $\lambda > 0$ and $\eta > 0$ close to zero, we can make the differences in psychological cost as close to zero as possible. This implies that each agent's psychological cost of disobeying the principal could be negligible as compared to his total material benefits.

4. The Theorem

We demonstrate the main theorem of this paper, according to which *the principal can implement any alternative that he wishes as the unique Nash equilibrium outcome*.

The Theorem: *The truthful strategy profile* s^* *is the unique Nash equilibrium in the game of* $(\Gamma,(U_i)_{i\in N})$.

Proof: It is clear from (2) and (3) that s^* is a Nash equilibrium; for every $i \in N$ and every $s_i \in S_i / \{s_i^*\}$:

$$V_i(s^*) - W_i(s^*) > V_i(s^* / s_i) - W_i(s^* / s_i), \text{ i.e., } U_i(s^*) > U_i(s^* / s_i).$$

Let us consider any other strategy profile $s \in S / \{s^*\}$, where there exist $i \in N$ and $j \in N / \{i\}$ such that

$$t_i(s_i) < \infty, \text{ and } t_i(s_i) \le t_j(s_j) \le t_h(s_h) \text{ for all } h \in N / \{i\}.$$

Let us choose $\varepsilon > 0$ close to zero. From (1), along with the definition of L_i, it follows that

$$V_i(s) - V_i(s / s_{i,t_j(s_j)+\varepsilon})$$

$$= \int_{t=t_i(s_i)}^{t_j(s_j)} \{v_i(g(s(t))) - v_i(g(s(t) / s_{i,t_j(s_j)+\varepsilon}(t)))\} d[1-\exp(-rt)]$$

$$+\int_{t=t_j(s_j)}^{t_j(s_j)+\varepsilon}\{v_i(g(s(t)))-v_i(g(s(t)/s_{i,t_j(s_j)+\varepsilon}(t)))\}d[1-\exp(-rt)]$$

$$=\int_{t=t_i(s_i)}^{t_j(s_j)}\{v_i(g(m^*/s_i(t)))-v_i(g(m^*))\}d[1-\exp(-rt)]$$

$$+\int_{t=t_j(s_j)}^{t_j(s_j)+\varepsilon}\{v_i(g(s(t)))-v_i(g(s(t)/m_i^*))\}d[1-\exp(-rt)]$$

$$\leq L_i[\exp\{-rt_j(s_j)\}-\exp(-r\{t_j(s_j)+\varepsilon\})],$$

which is approximated by

$$\varepsilon L_i r\exp(-rt_j(s_j)).$$

From (4), along with the sufficiently small $\varepsilon>0$, it follows that

$$W_i(s)-W_i(s/s_{i,t_j(s_j)+\varepsilon})>\varepsilon L_i r\exp(-rt_j(s_j)).$$

From these observations, we have shown that

$$U_i(s)-U_i(s/s_{i,t_j(s_j)+\varepsilon})$$

$$=V_i(s)-V_i(s/s_{i,t_j(s_j)+\varepsilon})-\{W_i(s)-W_i(s/s_{i,t_j(s_j)+\varepsilon})\}$$

$$<\varepsilon L_i r\exp(-rt_j(s_j))-\varepsilon L_i r\exp(-rt_j(s_j))=0,$$

which implies that s is not a Nash equilibrium.

Q.E.D.

In order to reduce his psychological cost, each agent may prefer postponing his first act of disobedience until after any other agent is disobedient. However, a difficulty is presented when postponing his disobedient announcement, because in this manner, he is caught between the reduction in psychological cost and the loss in material benefits, which are commonly caused by message changes from a disobedient one to an obedient one.

Expectation-based obedience can overcome this difficulty as follows. Suppose that any agent can avoid being the first person to be disobedient by postponing his disobedient announcement for a short interval. Expectation-based obedience implies that the probability that the decision-making process terminates during this interval is kept low enough to render the expected value of loss in the intrinsic utility less than the reduction in the psychological cost. Hence, this low probability can trigger a tail-chasing competition among the agents that perpetually edges their first acts of disobedience upward, thereby eliminating unwanted equilibria.

With respect to the functioning of tail-chasing competition, our model is closely related to the basic concept of the Abreu-Matsushima mechanism (Abreu and Matsushima [1992]). In the mechanism, each agent announces multiple messages[7] and is motivated to avoid being the first person who makes an announcement that is inconsistent with the first messages, triggering a tail-chasing competition among the agents.

There are substantive points of difference between our model and the Abreu-Matsushima mechanism; in order to trigger the tail-chasing competition, the Abreu-Matsushima mechanism uses any contractual device of side payments (or similar to this), stipulating that any agent is fined by a small amount of money if and only if he

[7] In the Abreu-Matsushima mechanism, the number of messages that each agent actually announces is fixed; this number is not fixed in our model since the decision-making process randomly terminates.

is the first person to make an announcement that is inconsistent with the first messages. The Abreu-Matsushima mechanism requires more complicated contractual devices[8] that incentivize the agents to make their first announcements truthful. Our model, on the other hand, does not use any such contractual device at all.

Throughout this paper, it was assumed that each agent cannot monitor the other agents' announcements until the process terminates. If we permit each agent to monitor them, we need to investigate a version of the repeated games and struggle with the multiplicity of equilibria implied by the folk theorem or some similar principle. By allowing monitoring in this manner, any agent may have an incentive to take the initiative to make a disobedient announcement at an early stage in order to free the other agents from the authorities' spell. I believe that this point would be substantial with respect to the issue of implementation and mind control, but is beyond the scope of the present paper.

[8] Matsushima (2008a, 2008b) introduced psychological costs into the Abreu-Matsushima mechanism in order to avoid this complexity.

References

Abreu, D. and H. Matsushima (1992): "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica* 60, 993–1008.

Akerlof, G. and W. Dickens (1982): "The Economic Consequences of Cognitive Dissonance," *American Economic Review* 72, 307–319.

Ash, S. (1955): Opinion and Social Pressure, *Scientific American* 193, 31–35.

Bébabou, R. (2007): "Groupthink: Collective Delusions in Organizations and Markets," mimeo.

Bernheim, D (1994): A Theory of Conformity, *Journal of Political Economy* 102, 841–877.

Cialdini, R. (2001): *Influence: Science and Practice*. Boston: Allyn & Bacon.

Charness, G. and M. Dufwenberg (2006): "Promises and Partnerships," *Econometrica* 74, 1579–1601.

Eliaz, K. (2002): "Fault Tolerant Implementation," *Review of Economic Studies* 69, 589–610.

Hofling, C., E. Brontzman, S. Dalrymple, N. Graves, and C. Piece (1966): "An Experimental Study of Nurse-Physician Relationships," *Journal of Nervous and Mental Disease* 143, 171–180.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games and Economic Behavior* 1, 60–79.

Glazer, J. and A. Rubinstein (1998): "Motives and Implementation: On the Design of

Mechanisms to Elicit Opinions," *Journal of Economic Theory* 79, 157–173.

Gneezy, U. (2005): "Deception: The Role of Consequences," *American Economic Review* 95, 384–394.

Maskin, E. (1999): "Nash equilibrium and Welfare Optimality," *Review of Economic Studies* 66, 23–38.

Maskin, E. and T. Sjöström (2002): "Implementation Theory," in *Handbook of Social Choice and Welfare Volume 1*, ed. by K. Arrow, A. Sen, and K. Suzumura. Amsterdam: Elsevier.

Matsushima, H. (2008a): "Role of Honesty in Full Implementation," *Journal of Economic Theory* 127, 353–359.

Matsushima, H. (2008b): "Behavioral Aspects of Implementation Theory," *Economics Letters* 100, 161–164.

Milgram, S. (1974): *Obedience to Authority: An Experimental View*. New York: Harper and Row.

Moore, J. (1992): "Implementation in Environments with Complete Information," in *Advances in Economic Theory: Sixth World Congress*, ed. by J. J. Laffont. Cambridge: Cambridge University Press.

Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*. Cambridge, MA: MIT Press.

Palfrey, T. (1992): "Implementation in Bayesian Equilibrium: The Multiple Equilibrium Problem in Mechanism Design," in *Advances in Economic Theory: Sixth World Congress*, ed. by J. J. Laffont. Cambridge: Cambridge University Press.

Zimbardo, P., C. Haney, W. Banks, and D. Jaffe (1977): The Psychology of

Imprisonment: Privation, Power, and Pathology, in *Contemporary Issues in Social Psychology*, ed. by J. Brigham and L. Wrightman. Cole Publishing Co.