早稲田大学現代政治経済研究所

A MECHANISM THAT OVERCOMES COORDINATION FAILURE BASED ON GRADUALISM, ENDOGENEITY, AND MODIFICATION

Yoshio Kamijo,School of Management, Kochi University of Technology
Hiroki Ozono, Faculty of Law, Economic and Humanities, Kagoshima University
Kazumi Shimizu, School of Political Science and Economics, Waseda University

No.E1401

Working Paper Series

Institute for Research in
Contemporary Political and Economic Affairs

Waseda University
169-8050 Tokyo,Japan

A MECHANISM THAT OVERCOMES COORDINATION FAILURE BASED ON GRADUALISM, ENDOGENEITY, AND MODIFICATION

Yoshio Kamijo[1], School of Management, Kochi University of Technology

Hiroki Ozono, Faculty of Law, Economic and Humanities, Kagoshima University

Kazumi Shimizu, School of Political Science and Economics, Waseda University

Abstract

We examine three tools that can enhance coordination success in a repeated multiple choice coordination game. Gradualism means that the game starts as an easy coordination problem and moves gradually to a more difficult one. Endogeneity implies that a gradual increase in the upper limit of coordination occurs only if coordination with the Pareto superior equilibrium in a stage game is attained. Modification requires that when they fail coordination, the level of the next coordination game is adjusted to an easier one. We find from laboratory experiment that a mechanism that combines these three, termed herein the GEM, works well.

JEL codes: C72, C91, C92, M54

Keyword: Coordination Failure, Minimum Effort Game, Experiment, Gradualism, Endogeneity, Modification

1. Introduction

The difficulty of coordinating actions among several persons is well described by Jean-Jacques Rousseau's allegoric story about stag hunting.[2] However, the seriousness of the coordination problem has been clarified through the findings of several laboratory experiments that show, in certain situations, that coordinating actions with a Pareto superior outcome remain difficult (Van Huyck, Battalio and Beil 1990, 1991; Cooper et al. 1989, 1990).

1 Corresponding Author. Email: yoshio.kamijo@gmail.com

2 Jean-Jacques Rousseau [1775]., *Discours sur l'origine et les fondemens de l'inégalité parmi les homes*, Amsterdam: Marc Michel Rey, 1755.

Coordination failure is also found in our daily decision making (e.g. chicken games, battle of sexes), in workplace decisions (Knez and Camerer 2006), in a firm's decision to seek a trading partner (Diamond 1982), in demand decisions in a multi-sector economy (Cooper and John 1988), and in country-level decisions such as reducing CO_2 and sanctioning illegal countries. Nevertheless, coordination success has enabled human society to develop profoundly, with major examples including the Glorious revolution (Weingast 1997) and the economic prosperity of medieval Genoese society (Greif 1993).

An early laboratory-based study of coordination problems focused on why coordination with a Nash equilibrium or a Pareto efficient equilibrium is rarely observed and examined why subjects choose one action from several salient ones (Pareto dominant, secure, etc.) and adjust it in repeated plays（Van Huyck, Battalio and Beil 1990, 1991）. This study offered a fresh perspective on how we can overcome coordination failure in a laboratory setting that might be beneficial to practical problems.[3] Recent studies tend to weigh a more attention to explore a direct or indirect way to overcome coordination failure. For example, Brands and Cooper (2006a, 2006b) examine whether a financial incentive that prompts an efficient action is effective in a laboratory setting and find that this incentive drastically shifts the choices of subjects to an efficient equilibrium. A similar observation is obtained from other studies like Guillen, Schwieren and Staffiero (2006) and Hamman, Rick and Weber (2007). In addition, this finding has been empirically demonstrated in the airport industry (Knez and Simester 2001).

Although the provision of a financial incentive seems to succeed, this is only one approach to overcoming coordination failure. Moreover, there is a negative aspect to adopting a financial incentive. It has often been emphasized from social psychology and recent experimental economics that an external incentive reduces the intrinsic motivation to perform altruistic or cooperative actions, thereby reducing overall cooperation level (Frey and Oberholzer-Gee

[3] For a survey of coordination games, see Devetag and Oltmann (2007).

1997; Frey and Jegen 2001; Lepper, Greene and Nisbet 1973). Brands and Cooper (2006a, 2006b) and Hamman, Rick, and Weber (2007), for example, showed experimentally that while coordination level increases after the introduction of a financial incentive and subjects achieve a more efficient equilibrium, this does not continue following its removal.

Based on the foregoing, in the present paper we investigate the driving forces behind coordination success outside of establishing an external incentive scheme. In this regard, previous research has found that communication among players (Cooper et al. 1989, 1992), the observability of others' choices (Brands and Cooper 2006b), leadership (Cooper 2006; Brandts, Cooper and Fatas 2007, 2011), and behavioral spillover from a similar problem (Devetag 2005; Weber 2006; Cason, Savikhin and Sheremeta 2012) are effective at reducing coordination failure. In this paper, we thus propose three tools to coordination success that can be effective in a highly anonymous environment (i.e., an environment without communication, without observing others, and without leadership). These three tools are (i) the gradual increase in the difficulty of the coordination problem (gradualism hereafter), (ii) the endogenous change in difficulty (endogeneity), and (iii) modification in the case of coordination failure (modification). Overall, we term a mechanism that combines these three tools "the GEM mechanism."

We examine these three tools through a repeated minimum effort game. A minimum effort game, sometimes called a weakest link game, is a coordination game that is difficult for subjects to coordinate to a Pareto efficient equilibrium. We choose a multiple-choice version of the minimum effort game for our study setting, not only because this is comparable with previous coordination problem experiments, but also because a multiple-choice version has higher external validity. In our game, each player chooses his or her level of effort from a limited set of numbers, say 0, 10, …, 60. All players that choose the same level of effort reside at the Nash equilibrium, but the Nash equilibria are Pareto ranked in the sense that 60 constitutes the "best" equilibrium and 0 the "worst" one. However, choosing 60 is risky because the payoff of the player that chooses 60 largely depends on others' choices. On the contrary, choosing 0 guarantees a positive payoff and this is therefore a secure choice. Given this context, coordinating for the best equilibrium is difficult without external assistance, which is where the proposed GEM mechanism is beneficial.

The first tool of the GEM mechanism is gradualism. This tool means that starting from an easy coordination problem and moving gradually to a more difficult but profitable one can eventually achieve coordination at the Pareto efficient equilibrium. Gradualism is based on the idea that people's success experience under a low demand condition may lead them to try a marginally more difficult coordination game. As a result, coordination level can be expected to increase under the "snowball effect."

The remaining question in the "success produces success" notion of gradualism is how we succeed in the first coordination and what happens if the coordination to the efficient equilibrium fails in an easy problem. This point is explained by the second tool, endogeneity, which can be encapsulated by the following phrase "If you succeed in this problem, the next one becomes harder but more profitable." Since endogenous change correlates the current choice with the future (better) situation, it enhances cooperation in the first period. In addition, if players fail to coordinate to an efficient equilibrium, they can try the same problem again and thus have a chance to coordinate on the easy problem once more.

The third tool is modification. In contrast to the second tool, the third modifies the level to an easier problem that had previously been overcome successfully. Modification not only allows subjects to reconstruct mutual trust after coordination failure but also prevents players from choosing inefficient actions through the threat of losing the future benefit. This tool has two distinct effects on coordination success in the long run; one is recovery after coordination failure and the other is the maintenance of efficient coordination.

By using a laboratory experiment, we find that, as expected, the GEM mechanism enhances coordination to a Pareto efficient equilibrium. Indeed, we show that most of the study group achieve and keep a Pareto efficient equilibrium in the most difficult and profitable minimum effort game. However, gradualism alone does not work well. Thus, our findings demonstrate that the combination of the three tools (i.e., the GEM mechanism) is important to achieving coordination success.

The remainder of this paper is organized as follows. In the next section, we review the aspects of gradualism, endogeneity, and modification that have been observed and discussed in studies in fields such as psychology, evolutionary biology, management, and experimental economics. In Section 3, we explain our basic model of a coordination game and the GEM mechanism. In Section 4, we describe the experimental design, and propose our hypotheses in Section 5. Section 6 reports the results of the experiment, while a discussion of these results and concluding remarks are presented in Section 7.

2. The literature on gradualism, endogeneity, and modification

2.1 Gradualism

The idea of gradualism in a coordination problem goes back to the finding of Van Huyck, Battalio and Beil (1990, 1991), who showed that subject behavior is heavily affected by past results (i.e., there is path-dependency in their behaviors). In fact, when a group of subjects achieves a high coordination level at the first time of asking, they tend to maintain or even increase it in a repeated play. By contrast, coordination failure tends to continue if the subjects fail to coordinate at an early stage.

In addition, Devetag (2005) and Cason, Savikhin, and Sheremeta (2012) report that a good precedent established in one game spills over to another albeit different game, implying that previous coordination success can be extended to a more complicated situation. Further, Weber (2006) suggests that coordination success in a large group, which is often observed in the real world, can be attained from the gradual change of the population from a small group to a larger one. In the same vein, Ye et al. (2011) consider gradual change in the difficulty and profitability of a coordination game and show that coordination success in the gradual change condition is greater compared with sharp change or the constant condition.

Kurzban, Rigdon, and Wilson (2008) use a repeated trust game in order to investigate whether the sender's strategy of gradually increasing his or her investment is effective and their results weakly confirm the use of this strategy. Further, Roberts and Renwick (2003) find that in a repeated prisoners' dilemma game, subjects often choose a low level of cooperation at the beginning and then gradually increase cooperation level. In both a trust game and a prisoner's dilemma game, gradualism thus serves to avoid taking a "leap of faith" and begins by "testing the water" (Roberts and Sherratt 1998), namely identifying the partner's type by testing if she/he cooperates for small stakes and mitigating the risk of trusting unconditionally the partner. As players come to trust others, cooperation level increases.

There exist various examples of practical situations where gradualism works well. For example, the enlargement of the EU over time has been a gradual integration process. As for tax reforms, although it is left unanswered whether economic policies work best when implemented quickly or gradually, there are significant advantages for a step-by-step procedure. In both cases, gradualism allows people to cope with problems by maintaining an appropriate balance between evolving environments and people's degree of adaptation.

2.2 Endogeneity

The effect of endogenous change has been investigated from early works in the fields of psychology and management to recent experimental economics studies. This body of research suggests that endogeneity has distinct effects on behavior—both before and after endogenous change. The "before" effect relates to motivation theory. When improvement occurs from their efforts, people recognize the positive correlation between their behavior and the outcome. According to expectancy theory (Vroom 1964), this correlation enhances the motivation of workers towards goal attainment and contributes to higher cooperation in the workplace. If we transfer this to the context of coordination problems, endogenous change can lead to belief sharing among subjects working towards goal achievement and thus facilitate a Pareto efficient coordination.

The "after" effect occurs for psychological reasons. The pioneering literature on procedural justice (Thibaut and Walker 1975) links people's concern about procedures to their motivation to affect their results, and thus defines procedural justice as the level of input or participation that procedures allow. This suggests that people are more likely to feel fairly treated when they participate in a decision-making process actively compared with when the outcome is given passively. A higher degree of fairness perception thus influences people's behavior and often leads to greater cooperation.

In addition, self-determination theory proposed by Ryan and Deci (2000) suggests that when people experience autonomous motivation, they adhere to their own choices. Although both procedural justice and self-determination theory do not necessarily predict the exact effect of the endogenous change on subject behavior, both suggest that people's behavior is completely different between an enforced situation and an endogenously self-determined situation.

A similar finding concerning the "after" effect of endogeneity is reported from experimental and empirical economics. By using a laboratory public goods game, Tyran and Feld (2006) show that exogenously imposing mild law does not achieve compliance but that if mild law is endogenously chosen, compliance is much improved. Dal Bo, Foster, and Putterman (2008), Sutter, Haigner, and Kocher (2010), and Markussen, Putterman, and Tyran (2010) also report experimental results consistent with the statement that endogenously chosen institutions show higher cooperation than exogenously enforced institutions even when these two institutions are the same. In addition, studies have found empirical evidence that democratic rights positively influence the satisfaction level of citizens (Frey and Stutzer 2005) and compliance to tax payments (Pommerehne and Weck-Hannemann 1996). These findings imply that an endogenous procedure pushes people's behavior in a socially desirable direction.

2.3 Modification

Modification can be defined as worldly wisdom in our daily lives. When we cannot achieve a task, we often reduce its complexity and restart with an easier one. Modifications include decreasing the number of steps or simplifying how the results are to be obtained. Along with our daily intuition, a number of studies in pedagogy and sports science have shown that modifications can improve people's intelligence as well as their physical capacity (Kuiper and Keizer, 1988; Houssart 2002; Sullivan, Mousley and Zevenbergen 2006).

3. Minimum effort coordination game and a GEM mechanism

3.1 Minimum effort coordination game

Let us introduce a four-person minimum effort coordination game. In this game, each player $i \in \{1,2,3,4\}$ simultaneously chooses his or her effort level $e_i \in \{0,10,20,\ldots,60\}$, which will be extracted from his or her endowment of 60 points, and each player obtains a minimum of four players' efforts multiplied by 2. Thus, i's payoff is

$$\pi_i = 60 - e_i + 2 \times \min_j e_j.$$

The payoff table of a player is summarized in Table I and our parameter selection follows Van Huyck, Battalio, and Beil (1990) with minor modifications in the range of effort levels available.

[Place Table I Here]

It is clear that the action profile (x, x, x, x) is an equilibrium for any selection of x and that $x = 60$ constitutes a Pareto efficient equilibrium. However, choosing 60 is risky because, for instance, if the other three players choose their efforts at random, the probability of their minimum being 0 is $1-(6/7)^3 \cong 0.370$ and the probability of their minimum being 60 is $(1/7)^3 \cong 0.003$, implying that a player that chooses 60 has a very low probability of obtaining 120 but about 40% probability of obtaining nothing. Therefore, even though choosing 60 constitutes a Pareto efficient equilibrium, this is not the sensible option without a communication or external enforcement system. Indeed, a number of experimental studies have found that in such a minimum effort coordination game, coordination to the Pareto efficient equilibrium is difficult when the number of group members is above three—even in a repeated game with fixed members (Engelmann and Normann 2010).

3.2 A GEM mechanism for a repeated minimum effort coordination game

We propose three tools to achieve coordination success in the minimum effort game. Our approach is based on existing findings that highlight that a good precedent for a slightly different situation still applies to a new coordination game, which allows subjects to coordinate to an efficient equilibrium (Devetag 2005; Weber 2006; Cason, Savikhin and Sheremeta 2012). In our experiment, the upper limit of effort level is set at 10 points in the first period and this gradually increases by period. Thus, the first period coordination game is a binary choice game between 0 and 10 (see Table I), the second period is one among 0, 10, and 20, the third period is one among 0, 10, 20, and 30, and so on. Thus, the difficulty and profitability of the game gradually increase.

Let $\mathrm{m}^{\mathrm{t}}(=0,10,20,\ldots,60)$ be the upper limit of effort level in period *t*. We also assume $\mathrm{m}^1=10$. Under an **exogenous gradualism** condition, the upper limit of the stage game is determined as follows:

$$m^t=\begin{cases}10\,t, t<6\\ 60, t\geq 6\end{cases}$$

It should be noted that a gradual increase in the bound occurs irrespective of the results of the coordination game in the previous period.

Why is gradual change expected to enhance coordination success in a multiple choice coordination game? Previous studies such as Van Huyck, Battalio, and Beil (1990, 1991) have argued that the difficulty in coordination arises from opposing two salient actions (and intermediate actions between the two); one is a secure but inefficient action "0" and the other is a Pareto efficient but risky action "60." Gradual change mitigates this difficulty through the success experience gained in the previous period. For example, subjects begin with a binary choice between 0 and 10 and achieve coordination success to 10 in the first period. Based on this previous success, subjects in the second period may consider the multiple-choice problem among 0, 10, and 20, such as a binary choice problem between 10 and 20, which is easier to reach Pareto efficient coordination. Successful subjects may then see the multiple-choice problem in the third period as a binary choice between 20 and 30. By repeating such coordination success, subjects can coordinate to 60 in the sixth period coordination game. In other words, under the gradualism condition, **present success produces future success**.

A shortcoming of this "success produces success" story to the Pareto efficient equilibrium is that the upper limit of the coordination problem increases by itself and thus this path cannot be expected to apply once coordination failure occurs. This limitation motivates the need for the upper limit to undergo endogenous change. In other words, we must specify that only when subjects succeed in coordination and make a good precedent does the upper limit in the next period gradually increase.

The second tool is the endogenous change in the gradual increases in the bound; in other words, the bound only increases when coordination success to an efficient equilibrium has occurred in the previous period. In addition to waiting for a good precedent, there are two reasons for endogenous change to enhance coordination success compared with under the exogenous gradualism condition. First, endogenous change can motivate subjects to cooperate, as it links their efforts and purposes/goals (see Vroom 1964). As our endogenous change setting correlates these two elements, we can expect it to drive participants to cooperate. In particular, this tool makes subjects coordinate to a Pareto efficient equilibrium in the first period (**enhancing coordination success in the first period**). Second, under endogenous change, transition from an easy task to a difficult one is a result of subjects' own choices rather than enforced by an external agent. According to self-determination theory (Ryan and Deci 2000), when people experience autonomous motivation, they feel a profound sense of choice and autonomy; as a result, they adhere to their own choices. Thus, endogenous change will maintain Pareto efficient cooperation in the forthcoming harder task because of the participant's self-determined choice in the previous easier task (see Thibaut and Walker 1975; Tyran and Feld 2006; Sutter, Haigner and Kocher 2010). It will enhance the "present success produces future success" process, that can be encouraged by the gradualism.

Let $\mathrm{Min}^{\mathrm{t}}$ be the group minimum in period *t*. Under a **gradualism with endogeneity** condition, the upper limit is determined as follows ($\mathrm{m}^1 = 10$ and for any $\mathrm{t} \geq 2$):

$$m^t = \begin{cases} m^{t-1} + 10, m^{t-1} \leq 50 \; and \; Min^{t-1} = m^{t-1} \\ \qquad m^{t-1}, otherwise \end{cases}$$

It should be noted that the bound is unchanged in the case of coordination failure.

The third tool concerns what happens after coordination failure: the level of the next coordination game is adjusted to become an easier one in which players experienced coordination success to the efficient equilibrium in the past. This tool plays two roles in coordination success. The first effect is recovery from coordination failure. After coordination failure in a harder problem, participants rebuild mutual trust through success in an easier problem, which leads to further coordination success in future problems (**the power of recovery**). Second, this tool serves to sanction the non-cooperator. After failure in a harder problem, the next one becomes an easier but less profitable coordination problem. Therefore, the current coordination failure results in a loss of the future benefit that could have been obtained through a difficult but profitable game. This effect will deter subjects from choosing inefficient actions (**the effect of deterrence**).

Under a **gradualism with endogeneity and modification** condition, the upper limit is determined as follows ($\mathrm{m}^1 = 10$ and for any $\mathrm{t} \geq 2$):

$$m^t = \begin{cases} m^t + 10, m^{t-1} \leq 50 \; and \; Min^{t-1} = m^{t-1} \\ Min^{t-1}, otherwise \end{cases}$$

Let T be the number of repetitions of the minimum effort game under these three conditions. Then, each constitutes a finite repetition of minimum effort games, which allows us to calculate the equilibrium of the super-games of the three conditions. Although these three super-games are different, their equilibrium predictions are similar. In each, there exists a subgame perfect equilibrium that leads to an efficient outcome of a one-shot minimum effort game in every period. Moreover, there also exists a subgame perfect equilibrium that leads to the worst outcome (zero equilibrium) of a one-shot minimum effort game in every period. Thus, there are no critical differences among these three according to standard game theory. Nonetheless, subjects behave differently in the three conditions.

4 Experimental design

4.1 Treatments

We conducted three types of minimum effort game experiments, as described in Section 3. These experiments are labeled G (for the gradualism condition), GE (gradualism with endogeneity condition), and GEM (gradualism with endogeneity and modification condition). In addition, as a control treatment, we conducted a typical repeated minimum effort game experiment where the bound is constant at 60 for every period. This condition is called CON. In every treatment, the length of the repetition of minimum effort games was 20; thus, $\mathrm{T} = 20$ for each treatment.

Each subject joined one condition and 52 subjects are recruited to each of G, GE, and GEM, and 36 subjects to CON. Subjects were separated into groups of four, and group members were fixed throughout the duration of the experiment according to a partner matching design. Thus, there were 13 independent groups for G, GE, and GEM, and nine independent groups for CON. The treatments and their subjects are summarized in Table II.

[Place Table II Here]

4.2 Subjects

We recruited 192 (=52 ×3 + 36) undergraduate students from various disciplines. All subjects were recruited from Waseda University via the Internet. Written informed consent was obtained from all subjects. We conducted the experiment in July 2012.

4.3 Procedures

In all treatments, subjects were randomly assigned to laboratory booths at the beginning of the experiment. These booths separated subjects in order to ensure that every individual made his or her decision anonymously and independently. Subjects were provided with written instructions that explained the game, payoffs, and procedures. In particular, we explained that the upper bound of effort level varies across periods. This means that in G, every subject knows that at the beginning of the first period, the bound in the second period is 20, that in the third is 30, and so on. We adopted this setting because GE and GEM do not work as we expected if subjects are not informed about the changing bounds.[4] The instructions used neutral wording, as is common practice in experimental economics. After reading the instructions, subjects were tested to confirm that they understood the rules and knew how to calculate their payoffs. We did not start the experiment until all participants had answered all questions correctly. Therefore, all subjects completely understood the rules of this game and were able to calculate their payoffs.

Subjects were then randomly and anonymously allocated to groups of four, and these groups played the minimum effort coordination game. Group composition remained the same throughout the 20 study periods in order to retain statistically independent groups. Each group

[4] This point is different from the instructions of Ye et al. (2011) about the exogenous gradualism condition in a binary choice setting. As explained in Section 7, this difference may affect the experimental results.

member had to determine his or her effort level on the computer screen simultaneously. After their decisions, feedback was provided to subjects, such as current payoff and their group's minimum in this period; however, this information did not include the effort level of the other three players. This limited feedback information makes it difficult to achieve coordination success (Berninghaus and Ehrhart 2001; Brands and Cooper 2006b). After each experiment, all subjects returned their questionnaires.

We used z-Tree software (Fischbacher 2007) to conduct the experiments. Each session took approximately 1 hour to complete on average. Subjects' earnings were the sum of points gained in all 20 periods exchanged at a rate of 10 points = 5 yen. Subjects were also paid a participation fee of 500 yen. The mean payment per subject was 1343 yen (= $17.00 evaluated at 1$ = 79 yen). The maximum was 1625 yen (= $20.57) and the minimum was 690 yen (= $8.73).[5]

5 Hypotheses development

The presented hypotheses about coordination success and failure in our four experimental treatments are based on the discussion in Sections 2 and 3. The first three hypotheses concern the performance of these four conditions. Since we predict a positive effect of each of the three tools for coordination success, we hypothesize that the order of performance measured by profit and effort is GEM, GE, G, and CON in descending order. Note that statements (a), (b), and (c) in each hypothesis relate to the effects of gradualism, endogenous change, and modification, respectively.

Hypothesis 1. Comparing the total profit per subject, (a) the total profit per subject in G is greater than that in CON, (b) the total profit per subject in GE is greater than that in G, and (c) the total profit per subject in GEM is greater than that in GE. That is, concerning the total profit

[5] Actually, 900~1000 yen is equivalent to a student's hourly wage. Thus, the stake was substantial.

per subject, we predict that GEM > GE > G > CON.

Hypothesis 2. Comparing the effort level in a certain period per subject, for $t \geq 6$, (a) the effort level in G is greater than that in CON, (b) the effort level in GE is greater than that in G, and (c) the effort level in GEM is greater than that in GE. That is, concerning the effort level at a certain period, we predict that GEM > GE > G > CON.

Hypothesis 3. Comparing the number of Nash equilibria per group, (a) the number of Nash equilibria per group in G is greater than that in CON, (b) the number of Nash equilibria per group in GE is greater than that in G, and (c) the number of Nash equilibria per group in GEM is greater than that in GE. That is, concerning the number of Nash equilibria, we predict that GEM > GE > G > CON.

Hypothesis 4 concerns the common feature among the three treatments. As explained in Section 3, the strength of gradualism is the "success-produces-success" process that, based on self-determination theory and procedural fairness, we expect to be more clearly observed in GE and GEM than in G.

Hypothesis 4 (success produces success). (a) The probability of achieving efficient coordination success immediately after succeeding in efficient coordination in the previous period is high in G, GE, and GEM and (b) this is more clearly observed in GE and GEM than in G.

Hypothesis 5 tests why the performance in GE and GEM is superior to that in G. As explained in Section 3, endogenous change correlates the effort and game structure in the next period, while subjects in this condition share the belief that even inexperienced others will cooperate.

Hypothesis 5 (making the first coordination success). Comparing the success rate in the first period, (a) the first period's coordination success rate in GE is greater than that in G and (b) the first period's coordination success rate in GEM is greater than that in G.

The final two behavioral hypotheses test why the performance in GEM is superior to that in GE. As explained earlier, modification not only allows subjects to rebuild mutual trust after coordination failure but also prevents subjects from choosing inefficient actions through the threat of losing a future benefit. Therefore, this tool has two distinct effects on coordination success in the long run; one is recovery after coordination failure and the other is retaining efficient coordination.

Hypothesis 6 (power of recovery). The success rate after coordination failure in the previous period in GEM is greater than that in GE.

Hypothesis 7 (effect of deterrence or sustainability of coordination). Comparing the minimum of a period and that of the subsequent period, (a) the number of events that the group minimum becomes less than the minimum in the previous period in GEM is less than that in GE and (b) the events that the group minimum becomes less than the minimum in the previous period occur in a later period in GEM than in GE.

Hypothesis 6 states that the endogenous condition affects recovery compared with the gradualism condition, but that this is clearer when the bound is adjusted downwards in the subsequent period. Hypothesis 7 states that the effect of deterrence can be measured or observed in two ways: one is the number of events that disrupt coordination and the other is when such an event occurs. Table III summarizes these seven hypotheses.

[Place Table III Here]

6 Results

6.1 Performance of the four treatments

We first check the distribution of the group minimums in the final period of each treatment. Table IV shows that the group minimums in GEM and GE are either 0 or 60 and that the majority of groups in each treatment attain 60 (eight out of 13 for GEM and seven out of 13 for GE). By contrast, the group minimum in G ranges from 0 to 60 but most of the group end up with 0 (seven out of 13). The group minimum in CON is uniformly spread from 0 to 50, which seems to be worse than the results observed in the literature. However, the direction of this findings makes sense considering the limited feedback in our experimental design (Brands and Cooper 2006b) and the variance reported in the experimental results of multiple versions of a minimum effort game in a previous study (Engelmann and Normann 2010). For the average group minimums in the final period, GEM > GE > CON > G, implying that the performance of GEM is higher than that of GE, which is higher than that of G, but that the performance of G is lower than that of CON.[6] These findings are investigated next.

[Place Table IV Here]

[Place Table V Here]

[6] Not every possible comparison of the four treatments is statistically significant after the Bonferroni adjustment.

Next, we examine the total profit across the 20 periods per subject in each treatment. Table V reports the average total profit and theoretical maximum. Note that the theoretical maximum of total profit in CON is greater than that in GEM, GE, and G because the upper limit of effort level in the latter three conditions starts at 10 and it reaches 60 by period 6 if full coordination continues. Despite the smaller theoretical maximum, average total profit in GEM and GE is greater than that in CON. However, average total profit in G is smaller than that in CON. As a result, GEM > GE > CON > G holds. The t-test with Bonferroni adjustment shows that only the difference between GEM and G is significant at the 5% level.

We further investigate Hypothesis 1 using individual-level analysis. The individual variation in profit is relative not only to treatment differences but also to differences among groups and periods. Thus, we use a nested analysis of variance (ANOVA) model in each period and for all 20 periods to attribute the amount of variance in profits and total profits explained by treatment variations. Individual is nested in group and group is nested in treatment. A nested ANOVA allows us to measure the explanatory power of treatment variation for profits by taking into account group influences. Table VI summarizes the statistical analysis using the nested ANOVA for certain periods (first, sixth, 10^{th}, 15^{th}, and 20^{th}) and all 20 periods.[7] This table shows that:

(a) Contrary to our prediction, average profits in CON are not significantly different from those in G in almost all periods and thus total profits across all 20 periods are not significantly different;

(b) The difference between GE and G is highly significant in every period and that in total profit between the two is also highly significant;

(c) The difference between GE and GEM is highly significant in all latter periods and that in total profit between the two is highly significant as well.

[7] Owing to space limitations, we only show the data from certain important periods. All results are presented in the web appendix. This also holds for Table VI.

[Place Table VI Here]

The transition of group minimums and subject effort levels in each treatment is reported in Figure I. We see that these transitions in GEM and GE are similar in periods 1 to 8, but that a sharp decrease (gradual increase) in both occurs in GE (GEM) thereafter. The transitions in G are lower than those of GE and CON, whereas the transition of group minimum in CON is constant or weakly increasing across periods and that of subject effort is gradually decreasing. This finding implies that under the CON condition, the large variance in subject effort in the first half of the investigated period shrinks over time and that group members tend to coordinate at an effort level between 0 and 50 (see Table IV). In fact, the number of Nash equilibria, which imply coordination success at a certain effort level, increases over time (see Figure II).

[Place Figure I Here]

[Place Figure II Here]

To confirm Hypothesis 2, we also conduct a nested ANOVA model (Table VII). In short, the results are consistent with those on the profits for the four treatments:

(a) Contrary to our prediction, average effort in CON is significantly higher than that in G in certain periods;

(b) The difference between GE and G is highly significant in every period; and

(c) The difference between GE and GEM is highly significant in the latter half of the study period.

[Place Table VII Here]

To examine Hypothesis 3, we illustrate the transition of Nash equilibria over time. Figure II shows that the ratio of Nash equilibria in GEM is steady across periods at approximately 0.8, the highest rate among the four conditions. The ratio in GE is also steady from periods 6 to 18 at approximately 0.65. For G, the ratio starts at approximately 0.5 and sharply drops to 0.15 in period 3 before gradually increasing to close to 0.5. Finally, the ratio in CON starts at 0 and gradually increases to above 0.5 after a few periods. These results lend support to Hypothesis 3.

[Place Table VIII Here]

Moreover, the chi-square analysis presented in Table VIII states while H3(a) is not supported, both H3(b) and H3(c) are verified: each *p* value is 0.1026 (χ squared value=2.67, df=1), 2.208e-12 (χ squared value=49.29, df=1), and 1.857e-05 (χ squared value=18.33, df=1).

6.2. Effect of gradualism

We next selected participants who experienced full coordination in the first period and examined their efforts in the second period in order to test Hypothesis 4. This sample selection aimed to avoid any cumulative effect on individual decision making over time. Across the three conditions, a very high average coordination rate (average individual effort divided by 20) was observed: 0.96 for GEM, 0.97 for GE, and 0.94 for G. If we selected participants who experienced full coordination in the first and second periods, a similar result was observed at the third period: 0.98 for GEM, 0.99 for GE, and 0.97 for G. These results strongly support Hypothesis 4. Thus, the "success produces success" process was observed in all three treatments.

6.3 Effect of endogeneity

In this subsection, we examine Hypothesis 5, the behavioral reason why endogenous change in difficulty level is better than exogenous change. We first assess coordination success in the first period. Table IX shows that in GEM, 12 of the 13 groups attained perfect coordination in the first period, in GE 11 of the 13 groups did so, and under G six of the 13 groups did. Fisher's exact test with Bonferroni adjustment shows that both H5(a) and H5(b) are supported at the 10% level. If we combine GEM and GE as endogenous treatments (i.e., 23 of the 26 groups attained perfect coordination in the first period), the difference between them and the exogenous treatment (G) is highly significant according to Fisher's exact test ($p = 0.0077$). This result means that the endogenous condition may facilitate a "good first impression" (full coordination in the first period) compared with G.

We can also add the positive effect of a "good first impression" for Pareto efficient coordination. Table X shows that no group that did not deliver a "good first impression" reached Pareto efficient coordination. Comparing the ratio of Pareto coordination success between groups that delivered a "good first impression" and those that did not provides a p value of 0.00044 according to Fisher's exact test. Indeed, the psychological literature (see MacRae, Stangor and Hewstone 1996) emphasizes the importance of initial experiences for building trust or reputation, implying that a "good first impression" could influence coordination.

It is worth noting that the discussions on Hypotheses 4 and 5 are closely related. Although the "success produces success" process occurs in all three treatments, GEM and GE are superior to G in terms of total profits and coordination success (see Table VII). This finding may suggest that the relatively high performances of GEM and GE could be due to other elements than the "success produces success" path. The difference in the ratio of coordination in the first period is large between G and other two, which causes the overall difference in performance levels.

6.4 Effect of modification

Compared with the other three treatments, the most evident characteristic of GEM is that the minimum of the previous period automatically becomes the upper limit of the next period. Therefore, we chose cases where participants experienced coordination failure *first* except in the "first period failure" case to examine how they made an effort in the next period (Hypothesis 6). There are two reasons for this case selection. First, this is expected to remove the "good first impression" effect. Second, since each participant traces a periodic path, her/his experience may exert a considerable effect on her/his effort decision, which could be cumulative over time. This case selection also allows us to remove this cumulative effect as much as possible, because the selected participants' experiences before the first failure is similar, namely coordination success.

Of these, five of the six cases in GEM attained full coordination immediately compared with two of the six in GE and none of the four in G. According to Fisher's exact test, the difference between every pair was not significant. If we focus on the performance of modification, we compare five of the six cases in GEM and two of the 10 in "not GEM" (i.e., GE and G). This time, Fisher's exact test shows that the difference is statistically significant ($p = 0.034$).

To further examine this hypothesis we employ the logistic model with fixed and random effects, a form of generalized linear mixed model, using individual data. Our principal aim is to examine how GEM and GE affect individual effort just after the first coordination failure by removing a possible "group effect" as a random effect.

Let the probability of the ith subject of the jth group's maximum effort choice be p_{ij}, while the dependent variable is defined as logit (p_{ij}), namely $\log(p_{ij}/1-p_{ij})$. The explanatory variables are GEM Dummy$_{ij}$ and G Dummy$_{ij}$. GEM Dummy$_{ij}$ (G Dummy$_{ij}$) is coded 1 if the ith subject of the jth group faces the GEM (G) treatment and 0 otherwise. Thus, we estimate the following equation:

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_1 \times \ GEM\ Dummy_{ij} + \beta_2 \times G\ Dummy_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where β_1 and β_2 are the fixed-effect coefficients, which are identical for the groups, e_{ij} is an individual-level error, γ_{00} is the average outcome for the population, and u_{0j} is the j group's specific effect.

The estimation shows that β_1 is 2.037 with a p value of 0.07 and β_2 is -1.8871 with a p value of 0.008. These results mean that the probability that participants in the GEM treatment choose the maximum effort is approximately 7.7 times as much as the probability that participants choose the maximum effort in the GE treatment. In the same way, participants in G are approximately 0.15 times as cooperative as those in GE. Although we need more samples to examine the "recovery effect" in greater detail, these findings suggest that modification positively affects recovery.

In terms of Hypothesis 7, Table XI shows that any difference in frequency about the deteriorated minimum was not observed across the three treatments. Indeed, deterioration occurred very rarely. This phenomenon may not be surprising because the minimum can be a "focal point" of the next period's coordination, and coordination in the previous minimum may be profitable for participants. Further, while in GEM the minimum decreased in periods 4, 8, 19, and 20, in GE, it declined in periods 4, 5, 9, 10, 12, 13, and 18. The Wilcoxon rank sum test thus shows that deterioration occurs significantly later in GEM than in GE ($p = 0.089$), lending support to this hypothesis.

6.5 Summary of the results

We conclude this section with a summary of the results of hypothesis testing:

- For the total profit per subject (Hypothesis 1), the effort level of subjects (Hypothesis 2), and the number of Nash equilibria (Hypothesis 3):
  - Gradualism **does not** matter (G > CON does not hold);
  - Endogenous change matters (GE > G and GE > CON hold); and
  - Modification matters (GEM > GE holds).
- The success probability after coordination success is quite high in GEM, GE, and G (Hypothesis 4 (a) holds), but this is not statistically different among G and GE (GEM) (Hypothesis 4 (b) does not hold).
- Comparing the success rate in the first period (Hypothesis 5):
  - That in GE is greater than that in G (Hypothesis 5 (a) holds); and
  - That in GEM is greater than that in G (Hypothesis 5 (b) holds).
- The success rate after coordination failure in the previous period in GEM is greater than that in GE (Hypothesis 6 holds).
- Comparing the minimum of a period and that of the subsequent period (Hypothesis 7):
  - The number of the events that the group minimum becomes less than the minimum in the previous period in GEM is **not** less than that in GE (Hypothesis 7 (a) does not hold); and
  - The events that the group minimum becomes less than the minimum in the previous period occur in a later period in GEM than in GE (Hypothesis 7 (b) holds).

Based on the results of the hypothesis testing, the differences in performance between G, GE, and GEM can be explained as follows. Since the "success produces success" mechanism is relevant in all three conditions (Hypothesis 4), the performance diversity between G and GE mainly occurs because of the "good first impression" effect. As predicted by Hypothesis 5, endogeneity has a positive effect on facilitating coordination success among inexperienced participants in the first stage, which causes the overall large difference between G and GE. The difference between GE and GEM is, as predicted, due to the two positive effects of modification, namely recovery from coordination failure and the sustainability of high coordination through the threat to future benefits. These two effects are supported by our statistical testing.

In addition, Figure I shows that the average minimum in GEM gradually increases after the sixth period, while that in GE drops in that period. This finding is consistent with the two modification effects. Moreover, the average minimum in GEM drops in the final period, which is unexpected since coordination to 60 is a Nash equilibrium. However, this finding is consistent with the explanation why GEM enhances the sustainability of coordination success, namely because of the threat of losing a future benefit by adjusting the bound to the lower value. Because there is no future period at the end of the experiment, such a threat does not work.

7 Discussion and conclusion

In this study, we found that the three tools proposed herein (gradualism, endogeneity, and modification) are effective at enhancing coordination success in a laboratory setting. These findings suggest that the GEM mechanism needs neither a monetary incentive nor a strong enforcement authority that monitors and punishes individuals as necessary, which, in turn, makes coordination feasible without communication or observation among participants. Thus, the GEM would theoretically work in a highly anonymous setting where members in society are unfamiliar with one another. All we need is a credible announcement of a current bound or target from a third party and the common understanding of how to change this bound. Given these requirements, the GEM mechanism could be applied to several real-life situations that suffer from coordination failure. In this sense, we propose an alternative answer to the question posed by Weber (2006): "If a large laboratory group cannot coordinate efficiently, how do large communities and firms do so?" While Weber (2006) suggests a gradual change in the population from a small group to a larger one, we rather propose a GEM system.

The basic idea of GEM, namely the step-by-step coordination of people's economic activities, is consistent with prior economic development theory, such as the O-ring theory of Kremer (1993). This author points out that a flaw of O-rings (that cost less than $1) caused the Challenger shuttle crash in 1986 and states that under a strong technology complementarity, production requires each of many activities to be carried out well in order to obtain a high value output. Many production tasks should therefore be fulfilled by keeping in step. If the O-ring theory is valid to explain an aspect of economic development, it may be possible that while a type of GEM system is adopted to match workers that have similar skills with the production tasks necessary in developed nations, this is not the case in poorer countries.

Specifically, we showed that gradualism alone (G) does not work well, whereas the GE and GEM conditions achieve relatively more efficient coordination success. The poor performance of gradualism alone in our experiment seems to contradict the experimental results of Ye et al. (2011), which find that gradualism alone actually works better compared with the constant condition (termed the HighStart condition in Ye et al. [2011]) under their binary choice minimum effort coordination game.

The presented experimental results are different from theirs for several reasons. The first is that the performance of the CON condition in a multiple minimum effort game can be better than that in a binary minimum effort game. In most groups under the CON condition, subjects keep coordinating towards their group minimum in the first period, which is consistent with a previous study that found path dependency in a repeated minimum effort game (Van Huyck, Battalio and Beil, 1990, 1991). In the multiple-choice case, subjects' effort levels in the first period range from 0 to 60, and thus the group minimum in the first period is also scattered across the same range. In most groups, the minimum in the first period continues to the end of the experiment. By contrast, in the binary choice case between 0 and 60, the minimum in most groups is 0 because of the limited feedback provided (this finding is common to both ours and Ye et al.'s [2011] designs).

As a thought experiment, let us consider the scenario that every subject randomly chooses his or her effort level from a set of available effort levels. Thus, while the expected minimum of four group members is 9.47 in the multiple-choice case, it is 3.75 in the binary choice setting (approximately 2.5 times as high).[8] Although subjects' actual behavior is non-random, the result of this thought experiment clearly shows that the performance in CON under multiple choice is superior to that under the binary choice model.

The difference between these comparisons of exogenous gradualism under the constant condition does not only occur because of the performance under this control condition. In fact, our exogenous gradualism condition performs far worse than that in Ye et al. (2011) (average contribution in the final period is 11/18 × 60 = 36.667 in Ye et al. (2011) compared with 17.692 in ours). As confirmed in subsection 6.2, the "success produces success" process is observed in G, which is similar to the results presented by Ye et al. (2011). However, the success rate in the first period differs markedly. While the coordination success rate in the first period is approximately 46% (six groups out of 13) in our G condition, it is over 72% (13 groups out of 18) in their gradualism condition, which is more comparable to the results of our GE condition (84%; 11 groups out of 13).

This finding is puzzling because the first-stage game in both conditions is a four-person

[8] The average minimum of group members in the first period of our CON condition is 22.2. By contrast, that in Ye et al. (2011) is 0 × (15/18) + 60 * (3/18) = 10 (of course, in order to compare their results and ours, we need some translation of their results. In their experiment, the subjects choose between 0 and 14 under the HighStart condition. Thus, we replace 14 with 60 in the calculation). The result of a thought experiment thus seems to be valid.

binary choice minimum effort game.[9] A possible explanation is subjects' knowledge about the future upper limit (in our experiment) or stake (in Ye et al. 2011). In contrast to our design, the subjects in Ye et al.'s (2011) paper were not informed of the stakes in the future period. Indeed, this design is more realistic in some situations, but it allows subjects to suppose that future stakes are determined by previous choices (i.e., subjects think about an endogenous change in stakes). As our experimental results under the G and GE conditions show, the rule of changes they presume largely affects their choices in the first period. This may explain why first period performance in the gradualism condition of Ye et al. (2011) is more similar to the results in our GE condition.[10]

Our experiments confirmed that the difference between G and GE (GEM) mainly occurs based on whether subjects succeed in the coordination of an efficient action (i.e., choosing 10) in the first stage. While most groups succeed in this coordination in the GE and GEM conditions, fewer than half of them do so in the G condition. By contrast, after success in the first period, the groups in all three conditions frequently tend to succeed in the next and subsequent periods. Thus, the "success produces success" path is observed in these three conditions. We find smaller differences among the performances of the groups in the three conditions when examining only those that succeeded in the first period. As MacRae, Stangor, and Hewstone (1996) state, a "good first impression," which is important for building trust or reputation, can positively influence coordination.

To conclude this study, we suggest another utility of the GEM mechanism: it could be applied to other problems in future research because gradualism, endogeneity, and modification can all help facilitate coordination. In dilemma scenarios, reciprocal strategies such as tit for tat (TFT) (Axelrod and Hamilton 1981; Axelrod 1984) and raise the stakes (RTS) (Roberts and Sherratt 1998) are often effective at attaining high cooperation. The TFT strategy is principally

[9] Of course, the experimental parameters are slightly different. While in the first period of Ye et al. (2011), the choice is made between 0 and 1/10 of the endowment, subjects in our experiment chose between 0 and 1/6 of the endowment.

[10] To verify this reasoning, we must carry out an experiment on the minimum effort coordination game where the rule is the same as G but subjects do not know that rule. This is an interesting topic but it does not fit the research purpose of the present paper.

for binary choice (cooperation or defection) prisoners' dilemma games rather than multiple or continuous choice games, whereas the RTS strategy, whose basic idea is the same as TFT, is applied to multiple-choice prisoners' dilemma situations (Roberts and Sherratt 1998).

Roberts and Renwick (2003) conduct a continuous choice repeated prisoners' dilemma experiment and find that some pairs of subjects maintain a high cooperation level by using a kind of RTS strategy. However, even such an RTS strategy is not expected to work in social dilemmas that are characterized by more than two participants. As the number of group members increases, it becomes rapidly difficult for them to maintain a high cooperation level—even under a reciprocal strategy—as the range of personality types (e.g., egoistic, altruistic, or reciprocal) rises (Dawes 1980).

The above argument implies that the achievement and maintenance of cooperation in social dilemma situations using only a TFT or an RTS strategy may be challenging. From this perspective, the GEM mechanism can guide people to use a kind of TFT or RTS strategy in two directions. First, it can demonstrate how to behave in repeated social dilemma situations. Second, it can create a shared belief that other people use a TFT or an RTS strategy. In other words, the GEM mechanism internalizes the behavioral properties of the TFT or RTS in society and forces society members to use a reciprocal strategy. Although the present study focused on a coordination problem, the GEM mechanism can also be applicable to cooperation problems (i.e., social dilemmas). Investigating this point theoretically and experimentally is a future research direction.

References

Axelrod, R., and W. D. Hamilton, "The evolution of cooperation," *Science,* 211 (1981), 1390 –1396.

Axelrod, R., *The evolution of cooperation*. (New York: Basic Books).

Berninghaus, K., and K.-M. Ehrhart, "Coordination and information: Recent experimental evidence ," Economics Letters 73 (2001), 345-351.

Brands, J., and D. Cooper, "A change would do you good …. An experimental study on how to overcome coordination failure in organization," *American Economic Review*, 96 (2006a), 669–693.

Brands, J., and D. Cooper, "Observability and overcoming coordination failure in organizations: An experimental study," *Experimental Economics,* 9 (2006b), 407–423.

Brandts, J., D. J. Cooper, and E. Fatas, "Leadership and overcoming coordination failure with asymmetric costs," *Experimental Economics,* 10 (2007), 269–284.

Brandts, J., D. J. Cooper, and E. Fatas, "Stand by me: Help, heterogeneity and commitment in experimental coordination games," mimeo, 2011.

Cason, T. N., A. C. Savikhin, and R. M. Sheremeta, "Behavioral spillovers in coordination games," *European Economic Review,* 56 (2012), 233–245.

Cooper, D. J., "Are experienced managers experts at overcoming coordination failure?," *The B.E. Journal of Economic Analysis & Policy*, 6 (2006), 6.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross, "Communication in the battle of the sexes game," *Rand Journal of Economics,* 20 (1989), 568–587.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross, "Selection criteria in coordination games: Some experimental results," *American Economic Review,* 80 (1990), 218–233.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross, "Communication in coordination games," *Quarterly Journal of Economics,* 107 (1992), 739–771.

Cooper, R., and A. John, "Coordinating coordination failures in Keynesian models," *Quarterly Journal of Economics* 103 (1988), 441–463.

Dal Bo, P., A. Foster, and L. Putterman, "Institutions and behavior: Experimental evidence on the effects of democracy," NBER Working Paper Series 13999, 2008.

Dawes, R. M., "Social dilemmas," *Annual Review of Psychology,* 31 (1980), 169–193.

Devetag, G., "Precedent transfer in coordination game: An experiment," *Economics Letters,* 89

(2005), 227–232.

Devetag, G., and A. Oltmann, “When and why? A critical survey on coordination failure in the laboratory,” *Experimental Economics,* 10 (2007), 331–344.

Diamond, P. A., “Aggregate demand management in search equilibrium,” *Journal of Political Economy,* 90 (1982), 881–894.

Engelmann, D., and H.-T. Normann, “Maximum effort in the minimum-effort game,” *Experimental Economics,* 13 (2010), 249–259.

Fischbacher, U., “Z-tree. Zurich toolbox for readymade economic experiments,” *Experimental Economics,* 10 (2007), 171–178.

Frey, B., and R. Jegen, “Motivation crowding theory,” *Journal of Economic Survey,* 15 (2001), 589–611.

Frey, B., and F. Oberholzer-Gee, “The cost of price incentives: An empirical analysis of motivation crowding-out,” *American Economic Review,* 87 (1997), 746–755.

Frey, B., A. and Stutzer, “Beyond outcomes: Measuring procedural utility,” *Oxford Economic Papers*, 57 (2005), 90–111.

Greif, A., “Contract enforceability and economic institutions in early trade: The Maghribi traders’ coalition,” *American Economic Review,* 83 (1993), 525–548.

Guillen, P., C. Schwieren, and G. Staffiero, “Why feed the Leviathan?,” *Public Choice* 130 (2006), 115–128.

Hamman, J., S. Rick, and A. Weber, “Solving coordination failure with “all-or-none” group-level incentives,” *Experimental Economics,* 10 (2007), 285–303.

Houssart, J., “Simplification and repetition of mathematical tasks: A recipe for success or failure?,” *Journal of Mathematical Behavior,* 21 (2002), 191–202.

Knez, M., and C. Camerer, “Creating expectational assets in the laboratory: Coordination in `weakest link’ games,” *Strategic Management Journal,* 15 (2006), 101–119.

Knez, M., and D. Simester, “Firm-wide incentives and mutual monitoring at Continental

Airlines," *Journal of Labor Economics,* 19 (2001), 743–772.

Kremer, M., "O-ring theory of economic development," *Quarterly Journal of Economics,* 108 (1993), 551–575.

Kuipers, H., and H. A. Keizer, "Overtraining in elite athletes: Review and directions for the future," *Sports Medicine,* 6 (1988), 79–92.

Kurzban, R., M. K. Rigdon, and B. J. Wilson, "Incremental approaches to establishing trust," *Experimental Economics,* 11 (2008), 370–389.

Lepper, M. R., D. Greene, and R. Nisbet, "Undermining children's intrinsic interest with extrinsic reward: A test of the 'overjustification' hypothesis," *Journal of Personality and Social Psychology,* 28 (1973), 129−137.

Lind, E. A., and T. R. Tyler, *The social psychology of procedural justice*. (New York: Plenum).

MacRae, C. N., C. Stangor, and M. Hewstone, *Stereotypes and stereotyping*. (UK: Guilford Press).

Markussen, T., L. Putterman, and J.-R. Tyran, "Self-organization for collective action: An experimental study of voting on formal, informal, and no sanction regimes," mimeo, 2011.

Pommerehne, W., H. and Weck-Hannemann, "Tax rates, tax administration and income evasion in Switzerland," *Public Choice,* 88 (1996), 161–170.

Roberts, G., and J. S. Renwick, "The development of cooperative relationships: An experiment," *Proceedings of the Royal Society,* 270 (2003), 2279–2283.

Roberts, G., and T. N. Sherratt, "Development of cooperative relationships through increasing investment," *Nature* (1998), 175–179.

Rousseau, Jean-Jacques, *Discours sur l'origine et les fondemens de l'inégalité parmi les homes*, Amsterdam: Marc Michel Rey, (1755).

Ryan, R. M., and E. L. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being," *American Psychologist,* 55 (2000), 68–78.

Sullivan, P., J. Mousley, and R. Zevenbergen, "Teacher actions to maximize mathematics learning opportunities in heterogeneous classrooms," *International Journal of Science and Mathematics Education,* 4 (2006), 117–143.

Sutter, M., S. Haigner, and M. G. Kocher, "Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations" *Review of Economics Studies,* 77 (2010), 1540–1566.

Tyran, J.-R., and L. P. Feld, "Achieving compliance when legal sanctions are non-deterrent," *Scandinavian Journal of Economics,* 108 (2006), 135–156.

Thibaut, J., and L. Walker, *Procedural justice*. (Hillsdale, NJ: Lawrence Erlbaum).

Van Huyck, J. B., R. C. Battalio, and R. O. Beil, "Tacit coordination games, strategic uncertainty, and coordination failure." *American Economic Review,* 80 (1990), 234–248

Van Huyck, J. B., R. C. Battalio, and R. O. Beil, "Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games," *Quarterly Journal of Economics,* 106 (1991), 885–910.

Vroom, V. H., *Work and motivation*. (New York: McGraw-Hill).

Weber, R., "Managing growth to achieve efficient coordination in large groups," *American Economic Review,* 96 (2006), 1114–1126.

Weingast, B. R. "The political foundations of democracy and the rule of law," *American Political Science Review*, 91 (1997), 245–263.

Ye, M., S. Asher, L. Casaburi, and P. Nikolov, "One step at a time: Does gradualism build a coordination?," mimeo, 2011.

The other three's minimum

First Period Game

		0	10	20	30	40	50	60
	0	60	60	60	60	60	60	60
	10	50	70	70	70	70	70	70
Your	20	40	60	80	80	80	80	80
Effort	30	30	50	70	90	90	90	90
	40	20	40	60	80	100	100	100
	50	10	30	50	70	90	110	110
	60	0	20	40	60	80	100	120

Table I: Payoff table for a minimum effort game

The diagonal line represents the best response to the other three's minimum.

	# of repetitions	Group size	# of subjects	# of groups
G (gradualism)	20	4	52 for each treatment	13 for each treatment
GE (gradualism with endogeneity)				
GEM (gradualism with endogeneity and modification)				
CON (bound fixed at 60)			36	9

Table II: Summary of treatments

	Performance (Profit, effort level, # of NEs)	Success produces success	Making a good first impression	Power of recovery	Sustainability of coordination
CON	×	×	×	×	○
G	△	○	△	×	○
GE	○	◎	○	○	○
GEM	◎	◎	○	◎	◎
Hypothesis	1,2,3	4	5	6	7

Table III: Summary of hypotheses

	0	10	20	30	40	50	60	AVE
GEM	4	0	0	0	0	1	8	40.769
GE	4	2	0	0	0	0	7	33.846
G	7	1	1	0	2	0	2	17.692
CON	2	1	2	1	1	1	0	25.555

Table IV: Distribution of group minimums in the final period (period 20)

	GEM	GE	G	CON
Average of experimental data	1929.231	1775.192	1463.462	1528.056
Theoretical maximum	2250	2250	2250	2400

Table V: Average total profit across the 20 periods per subject

	1	6	10	15	20	Total
GEM	68.7	97.3	101.3	105.6	97.7	1929.2
GE	67.3	94.4	89.0	91.7	92.7	1775.2
G	61.2	73.3	67.3	76.3	71.0	1463.5
CON	65.8	71.9	74.2	80.6	81.4	1528.1
G vs. CON	n.s.	n.s.	†	n.s.	***	n.s.
GE vs. G	***	***	***	***	***	***
GEN vs. GE	n.s.	***	***	***	***	***

Table VI: Total profit and average profit for given periods and the results of the nested ANOVA

(`***' - *p*=0; `**' - *p*<0.001; `*' - *p*<0.01; `.' - *p*<0.05; `†' - *p*<0.1; `n.s.' – Not Significant)

	1	6	10	15	20
GEM	9.8	39.6	43.3	46.7	43.8
GE	9.6	39.4	38.7	35.9	35.0
G	8.1	26.7	25.0	23.7	24.4
CON	38.6	32.5	30.2	28.3	29.7
G vs. CON	***	†	n.s.	†	n.s.
GE vs. G	†	***	***	***	***
GEN vs. GE	n.s.	n.s.	***	***	***

Table VII: Average effort for given periods and the results of the nested ANOVA

(`***' - *p*=0; `**' - *p*<0.001; `*' - *p*<0.01; `.' - *p*<0.05; `†' - *p*<0.1; `n.s.' – Not Significant)

	Nash equilibrium	Not Nash equilibrium	Success rate
GEM	203	47	0.81
GE	166	94	0.64
G	85	175	0.33
CON	45	135	0.25
G vs. CON			n.s.
GE vs. G			***
GEN vs. GE			***

Table VIII: The number of Nash equilibrium outcomes and results of chi-square analysis

(`***' - *p*=0; `**' - *p*<0.001; `*' - *p*<0.01; `.' - *p*<0.05; `†' - *p*<0.1; `n.s.' – Not Significant)

	Success	Failure
GEM	12	1
GE	11	2
G	6	7

Table IX: The number of coordination successes to the Pareto efficient equilibrium in the first period

	Attaining 60 equilibrium	Not attaining 60 equilibrium
Good first impression	19	10
Bad first impression	0	10

Table X: Effect of "first impression" on Pareto efficient coordination

	Deteriorated	Maintained	Ameliorated	Deterioration rate
GEM	9	667	312	0.99
GE	8	619	361	0.99
G	17	542	429	0.98

Table XI: Change in minimum effort across the three conditions

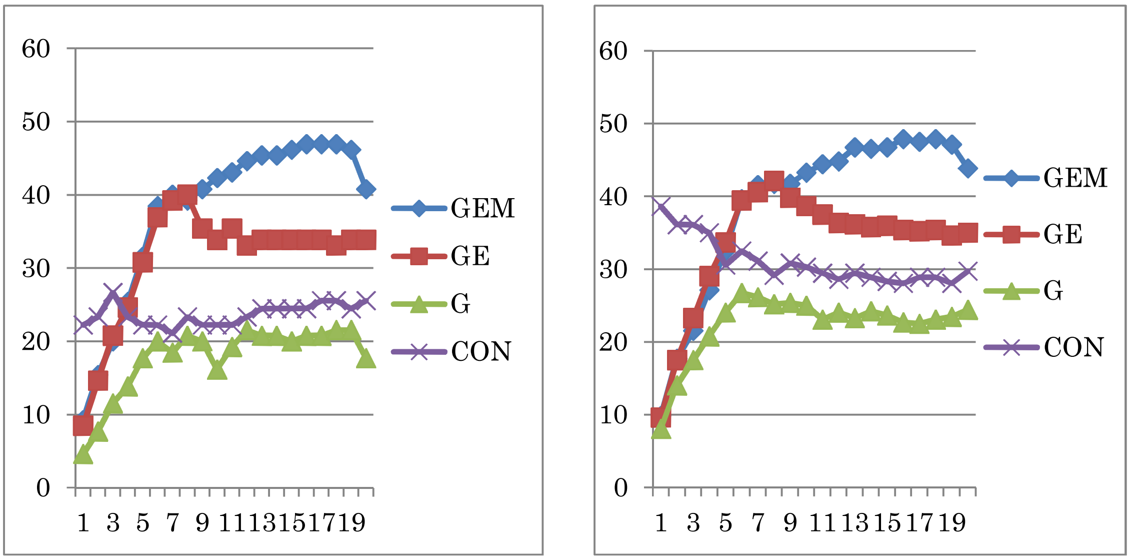

Figure I: Left panel is the transition of the average group minimum of each treatment; Right panel is the transition of the average subject's effort of each treatment

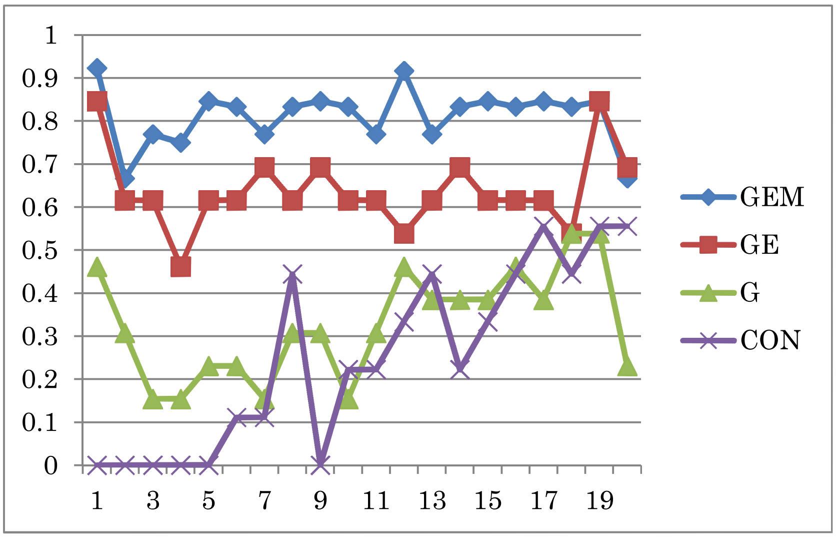

Figure II: The transition of the ratio of Nash equilibria for each treatment