One Step at a Time: Does Gradualism Build Coordination?*

Maoliang Ye[†], Sam Asher[‡], Lorenzo Casaburi[§], Plamen Nikolov[**]

[*] We would like to thank Alberto Alesina, Jim Alt, Nejat Anbarci, Abhijit Banerjee, Max Bazerman, Iris Bohnet, Hannah Bowles, Raj Chetty, Stephen Coate, David Cutler, Sreedhari Desai, Ernst Fehr, John Friedman, Roland Fryer, Drew Fudenberg, Edward Glaeser, Francesca Gino, Ben Goodrich, Haifeng Huang, Torben Iversen, Garett Jones, Yuichiro Kamada, Larry Katz, Judd Kessler, David Laibson, Randall Lewis, Jeffrey Liebman, Erzo F. P. Luttmer, Brigitte Madrian, Louis Putterman, Alvin Roth, Kenneth Shepsle, Andrei Shleifer, Monica Singhal, Dustin Tingley, Yan Yu, Tristan Zajonc, Richard Zeckhauser, Yao Zeng, and participants at Public Choice Society 2011 Meeting in San Antonio, and seminars at Harvard Economics Department, Government Department, Kennedy School and Business School for valuable comments and suggestions at different stages of this project. Maoliang Ye also thanks the Institute for Quantitative Social Sciences at Harvard University for financial supports, the School of Economics at Renmin University of China for generously providing the laboratory, Ze Yu for coordinating the use of laboratory, and Yuqi Zheng for excellent research assistance. All errors are our own.

[‡] Harvard University, Department of Economics, Cambridge, MA 02138.
[§] Harvard University, Department of Economics, Cambridge, MA 02138.
[**] Harvard University, Department of Economics, Cambridge, MA 02138.

One Step at a Time:
Does Gradualism Build Coordination?

Abstract

We study how gradualism -- increasing required levels ("thresholds") of contributions slowly over time rather than requiring a high level of contribution immediately -- affects individuals' decisions to contribute to a public project. Using a laboratory binary choice minimum-effort coordination game, we randomly assign participants to three treatments: starting and continuing at a high threshold, starting at a low threshold but jumping to a high threshold after a few periods, and starting at a low threshold and gradually increasing the threshold over time (the "gradualism" treatment). We find that individuals coordinate most successfully at the high threshold in the gradualism treatment relative to the other two groups. We propose a theory based on belief updating to explain why gradualism works. We also discuss alternative explanations such as reinforcement learning, conditional cooperation, inertia, preference for consistency, and limited attention. Our findings point to a simple, voluntary mechanism to promote successful coordination when the capacity to impose sanctions is limited.

JEL Classifications: C91; C92; D03; D71; D81; H41

Keywords: Gradualism; Coordination; Cooperation; Public Goods; Belief-based Learning; Laboratory Experiment

1. Introduction

Cooperation and coordination[1] are important for daily economic, social and political activities (Schelling, 1960; Arrow, 1974; Cooper & John, 1988; Weingast, 1997). Yet, coordination failure is common (Van Huyck et al., 1990, 1991; Cooper et al., 1990; Knez & Camerer, 1994, 2000; Cachon & Camerer, 1996). Some studies find sanction institutions, social pressure and reputation to be mechanisms that promote cooperation (e.g., Olson, 1971; Ostrom, Walker & Gardner, 1992; Fehr & Gächter, 2000; Masclet et al., 2003; Gächter & Herrmann, 2010; Bochet et al., 2006; Carpenter, 2007).[2] Others explore methods to facilitate coordination when sanction and social pressure cannot be imposed, such as repetition with fixed group members (Clark & Sefton, 2001), complete information structure (Brandts & Cooper, 2006a), communication (Cooper et al., 1992; Charness, 2000; Weber et al., 2001; Duffy & Feltovich, 2002; Chaudhuri et al., 2009), between-group competition (Bornstein et al., 2002; Riechmann & Weimann, 2008), and gradual organizational growth (Weber, 2006).

In this paper we study gradualism within a fixed-size group, a natural and voluntary mechanism to promote coordination where sanction and social pressure are absent. We refer to gradualism as the hypothesis that allowing agents to coordinate first on small and easy-to-achieve goals facilitates later coordination on otherwise hard-to-achieve outcomes ("public goods"). This hypothesis has a long history in the domain of political science and international relations. Abbott and Snidal (2002) present a recent analysis highlighting the role that a gradualist approach played in the development of the 1997 OECD Anti Bribery Convention. The authors study how breaking the final goal into a series of steps allowed the players to overcome the challenges emerged in previous *big-bang* approaches. Other real world examples of gradualism in coordination include the development of the World Trade Organization and the

[1] In the experimental literature, generally "cooperation" refers to choosing cooperative action when non-cooperative action is predicted by equilibrium in one-shot game based on self-interest (i.e., free-riding or deviating, rather than contributing or cooperating, is the best response given other others cooperating). The term is mostly used in public good games and prisoners' dilemma games. "Coordination" refers to achieving a high efficient equilibrium rather than a low inefficient equilibrium when there are multiple equilibria; it is used in coordination games. In this paper, we use "coordination" in the traditional sense and focus on coordination games with Pareto-ranked equilibria rather than other coordination games. But we employ the term "cooperation" more broadly than the tradition, one that also includes cooperative actions in coordination game: achieving a high equilibrium involves cooperative actions of all players.

[2] Arguably, these mechanisms can also help establish successful coordination.

European Union,[3] which are the result of decades of negotiation rounds, international arms reduction agreements, and regional and international environmental cooperation. Finally, gradualism finds its application in intra-organization team building (e.g. new employees may be given small initial tasks to help build coordination and to make sure that they can coordinate well in larger tasks later).

Mimicking a typical coordination setting in the real world, we conduct a computer-based laboratory experiment with repeated interactions. In each period, participants can only choose whether to contribute or not to the group's common pool, with the size of the contribution (which we call *threshold* or *stake) fixed for each period and allowed to vary across periods*. Each member realizes an extra return only when all group members contribute to the pool; otherwise each ends up with only the points that she does not give. This set-up is generally referred to as the minimum-effort or weak-link coordination games: the payoff depends on one's effort and the minimum effort of group members. Our setting simplifies the payoff function[4] and is similar to a standard discrete public good game, except that our set-up does not offer an opportunity for any individual to free ride.[5] Specifically, it is a multi-period stage hunt game due to the binary-choice feature in each period.

To study gradualism, we introduce three main treatments of threshold patterns, which differ in the first 6 periods but have an identical threshold for the final 6 periods. The first treatment has a constant high threshold for all 12 periods; we call this "High Start." The second treatment has a constant low threshold for the first 6 periods and jumps to the high threshold for the final 6 periods; we call this "Big Jump." Finally, in the "Gradualism" treatment, the threshold gradually increases in each of the first six periods until it reaches the high threshold by period 7. (See

[3] Increased coordination level and organization growth are two key features in EU development. The former is the focus of our study; Weber (2006) examines the latter. A detailed introduction of the EU history can be found at: http://europa.eu/about-eu/eu-history/index_en.htm.

[4] The literature of these games generally uses a complex matrix of payoff: there are 7 choices of actions, and the payoff depends on own action and the minimum action of others. See Van Huyck et al. (1990), Knez and Camerer (1994, 2000), Cachon and Camerer (1996), Weber (2006), and Chaudhuri et al. (2009).

[5] Free riding is generally a necessary feature of a public good game in the experimental economics literature. However, our set-up also belongs to the "weakest-link public goods game" as described by several theoretical papers on public goods (e.g., Hirshleifer, 1983; Cornes & Hartley, 2007). Cornes and Hartley (2007) provide a general social composition function of public goods with input of individual gifts, which incorporate standard continuous public goods, standard discrete public goods, weak-link (and weakest-link) public goods, good-shot (and best-shot) public goods, etc.

Figure 1 for a graphical illustration.) Exploiting this design, we study the effect of gradualism on coordination at a high level of threshold.

Consistently with the stated hypothesis, our experimental results show that the gradualism treatment attains significantly more successful coordination at a high level of threshold. In our lab experiment, subjects in the "Gradualism" treatment are more likely to contribute in the final high threshold level periods than the other two main treatments. In terms of magnitude, the effects are quite large – for example, in the end period, 61.1% of the "Gradualism" groups successfully coordinate, while only 16.7% and 33.3% of "High Start" and "Big Jump" groups do, respectively.

Our experimental results also suggest an externality of coordination building (or collapse) across different social groups. Those treated in the gradualism setting are about 10 percentage points more likely to cooperate upon entering a new environment. However, when they find their cooperation does not get rewarded in a new environment (because the new group members may have been treated differently and have different coordination outcomes in the first stage), they tend to become less cooperative.

We propose a simple theoretical framework of belief-based learning which explains our empirical findings. Participants have prior beliefs about others' actions before the game starts, and update the beliefs according to the outcome in each period. A low threshold level at the beginning makes it cheap to attempt coordination in the face of uncertainty, thus giving them stronger beliefs that others would try to cooperate. Similarly, at each level where they manage to coordinate, they reinforce their beliefs about the likelihood of future successful coordination at slightly higher thresholds. On the contrary, if the threshold starts at a high level, initial failures undermine future cooperation. Finally, in the presence of big jumps, previous coordination at low thresholds does not affect players' posteriors on actions at substantially higher thresholds levels.

We also discuss alternative explanations of the gradualism mechanism. A naïve reinforcement learning model, standard conditional cooperation, inertia, and preference for consistency have troubles explaining the "Big Jump" treatment. Limited attention or bounded awareness is not likely to apply for our experimental design. We discuss all these in depth in Section 5.

Although numerous studies study public good games[6] and coordination games, to the best of our knowledge, ours is the first study that clearly tests the role of exogenous gradualism in coordination within a given group and employs underlying behavioral theories. It is also among the first studies about how cooperation evolves over time with varying levels of stakes in multi-period experiments.

It is worth pointing out that our study, adopting an exogenous (versus endogenous) setting of threshold path, tries to answer a mechanism design question from the perspective of social planner (e.g., an employer who cares about the performance of the firm or organization, or a benevolent government): what is the optimal path to build successful coordination at a high threshold? Our findings point to a simple, voluntary mechanism to promote coordination when the capacity to impose sanctions is limited.

Compared to the existing literature on coordination games, this paper also suggests a case where limiting choices can improve social welfare. The literature on coordination games generally uses a complex payoff structure in which each individual has as many as seven choices for actions in each period (e.g., Van Huyck et al., 1990; Knez & Camerer, 1994, 2000; Cachon & Camerer 1996; Weber, 2006; Chaudhuri et al., 2009). For those with higher willingness to give in the first period, once they observe their high-cooperative actions are harmed by low-cooperative actions of their partners, they reduce the level of cooperation. After just several periods, an inefficient outcome is attained rather than a high efficient outcome, and the groups are then trapped in this low equilibrium. In contrast, in our experiment the participants are restricted to two choices in each period: giving a specified amount or not giving. A gradualism institution, which increases the specified amount gradually, maintains participants' high willingness to give even when the specified amount becomes substantial. As shown in this study, limiting choices in each period (but without mandatory or semi-mandatory institutions, e.g., sanction, punishment, and social pressure), plus a well-designed institutional path, helps participants reach a social optimal outcome.

The rest of the paper is organized as follows. In Section 2, we discuss the related literature with more details. In Section 3, we detail the experimental design. In Section 4, we present the graphs, tables and regressions, as well as explanations of the major results. Section 5 proposes

[6] For example, Palfrey and Rosenthal, 1984; Isaac, Schmidtz, and Walker, 1989; Bagnoli and Lipman, 1989, 1991; Marks and Croson, 1998, 1999. For an early literature survey, see Ledyard (1995).

micro theories behind gradualism. Section 6 discusses potential extensions for future research. Section 7 concludes.

2. Literature

Our paper is closely related to two themes in the literature. The first theme explores comparative statics of thresholds in minimum-effort coordination games and discrete public good games. The second strand explores theories and experimental results about gradualism in prisoners' dilemma games, public goods games, coordination games and trust games.

Schram, Offerman and Sonnemans (2008) study the comparative statics of thresholds in discrete public good provision, and find that a higher threshold lowers the rate of successful provision. Through fixing the required number of contributors and change the group size, the authors show that a smaller group size means a higher threshold.

In a laboratory dynamic coordination experiment, Weber (2006) studies the dynamics of organizational growth and finds that gradually growing the size of group leads to more successful coordination in a large group versus starting with a large group. Our study differs from Weber (2006) in four major ways: (1) we explore gradualism in coordination within a given fixed-size group; (2) in our experiment, the choice set in each period is binary, and the payoff structure is much simpler than the one in Weber's study; (3) We have a third main treatment "Big Jump," which helps us mimic some other important coordination processes in the real world, and distinguish alternative explanations of gradualism; and (4) our theoretical model incorporates the results of his study, while his cannot explain ours.

Our experimental design and results show much more clearly the efficiency gains of gradualism than the study by Andreoni and Samuelson (2006). They examine a twice-played prisoners' dilemma in which the number of total stakes in the two periods is fixed, while the distribution of these stakes can be varied across period. Both theoretical and experimental results show that it is best to "start small," reserving most of the stakes for the second period. However, cooperation is low for the period with a high stake,[7] which shows their gradualism setting

[7] When the relative stake of period 2 is high, there is more cooperation in period 1 but less cooperation in period 2; when the relative stake of period 2 is low, less cooperation in period1 but more cooperation in period 2.

actually does not largely help improve cooperation at a high level of difficulty and stake and cannot serve as an effective tool to build cooperation.[8]

Offerman and van der Veen (2010) study whether governmental subsidies to promote public good provision should be abruptly introduced or gradually increased, i.e., given the benefit of the public good, whether the individual cost of providing the public good should be decreased sharply or gradually. Their result favors an immediate increase of subsidy: when the final subsidy level is substantial, the effect of a quick increase is much stronger than that of a gradual increase. Our study differs from theirs in the following important ways. First, their motivation is about how to use subsidies to stimulate cooperation after cooperation failures at the beginning. Our mechanism, on the other hand, does not need governmental subsidies to promote cooperative behavior; instead we use a low threshold (stake) to encourage earlier cooperation and a gradual increasing stake path to sustain cooperation at higher levels. Second, in our experiment, what may change is the stake level, which indicates both the cost and the benefit of the public good. Third, our stake paths are non-decreasing, while their cost paths are non-increasing.

There are some studies on the monotone game, a multi-period game in which players are constrained to choose strategies that are non-decreasing over time, i.e., to increase contributions over time (e.g., Gale, 1995, 2001; Lockwood & Thomas, 2002; Choi et al., 2008). In contrast to these studies, we employ a more natural setting and force the threshold (indicating difficulty and stake) instead of forcing the contribution decision to be non-decreasing.

Pitchford and Snyder (2004) develop a model where the sequence of gradually smaller investments solves the holdup problem when the buyer's ability to hold up a seller's investment is severe. However, Kurzban et al. (2008) contradict this prediction by showing that participants prefer starting with smaller levels of investment and increasing it, rather than the other way around.

In two theoretical papers, Watson (1999, 2002) shows "starting small and increasing interactions over time" is an equilibrium under renegotiation and commitment conditions. In our study, the stake levels over periods are exogenous, which differs from Watson (1999); and individuals do not know the path of stake levels in later periods, which differs from Watson (2002).

[8] One example (and motivation) for their study is as follows: "An employer forming a new team of workers may give them small initial tasks, to help build cooperation, followed by larger tasks that can take advantage of that cooperation." According to their results, actually the workers will fail to reach successful cooperation for larger tasks.

3. Experimental Design

Our study is a variation of the discrete multi-period public goods games, and can be classified as a minimum-effort coordination game with a much simpler payoff structure than the standard coordination games in the literature.

3.1 Sample and Payoff Structure

The laboratory experiment was conducted at Renmin University of China in Beijing, China in July 2010 with 256 participants recruited via the Bulletin Board System (BBS) and posters at the university. Most participants were students from this university and universities nearby, and people living nearby. All laboratory sessions were computerized using the z-Tree experiment software package (Fischbacher, 2007). Both the instructions and the information shown on the computer screen were in Chinese.

There were 18 sessions. In each session, we randomly assigned participants to groups of 4,[9] so our sample consisted of 64 groups in total. The experiment included two stages: the first stage comprised 12 periods, while the second one had 8 periods. In each period, we endowed participants with 20 points and asked them to give a certain number of points to their assigned groups' common pools. The required number could vary across periods, and each participant could only choose either "not to give" or "to give the *exact* points required," which we refer to as *threshold* or *stake.* If all members in a group gave,[10] then each member not only got the points she had given back, but also gained an extra return which equaled the required number of points (i.e., the threshold). If not all group members contributed, then each member finished each period only with her points remaining, and the points she gave out were wasted.

In sum, the following formula describes each period earnings:

[9] For coordination games, four is considered as a small or moderate group size. For public goods games, Croson and Marks (2000) show in a meta-analysis study that the most frequently used group sizes are 4, 5, and 7. In their own studies (Marks & Croson, 1998, 1999), each group consists of 4 players.

[10] Requiring all members to give makes the game a minimum-effort coordination game. Other reasons for this requirement are: First, it makes cooperation more difficult given the small group size of 4, and a higher difficulty level is where we may need gradualism; second, it makes the theory simple (for each single period, there are only two pure strategy equilibria, one with all group members giving, the other with no group members giving.) In an accompanying paper coming soon, the role of gradualism in voluntary provision of public goods is studied by allowing free riding in the game.

$$Earning_{i,t} = \begin{cases} 20 + Th_t, & if\ A_{i,t} = C\ \ and\ A_{j,t} = C, \forall j \neq i \\ 20, & if\ A_{i,t} = D \\ 20 - Th_t, & if\ A_{i,t} = C\ \ and\ \exists j \neq i,\ s.t.\ A_{j,t} = D \end{cases}$$

Where $Earning_{i,t}$ is i's earning in period t, Th_t is threshold (stake) at t. $A_{i,t}$ and $A_{j,t}$ are the actions of i and j at t, respectively (i and j are in the same group.) C represents "cooperate" ("give"), while D represents "deviate" ("not give").

The final payment is the total earning accumulated over periods plus a show-up fee, and the exchange rate is 40 points per *yuan*.[11] The summary of the experimental design is shown in Table 3.

3.2 Treatment Group Assignments

Our experiment consisted of three main treatments: "High Start," "Big Jump" and "Gradualism." All groups in the three main treatments faced the same threshold in the second half (periods 7-12) of the first stage, but the threshold paths for them differed in the first half (periods 1-6), which may have imposed an income effect in the laboratory. To estimate how much of the difference in performance among the three main treatments in the second half of the first stage is driven by an income effect, we introduced a variant of the "HighStart" treatment, namely the "High Show-up Fee" treatment, which we describe in more detail below. In 8 of the 18 sessions, 12 participants were randomly assigned into the three main treatments; in the remaining 10 sessions, 16 participants were randomly assigned into the four treatments (three main treatments and one supplementary treatment). In total, we have 18, 18, 18 and 10 groups in "High Start," "Big Jump," "Gradualism," and "High Show-up Fee" treatments, respectively.

In the first stage, the thresholds over 12 periods are shown in Figure 1: for the "High Start" treatment, the thresholds are always at the highest level, which is 14 for 16 sessions and 12 for 2 sessions;[12] for the "Big Jump" treatment, they are 2 for the first 6 periods and set at the highest threshold for the next 6 periods; for the "Gradualism" treatment, they increase from 2 to 12 with a step of 2 for the first 6 periods and fix at the highest threshold for the next 6 periods. The show-up fees for these three treatments were 400 points for each individual. The "High Show-up Fee" treatment is the same with the "High Start" treatment, except that the show-up fee is 480 instead

[11] The *yuan/USD* exchange rate was about 6.7.
[12] We calibrated the highest threshold level using 12 and 14, and finally decided to choose 14 in most sessions. To make full use of the samples, in the following analysis we pool all 18 sessions together.

of 400 points. The extra 80 points are sufficient to capture potential earning differences accumulated from periods 1-6 and thus to isolate the effect generated by an income effect by comparing the “High Show-up Fee” vs. the “High Start” treatment (we discuss this in detail in Section 4.)

When subjects enter the second stage of the game, they are randomly reshuffled into groups of four. New group members may not necessarily come from the same treatment in the first stage; this rule is made common knowledge. Within the second stage, group compositions are fixed, and thresholds are all set at the highest threshold for all periods and all groups, i.e., those treated in different treatments in the first stage face the same threshold in each period of the second stage.

The information structure is as follows.

Participants know that there are 2 stages, but not the exact number of periods in each stage. Instead, they are told that the experiment will last for 30 minutes to one hour, including the time for signing up, reading of instructions, a quiz designed to make sure that they understand the experimental rule, and final payment. We chose this design for two reasons. First, it is to eliminate backward induction in theory and a potential end-of-game effect, although backward induction is not well supported empirically in the literature, and we argue that in minimum-effort coordination games the end-of-game effect does not exist or will be just minor (see Section 6.6). Second, in many cases of the real world, people do not know the exact number of cooperation and coordination opportunities (e.g., how many times they will meet each other, how many projects they will have).

At the beginning of each period they know the threshold of the current period but not those of future periods. In many cases of the real world, the levels of future interactions are unknown *ex ante*.

At the end of each period they know whether all 4 group members (including himself or herself) gave the required points of that period, but not the total number of group members who gave (if fewer than 4 members gave). This is consistent with the literature of minimum-effort coordination games (e.g., Van Huyck et al., 1990), in which the only common historical data available to the participants is the minimum.[13] We can find further supports for this design from the literature of contract theory, in which imperfect observation of efforts is common. By

[13] In our experiment, if all 4 members give the required amount, then the minimum is that amount (or coded in a binary way, “1”); otherwise the minimum is 0.

adopting this design we can also increase the difficulty of coordination given other aspects of the experiment.

Communications across players is not allowed. We chose this design for two reasons. First, communication is often impossible or ineffective in the real world. Second, this design makes coordination more difficult.[14]

At the start of the second stage, each player is notified that they enter a new random group. At the end of each stage, each player is told how many points she has accumulated to date.

At the end of the experiment, we asked participants to complete a brief survey. The survey collected information on age, gender, nationality, education level, concentration at school, working status and income, in addition to eliciting risk preferences over lotteries (see Appendix).

4. Results

In this section we present our findings in summary tables, figures and regressions. We analyze the effect of treatments on the following three outcome variables per period: whether a group coordinates successfully, whether an individual contributes, and the individual's earning.

Table 1 contains basic summary characteristics of the participant pool. The participants are generally young with an average age around 22, since 91% of them are college or graduate students. 41% are male. 12% are (or were) majored in economics, 16% in other social sciences, 27% in business, and the remaining in other disciplines. The average individual annual income in the year of 2009 falls between 5,000 *yuan* and 10,000 *yuan*.

Table 2 checks how randomization worked in assigning participants into different treatments. The default category is the "Gradualism" treatment and the regressions do not have other control variables, so the constant term indicates the mean values of dependent variables for the "Gradualism" treatment. There are no significant differences in subjects' characteristics across treatments, except that participants in the "Big Jump" treatment have higher self-reported family economic status, that those in the "High Start" and "High Show-up Fee" treatments have higher risk aversion indexes, and that those in the "High Start" treatment are more likely to be students. This shows that randomization did very well, although not perfectly.

[14] Ostrom (2010) summarizes that communication improves cooperation. Charness (2000) shows that communication helps coordination in small groups, while Weber et al. (2001) and Chaudhuri et al. (2009) find that large group coordination is still difficult even with communication.

Table 3 contains the summary of designs and performances of all four treatments in the first stage. Clearly, the "Gradualism" treatment has better performances of coordination for periods 7-12 in the first stage, which shows that gradualism does promote coordination at a high threshold level. For example, in period 7, 66.7% of "Gradualism" groups have coordinated successfully, while the success rates of "High Start," "Big Jump," and "High Show-up Fee" are only 16.7%, 33.3% and 30%, respectively. Once successful coordination has been established in period 7, generally it can be maintained subsequently (actually, only one "Gradualism" group failed to do this.)

In Table 4, Panel A provides summary statistics of the major outcome variables in the first stage: fraction of group members contributing, average number of group members contributing, success rates of groups (whether all 4 group members contribute) and individual earning per period. We also provide summaries of individual earning up to period 6 (i.e., earning accumulated from period 1 to 6, not including the show-up fee) for each treatment. On average, participants in the "Gradualism" treatment have the highest earning up to period 6 than all other three treatments. But the differences in means (and medians) are much smaller than 80 points, the difference of show-up fee between the "High Show-up Fee" treatment and other three treatments. For example, the average earning by period 6 is 143.94 for the "Gradualism" treatment, and 112.42 for the "High Start" treatment; and the median is 162 and 106, respectively. This shows that a show-up fee difference of 80 points between the "High Start" and "High Show-up Fee" treatments is large enough to capture the potential income differences at the start of period 7 between "High Start," "Big Jump," and "Gradualism" treatments. Actually, if we add up the show-up fee, participants in the "Gradualism" treatment earn less than those in the "High Show-up Fee" treatment on average by period 6. So if we still find the "Gradualism" treatment has better performance than the "High Show-up Fee" treatment in periods 7 to 12 in stage 1, then the difference should not be driven by a potential income effect from the first 6 periods.[15] Panel B shows the overall performance in both stages for all participants. An average participant ends the experiment with an average around 857 points, and earns 21-22 *yuan* (around 3 dollars), which affords ordinary meals for 1-2 days on campus.

[15] Real world income may also affect individuals' decisions. By randomly assigning participants into various treatments, we rule out the possibility that the differences of performance are due to real world income.

In the figures and regressions below, we break out these summary statistics into performance by period to gain a clearer understanding of the different effects of treatments on behaviors and outcomes.

Figure 2 shows the average group success rates in each period for four different treatments respectively, and clearly provides us a stark result: gradualism does help build coordination. A group is considered as successful in coordination if all 4 members give the required amount in that period. In period 1, most “Big Jump” and “Gradualism” groups coordinate successfully at the low threshold, while only 3 out of 18 (or 10) of the “High Start” (or “High Show-up Fee”) groups are successful at the high threshold. Interestingly, there are both 5 (out of 18) groups in “Big Jump” and “Gradualism” treatments respectively, which are unsuccessful at the low threshold in period 1 when the threshold is only 2.

A large gap between “Gradualism” and “Big Jump” treatments emerges in period 7 when the threshold increases significantly from 2 to 14 (or 12) for the “Big Jump” treatment. Both treatments have high success rates of approximately 70% for the first 6 periods. But the success rate of the “Big Jump” treatment decreases sharply to only 33.3% from period 7 (i.e., more than half of the previously successful “Big Jump” groups now fail), while that of the “Gradualism” groups remains at a high level of above 60% (only one previous successful “Gradualism” group fails during the last 6 periods.) The “High Show-up Fee” treatment has a success rate of 30% for all 12 periods, and about 16.7% of “High Start” groups succeed for almost all periods.

Interestingly, once a group has failed in coordination, it almost never becomes successful subsequently. The only two exceptions are one “Big Jump” group that fails in period 5 but succeeds in period 6, and one “High Start” group that fails in period 1 but succeeds in period 2. This finding is consistent with Weber et al. (2001) and Weber (2006), in which they find once groups have reached an inefficient outcome, they are unable subsequently to reach a more efficient outcome.[16]

Conversely, once a group has succeeded, it almost always remains successful unless the threshold increases sharply (i.e., from period 6 to period 7 for “Big Jump” treatment). Conditional on successful coordination in period 1, “High Start,” “Gradualism” and “High Show-up Fee” groups almost are all successful in following periods (except that there are two “Gradualism” groups which failed to maintain it until the end of stage 1), but since the “High

[16] However, changing incentives can improve coordination (Berninghaus & Ehrhart, 1998; Bornstein, Gneezy & Nagel, 2002; Brandts & Cooper, 2006b).

Start” and “High Show-up Fee” treatments have a much lower success rate in period 1 than the “Gradualism” treatment, on average they perform much worse than the “Gradualism” treatment at a high threshold.

Figure 3 shows the fractions of individuals contributing over periods in the first stage for the four treatments. “High Start” and “High Show-up Fee” treatments start with average contribution rates above 60%, but decrease quickly over the initial periods and end at about 20% and 40%, respectively. “Big Jump” and “Gradualism” treatments start with a high average contribution rate above 90%, decrease slightly over the first 6 periods to a rate of about 80%. However, the contribution rate decreases sharply from period 6 to 7 for the “Big Jump” treatment (when threshold is sharply raised from 2 to 14), while that for the “Gradualism” treatment remains high. Although the contribution rates of all treatments generally decrease over time (even for the “Gradualism” treatment), the results are consistent with Figure 2: except for the decrease of contribution rate from period 6 to 7 for “Big Jump” groups, the decrease of contribution rates is almost all caused by those who give up cooperating when their cooperative actions at previous periods have not been rewarded because of coordination failure, while those have succeeded keep cooperating. The differences among the four treatments in Figure 2 (regarding group success rate) is much more stark than those in Figure 3 (regarding fractions of individuals contributing), since you need all 4 members to give at the same time to make the group coordination successful, which is much more difficult than asking only one person to give. This is why coordinating at the same pace is so important, and gradualism helps address this challenge significantly.

Figure 4 shows the average individual earning in each period. The results also show that the “Gradualism” treatment works best.

Table 5 shows the formal regression results for periods 7-12 in the first stage, when all treatment groups face the same high threshold level. The default category is the “Gradualism” treatment and the regressions do not have other control variables, so the constant terms indicate the mean values of dependent variables for the “Gradualism” treatment. Dependent variables are a dummy indicating whether an individual gives or not in period 7 and 12 (in Column 1 and 2, respectively), an individual’s earning in period 7 and 12 (in Column 3 and 4), and a dummy indicating whether the group has successful coordination: we consider period 7 in Column 5, period 12 in Column 6, and since this dummy is our main variable of interest, we further examine all six periods from period 7 to 12 in Column 7. All standard errors are clustered at the appropriate level. It shows that the differences between the "Gradualism" treatment and other treatments are mostly large and significant. F-test (unreported) shows that the differences between "High Start" and "High Show-up Fee" treatments are statistically insignificant, although it might be due to a relatively small sample. Moreover, the large and significant differences between "Gradualism" and "High Show-up Fee" treatments suggest that the advantage of gradualism is not driven by the income effect in the laboratory. We also employ probit and logit specifications when the dependent variable is whether an individual contributes or whether a group reaches successful coordination, and the results are very similar with those OLS results in Table 5 (Column (1), (2), and (5)-(7)).[17] When the dependent variables are contribution and earning, which are at the individual level, additional regressions[18] with survey controls show very similar results with those in Column (1)-(4).[19]

In Table 6 we examine whether the treatments in the first stage have effects on behaviors and outcomes in the second stage when participants enter a new group. Note that everyone knows that the new group members can come from any of the four treatments in the first stage. Since the group formations are different from those in the first stage, we only look at individuals' contribution and earning in each period of the second stage, rather than the group-level coordination results. Interestingly, those in the "Gradualism" treatment in the first stage are more likely to contribute in the first period of the second stage. It is important to keep in mind that group assignment in stage 2 is a new randomization, and participants know that they are playing in a new group. Thus, this finding is indicative of a new effect, namely an inter-stage effect of having been treated in a gradualism environment, which is similar with the finding of Bohnet and Huck (2004).[20] However, this effect disappears over the course of the second stage, suggesting a learning process where the behaviors of the different treatment groups converge as they observe the play of their new group members. For example, those in the "Gradualism" treatment find their new group members are not as cooperative as those in the first stage, thus becoming less willing to contribute in following periods.

---

[17] Available upon request.

[18] Available upon request.

[19] There may be a concern that economics and business students play differently with other students. We address this concern by examining actions in the first period, as well as actions in the second period conditional on the coordination outcome of the first period, and find no differences between economics/business students and other students.

[20] They study a trust game and find that trustees are more trustworthy in the stranger environment in the second stage after having been exposed to a partner in the first stage.

These findings suggest that the gradualism setting can induce a long-run cooperative behavior, most likely through a history of successful cooperation and an increased level of trust.[21] However, when they find their trust does not get rewards in a new environment, they tend to become less cooperative. This shows an externality of coordination building (or collapse) across different social groups.

5. Toward a Theory: Why Gradualism Works?

5.1 Belief-based Learning

Here we use a simple model with some assumptions to capture the most important feature of gradualism and explain why gradualism works. The main aspects of our model are belief updating with bounded rationality (especially level-1 reasoning), and standard self-interested preferences with risk aversion. Non-traditional (e.g., other-regarding) preferences, as well as other potential learning features which might be consistent with our results, are not incorporated into our model, since we do not need them to explain our main experimental results. A general learning model of dynamic games is beyond the scope of this paper.

There are two main types of learning models: belief-based learning models and reinforcement learning models, which are both incorporated in a general model of experience-weighted attraction (EWA) learning model (Camerer & Ho, 1999).[22] Belief-based learning models assume that players keep track of the history of previous play by other players and form some belief about what others will do in the future based on past observations. Then they choose a best-response that maximizes their expected payoffs given the beliefs they formed. In contrast, reinforcement models assume that players do not have beliefs about what other players will do, and the propensity to choose a strategy depends in some way on its past payoffs rather than the history of play that created those payoffs. Reinforcement learning models do not apply well for our experimental results, since many participants in the "Big Jump" treatment who benefit from

[21] It is possible that those in the "Gradualism" treatment contribute significantly more in first period of second stage just because of a wealth effect coming from their higher average earning from the preceding stage. Under the assumption of decreasing absolute risk aversion, this greater wealth could produce a greater willingness to contribute. While possible, we do believe that any income effect would be small, certainly not enough to generate the observed discrepancy.

[22] There are other models such as the quantile response equilibrium (QRE) model (McKelvey & Palfrey, 1995). As suggested by Brandts and Cooper (2006b), given the strong dynamics and history dependence in our experimental data, static models such as QRE are not good candidates. Thus we focus on learning models in which some players have bounded rationality and learn how to respond from their experiences.

cooperating by period 6 in stage 1 give up the passive strategy of continuous cooperation from period 7. Thus we propose a belief-based learning model, presented below, which matches our experimental results very well.

Crawford (1995) and Weber (2005, 2006) adopt an adaptive dynamics model, in which every participant employs a linear latent strategy weighting own action and the minimum action of group members in the previous period.[23] We do not use their model for two reasons. First, in their experiments, there are 7 action choices in each period, making a linear continuous latent strategy more relevant; while in our experiment, there are only 2 choices in each period, which is "more discrete." Second, in their experiments, the game is repeated identically for all periods, while in our experiment, the threshold changes over the first 6 periods of stage 1 for the "Gradualism" treatment, and from period 6 to 7 for the "Big Jump" treatment, so a linear weighting of previous actions cannot well predict and guide actions in the current period.

Our simple belief-based learning model borrows the form of that in Van Huyck et al. (1990), with features of a prior belief about opponents' actions and a belief updating process based on the coordination result in the previous period, but differs from their model in two main aspects. First, in our model players' choice set in each period is binary (either 0 or the required amount, i.e., the threshold) and thus in each period players play an *n*-person symmetric coordination game with two pure strategies. Second, players in our model are restricted to level-1 thinking on their beliefs and are supposed to be myopic, that is, on one hand, players know their opponents' choice rules and can form expectations on opponents' chosen strategies based on their beliefs, but they do not know that their opponents know the choice rules of themselves and so on; on the other hand, when playing the coordination game for one period, players never know whether the game would be repeated in the future, needless to say that they do not know the future path of thresholds as well. Because of the restriction of level-1 thinking players cannot conduct deductive thinking and thus belief-updating across periods makes sense, and due to their myopia players do not conduct backward induction to solve the game so that we can focus on the one-period game and then analyze the process of belief dynamics.

---

[23] Specifically, in their models, play i's latent strategy in period t is given by the following formula: $a_{i,t} = (1-\beta)x_{i,t-1} + \beta y_{t-1} + \varepsilon_{i,t}$, where $x_{i,t-1}$ is i's discrete action in period $t-1$, y_{t-1} is the minimum action of group members in period $t-1$, and $0 < \beta < 1$. The $\varepsilon_{i,t}$ are distributed normally, with mean zero and variance σ_t^2. Play i's discrete action in period t ($x_{i,t}$) is determined by $a_{i,t}$ (e.g., by rounding the latent variable to the nearest discrete action).

Also, as the strategy set in each period is binary in our experiment, we further adopt a concept of “latent action.” Latent action refers to the amount of points one would like to give if there is no binary restriction on the choice set. In other words, latent action is the largest possible amount of points one would like to give in any period, and we call it “reserved threshold” in our context for clearance. If the reserved threshold is greater than the set threshold in the current period, one would give; otherwise she would not give. If one believes all opponents’ reserved thresholds are equal to or greater than the current threshold, she would give. It is important to point out that, from the perspective of a myopic player with level-1 thinking, her own reserved threshold is determined by her choice rule that relies on her belief of opponents, which should be past-dependent, and thus one’s own reserved threshold may change over periods; however, in her mind her opponents’ reserved thresholds must be static because she does not know that her opponents indeed follow the same process to update their beliefs as well as reserved thresholds, and she can only update her belief on these static reserved thresholds of her opponents. The convenient assumption of level-1 thinking play a key role here: agents take their opponents’ reserved thresholds as constant but unknown parameters, which enables a gradually refining but consistent belief updating process on the distribution of these parameters. If we allow level-2 thinking, then there will not be constant reserved thresholds, and the model will become much more complicated.

Suppose there are I players who are all risk averse. Each player i’s belief about every opponent j’s reserved threshold is B_j^i, which means that she believes opponent j will only give when the set threshold in the current period Th_t is less than B_j^i, otherwise that opponent will not give. Each player i’s belief about opponent j’s reserved threshold B_j^i follows a cumulated distribution function $F_{j,t}^i()$, $j \neq i$, whose support is $[LB_t^i, UB_t^i]$. At the start of period 1, we suppose the supports to $F_{j,1}^i()$ for all i, j are $[0, UB]$, where UB is the highest possible level of reserved threshold,[24] i.e., the highest possible number of points one would like to give in players’ minds, which is not necessarily less than the endowment in each period. As discussed before, each player i is restricted to level-1 thinking, that is, she does not know that her opponents know her own cumulated distribution function $F_{j,t}^i()$, so that she does not take her opponents’ response to her own choice into account when she makes her own decision.

[24] Setting $UB \to \infty$ will not change our main results, but will make the analysis of optimal threshold path more complicated.

Now we study under what conditions a player would like to give in period t when the current threshold is Th_t, and in what manners she updates her belief. From the perspective of player i, she will give if and only if she believes the probability that all her opponents will give exceeds a certain value $\theta_i(Th_t)$, which denotes her risk attitude. In our framework of coordination game, under the assumption of risk aversion, it is easy to see that $0.5 < \theta_i(Th_t) < 1$ for $Th_t > 0$ and $\partial\theta_i(Th_t)/\partial Th_t > 0$. In this sense, assuming independent distributions of $F_{j,t}^i()$ for all j, as each player is restricted to level-1 thinking and is myopic, player i will give in period t if and only if:

$$\Pr(\bigcap_{j\neq i} B_j^i \geq Th_t) = \prod_{j\neq i}(1 - F_{j,\,t}^{\,i}(Th_t)) \geq \theta_i(Th_t) \tag{1.1}$$

For convenience, we define a new concept of "social reserved threshold," the minimum of the "reserved thresholds" of each player' opponents, which only exists in her belief. Specifically, player i's social reserved threshold is B^i that implies she believes that all her opponents will give only when the set threshold in the current period Th_t is lower than B^i, otherwise at least one opponent will not give. Since B_j^i is static in player i's mind, B^i must be also static in her mind, and thus a cumulated distribution function of B^i is well defined. From (1.1) we know that the cumulated distribution function of B^i can be expressed as follows:

$$F_t^{\,i}(Th_t) = 1 - \Pr^i(\bigcap_{j\neq i} B_j^i \geq Th_t) = 1 - \prod_{j\neq i}(1 - F_{j,t}^{\,i}(Th_t)) \tag{1.2}$$

It is easy to show that the supports of $F_1^{\,i}()$ for all i are still $[0, UB]$.

Therefore, player i will give in period t if and only if:

$$1 - F_t^{\,i}(Th_t) \geq \theta_i(Th_t) \tag{1.3}$$

When Th_1 is too high such that (1.3) does not hold, i chooses not to give; otherwise she gives. This explains why "Gradualism" and "Big Jump" groups coordinate much better than "High Start" and "High Show-up Fee" groups in period 1 of stage 1.

Next, we study the manner in which each player updates her belief. In a given period t, if (1.3) is violated, in other words, if player i choose not to give, then she cannot know whether or not her opponents have given due to the information structure, and hence she cannot update her belief on B^i at all. However, if (1.3) is satisfied (i.e., player i chooses to give), there must be the following two cases.

The belief updating process is as follows. It is worth noting that the belief updating process does not follow the traditional Bayesian rule. As discussed above, the social reserved threshold is a determinant in a player's mind, and thus it is regarded as a random variable only in the sense that a player does not know its exact value. Specifically, a player updates the support of distribution of social reserved threshold based on the outcome in each period, which is characterized in the following cases.

Case 1: all her opponents have given and thus the coordination succeeds. As player i is restricted to level-1 thinking and B^i is static in her mind, she now surely knows that $1-F_t^i(Th_t)=1$, in other words, $F_t^i(Th_t)=0$. This implies that player i can naturally update the support of $F_t^i()$ from $[LB_t^i, UB_t^i]$ to $[Th_t, UB_t^i]$.

Case 2: at least one of her opponents has not given and thus the coordination fails. Similarly, she now surely knows that $1-F_t^i(Th_t)=0$, in other words, $F_t^i(Th_t)=1$. This implies that player i can naturally update the support of $F_t^i()$ from $[LB_t^i, UB_t^i]$ to $[LB_t^i, Th_t]$.

Conditional on successful coordination at t, for $Th_{t+1} \geq Th_t$, if Th_{t+1} is close enough to Th_t, and far enough from UB_t^i, then (1.3) still holds, so i would continue cooperating until (1.3) is violated, that is why gradualism works. This idea will be formally presented in Proposition 1 below. However, if $Th_{t+1} - LB_t^i$ is too large, than there is a high probability that at least one of opponents' "reserved thresholds" is lower than Th_{t+1}, so (1) does not hold anymore, and i will choose not to give. This explains why the "Gradualism" treatment performs well, while the "Big Jump" treatment largely fails when the threshold jumps from a low 2 to a high 14.

Following the belief-updating rule, we can show the following lemma that is important in characterizing our final results.

Lemma 1 (Belief Update) For any player i, there is $[LB_{t+1}^i, UB_{t+1}^i] \subseteq [LB_t^i, UB_t^i]$ for any period t.

Proof: See Appendix A.

Based on our decision rule and belief updating rule for each player that can be naturally interpreted under the assumptions of binary choices, level-1 thinking, and myopia, we can eventually get the following results regarding our experimental findings.

Lemma 2 (Gradualism) There always exists a threshold $Th_t > \mathrm{Min}_i\, LB_t^i$ to guarantee successful coordination if $UB_t^i > LB_t^i$ for all i for any period t.

Proof: See Appendix A.

Proposition 1 (Gradualism) There exists an infinite path of threshold $\{Th_t \mid Th_{t+1} \geq Th_t\}$ to guarantee successful coordination from period 1.

Proof: See Appendix A.

Proposition 2 (Memory of Failure) In the path of the game, if there are and only are fewer than *I* players who have given in any period *t* with a threshold Th_t, the future threshold that can make coordination successful must be $Th_\tau < Th_t$ for all $\tau > t$.

Proof: See Appendix A.

Our model can also explain the main results of Weber (2006), who studies gradualism in organizational growth. Staring with a large group size makes one believe that there is a low probability that all partners will choose a high action, so it is less likely to reach a high equilibrium. On the contrary, starting with a small group size will give one a high belief that all partners will choose a high action.[25] And if a new group member knows the previous coordination history of the group, she will have a high belief that all partners will choose a high action (and other partners, knowing that the entrant knows the history, will also have a high belief that the entrant will choose a high action), so they can maintain a successful coordination when the group enlarges. However, if the entrant cannot observe the history, a high trust among the entrant and incumbents cannot be established, leading to a coordination failure when the group expands.

The model can also explain the main results of other minimum-effort coordination games in the literature, e.g., Van Huyck et al. (1990), but not some minor results such as the "overshooting" phenomenon that some participants play below the minimum of the preceding period (see pp 241, Section V of their paper). Incorporating non-standard preferences into our simple model can explain these minor results.

Although we have shown (and will further show below) that belief-based learning is the most plausible reason of the success of gradualism, we do not formally test it in this study. Some recent papers confirm that the majority of, although not all, participants behave consistently with their beliefs (e.g., Nyarko & Schotter, 2002; Costa-Gomes & Weizsäcker, 2008; Rey-Biel, 2009). Direct belief elicitation methods, especially in an incentive-compatible way, become

[25] Note that in formula (1.1), the larger the group size, the smaller the probability one perceives that all partners will give at least a certain amount.

increasingly popular in experimental economics (e.g., Offerman et al., 1996, 2001; Nyarko & Schotter, 2002; Costa-Gomes & Weizsäcker, 2008; Rey-Biel, 2009; Hyndman et al., 2009; Tingley & Wang, 2010). We finally gave up doing this because our main purpose of this study is to cleanly test whether gradualism works, and we had a concern that asking belief questions in each period would have undesirable effects on participants' playing which may contaminate our results. The literature supports our concern: several other studies (Costa-Gomes & Weizsäcker, 2008; and especially, Rutström & Wilcox, 2004, 2009) do show that eliciting players' beliefs induce more sophistication and higher-order rationality, thus affecting their actions.

5.2 Other Explanations

5.2.1 Reinforcement Learning

As discussed above, a naïve reinforcement learning model predicts that participants who benefit from cooperating by period 6 in the first stage should adopt the passive strategy of continuous cooperation from period 7. But this contracts the findings regarding the "Big Jump" treatment.

5.2.2 Conditional Cooperation

Conditional cooperation assumes that players are willing to cooperate if others cooperate as well (Fehr & Gächter, 2000). Based on the same reason as above, a simple story of conditional cooperation does not work, either. A more sophisticated version of conditional cooperation may fit the results better, as described below.

Some participants might reciprocate by contributing in response to the cooperative actions of others. Such reciprocity also opens up the possibility for strategic behavior, where a person might contribute in early periods to encourage others to do the same. More participants may choose this strategy and try an "initial venture" when thresholds are lower than when thresholds are higher at the beginning. This explains why the "Gradualism" and "Big Jump" treatments work better than the "High Start" treatment.

However, for the theory of conditional cooperation to better explain why the "Big Jump" treatment does not work as well as the "Gradualism" treatment, some refinements of conditional cooperation is desired: conditional cooperation may not be absolute and unlimited, and may apply only when the thresholds do not have big changes over time.

Unlimited (or strong) conditional cooperation:

$$A_{i,t} = \begin{cases} C, & if \; A_{-i,t-1} = C \\ D, & if \; A_{-i,t-1} = D \end{cases}$$

Where C and D represent "cooperate" ("give") and "deviate" ("not give"), respectively.[26]

Limited (or weak) conditional cooperation:

$$A_{i,t} = \begin{cases} C, & if \; A_{-i,t-1} = C \; and \; Th_t - Th_{t-1} \le \delta_1 \\ D, & if \; A_{-i,t-1} = D \; and \; Th_t - Th_{t-1} \ge -\delta_2 \\ C \text{ or } D, & \text{otherwise} \end{cases}$$

Where δ_1 and δ_2 are positive constants. The decision rule is as follows: if all partners cooperated at $t-1$, then cooperate if the threshold increases less than a given amount, keeps the same, or decreases; if at least one partner deviated at $t-1$, then deviate if the threshold decrease less than a given amount, keeps the same, or increases; otherwise "conditional cooperation" is not used, and the self-regarding preference dominates.

Strong conditional cooperation may apply for some participants, but for others only the weak one applies. When the threshold increases by a large amaount (say, threshold increases from 2 to 14 in our experiment), those with strong conditional cooperation contribute if their partners all contributed in the previous period (when the threshold was at a low level of 2), while those with weak conditional cooperation may not keep contributing even if their partners all contributed in the previous period. This explains why the "Big Jump" treatment does NOT work well in building coordination at a high level of threshold. In contrast, for the "Gradualism" treatment, the threshold changes slowly, which makes those with weak conditional cooperation still willing to contribute when the threshold becomes higher and higher.

5.2.3 Inertia

Psychological inertia[27] refers to indisposition to change. It is not successful at explaining why "Big Jump" groups cannot maintain successful coordination when the threshold increases quickly.

5.2.4 Preference for Consistency

[26] According to our experimental design, when an individual does not give in a certain period, she cannot know whether all other 3 members give or not even after the outcome is revealed.

[27] Inertia has been studied in behavioral public finance issues such as individual saving behavior (e.g., Madrian & Shea, 2001; Thaler & Benartzi, 2004).

A similar psychological theory is preference for consistency (Cialdini et al., 1995), which means that once people make a choice or take a stand, they encounter personal and interpersonal pressures to behave consistently with that commitment. Similar with inertia, it cannot explain why “Big Jump” groups fail to maintain successful coordination.

5.2.5 Limited Attention and Bounded Awareness

Since the threshold step in the “Gradualism” treatment is small, is the gradualism effect due to limited attention (for a review, see DellaVigna, 2009), or bounded awareness (Gino & Bazerman, 2009)? Our answer is no. First, participants were doing the single coordination task in our experiment, so there was no other information or task which would distract their “limited attention.” Second, although the absolute increment of the threshold in the gradualism treatment is small (2 points per period), it is significantly larger than those in the experiments of Gino and Bazerman (2009) and Offerman and der Veen (2010), which adopt much more gradual settings. Actually, in our experiment, the step of 2 is quite a large increase, ranging between 100% of the threshold size in period 1 and 16.7% of threshold size in period 6.

6. Extensions

In this experiment we focus on certain settings of gradualism. Below we discuss potential extensions for future studies.

6.1 Information Structure

In this paper we choose a certain information structure that parallels many cases in the real world, as discussed in Section 3. For example, participants do not know the exact number of periods; they only know the current threshold but not future thresholds;[28] they only know whether all group members (including himself or herself) gave, but not the total number of group members who gave (if not all members gave); they only know the dynamics of threshold in their

[28] For the “High Start” groups, some participants may want to wait the threshold to decline before giving, so they do not give in the first period, but they then find that the threshold is always 14 and that they have missed the chance of future successful coordination. If they know *ex ante* that the thresholds are always 14 and that coordination in the first period is pivotal, more of them may cooperate in the first period, thus having a higher rate of successful coordination in that period and the following periods. The same argument may apply for the “Big Jump” groups in period 7 when the threshold jumps to the high level. Thus this new information structure may shrink the role of gradualism. However, even they know the threshold dynamics *ex ante*, participants may not be so rational and sophisticated to try cooperating in the first period. Thus how the effect of gradualism changes is still a pending empirical question.

own group (*ex post*), but not that of other groups (even *ex post*), although they know that other groups' paths may be different from theirs. How the results may change with a different information structure, which may fit other cases in the real world, deserves future exploration.

6.2 Free Riding (The Case of Standard Discrete Public Goods)

In this experiment we require all group members to contribute to get the benefits. What if we require fewer than all group members, say, 3 out of 4, to contribute to receive the benefits of public goods (i.e., what if we introduce the possibility of free riding as in standard public goods games?) The theoretical framework in Section 5.1,[29] plus some elements of social preference, can still work here. Assuming m is the minimum required number of contributing group members to receive the benefit, a self-interested individual will give if and only if she has a high belief that exactly $m-1$ of his or her partners will give. But according to the literature on public goods game, even one believes that there is a high probability that at least m partners will give, she may still give, i.e., some may choose not to free ride. Thus we may need to add some elements of social preference (e.g., conditional cooperation) into the simple model in Section 5.1 for the purpose of the public goods case. Still, strategy uncertainty about others' actions also exists in this discrete public goods game, and gradualism helps reduce strategy uncertainty at the high threshold level, so we postulate that gradualism still works in discrete public goods game with a reasonable group size and a reasonable payoff structure. This hypothesis will be tested in an accompanying paper coming soon.

6.3 Group Size

Coordination becomes difficult when group size increases in one-shot minimum-effort coordination games (Weber, 2006). The effect of group size can also be incorporated into our theoretical model: holding others constant, a larger group size reduces one's belief that all other group members will give, as shown by formula (1.1) in Section 5.1. A group size of 4 is not too

[29] The model in Section 5.1 only needs self-interested preference (with risk aversion). This is largely due to the payoff structure of minimum-effort (weak-link) coordination game: A self-interest individual will give if and only if she believes that there is a high probability that all partners will give. Although this strategy is also predicted by conditional cooperation, theoretically we do not need social preference like conditional cooperation to lead to this strategy, so it is hard to distinguish the motivation of conditional cooperation from a self-interest motivation in this case of coordination game. However, as discussed in Section 5.2, a simple version of conditional cooperation cannot explain the results regarding "Big Jump" groups, although we still cannot rule out a more sophisticated version of conditional cooperation.

large, so successful coordination at a high threshold of 14 is possible for those treated in a gradualism process, as shown in our results. A slightly larger group size makes coordination more difficult and gradualism more useful. But if the group size is too large, even gradualism may not help. On the other side, a smaller group size (2 or 3) makes coordination easier and gradualism less useful. In sum, the effect of gradualism will show an inverted-U relationship with group size. This hypothesis can be tested in the future.

6.4 Highest Threshold Level

In this experiment, the highest threshold level is 14,[30] given an endowment of 20 per period. How will gradualism perform when we change this level? Similar with group size, a higher (lower) threshold level indicates a more (less) difficult coordination task. This can also be shown by formula (1.1) in Section 5.1. So based on the reason as Section 6.3, we conjecture that the effect of gradualism will show an inverted-U relationship with the highest threshold level.[31]

6.5 Rebuilding Coordination after Coordination Collapses?

This paper tests whether gradualism builds coordination for a newly formed group. Is gradualism also helpful for groups in which coordination and trust have already collapsed? For example, coordination at a high threshold may collapse due to the treatment of "High Start" or "Big Jump," or other reasons (e.g., conflicts between two countries or ethnic groups).

6.6 Is There an End-of-game Effect in Coordination Games?

The literature on multi-period prisoners' dilemma games and public goods games shows an end-of-game effect. Does such an effect exist in coordination games, thus making gradualism less useful in the final period? Our belief-based learning model in Section 5 predicts no end-of-game effect. Unfortunately, our experimental design in this paper does not allow us to test this,[32] since participants are not informed when the last period comes. However, we conjecture that the end-of-game effect will only be minor or not exist. The reason is as follows. In public goods games and prisoners' dilemma games, a best response for self-interested individuals is deviating (or free riding) given that the partner(s) is (are) cooperating (contributing), so a self-interested individual

[30] In 2 out of the 18 sessions, the highest threshold level is 12.
[31] The turning point may not be necessarily smaller than the endowment per period, 20.
[32] In this experiment, participants do not know it is the end point when the last period comes, which fits many cases in the real world.

will try to exploit his or her partner(s) by cooperating before the end period, thus inducing the partner(s) to cooperate in the end period, while she gains by deviating in the end period.[33] On the contrary, in minimum-effort coordination games, the best response for self-interested individuals to a cooperative action is cooperating, so deviating in the last period makes no sense as long as you have a high belief about others' cooperative behavior through successful coordination at the same or a close level of threshold. Thus the end-of-game effect should be much less relevant for minimum-effort coordination games.[34] Empirically examining whether such an effect exists and affects the role of gradualism can be done in a future study.

7. Conclusion

The findings in this paper suggest that gradualism -- defined as increasing step-by-step the threshold level required for coordination -- can serve as a powerful mechanism for achieving socially optimal outcomes. In a laboratory setting, gradualism significantly outperforms alternative paths to coordinated behavior. We also find an externality of coordination building (or collapse) across different social groups. Those treated in the gradualism setting are more likely to cooperate upon entering a new environment than those treated differently. However, when they find their cooperation does not get rewards in a new environment, they tend to become less cooperative as well.

We propose a simple theoretical framework of belief-based learning to explain why gradualism works. This framework is consistent with results of our experiment and of Weber (2005, 2006). Other potential explanations of the gradualism mechanism, such as naïve reinforcement learning, standard conditional cooperation, inertia, preference for consistency, and limited attention, do a poor job of explaining the main findings. A sophisticated version of conditional cooperation may better explain the results. Due to the weakest-link feature of our

[33] As long as a self-regarding component of one's preference dominates, this statement holds even if we allow certain social preferences.

[34] As discussed in the introduction section, the general literature on minimum-effort coordination games, employing a complex payoff matrix of 7×7, finds that even with a small group size like 3 (more apparently with a group size of 6), the game has been largely converged to the least efficient outcome from period 5, and the groups are then trapped in this low equilibrium (Weber, 2006). However, this is different from the end-of-game effect in prisoners' dilemma games and public goods games. Moreover, it should be noted that this result is due to the "freedom" of choosing from 7 actions in each period, as discussed in the introduction. Our experiment differs from these studies, and further contributes to the literature in the sense that by limiting choices in each period and by a gradualism institution, convergence towards an inefficient outcome can be avoided, and the social optimal can be attained.

experiment, since the best self-interested response given others cooperating is to cooperate as well, actually we do not need conditional cooperation to explain our results.[35] But this does not rule out a sophisticated version of conditional cooperation. Further distinguishing the belief-based learning (with self-interested preference), conditional cooperation and other potential theoretical explanations is a topic for future research.

Although we show that belief-based learning is the most likely mechanism for the success of gradualism, we do not formally test it in this study. Although imperfect, we may still use the incentive-compatible belief elicitation method to suggestively show whether belief updating is occurring and whether agents play according to their beliefs. We can also test whether gradualism helps most for people who do not know their group members than for those who are friends of each other before the experiment, since the latter group already has a higher trust level even without gradualism.

In this experiment we focus on certain settings of gradualism. There are many ways to extend this study in the future. Allowing communication and free riding, changing the group size and the highest threshold level, adopting a different information structure and a more complex dynamic path of thresholds, are all important dimensions for future study. It will be also interesting to examine whether gradualism helps rebuild coordination after it collapses, and whether an end-of-game effect exists and affects the role of gradualism in coordination.

Compared to the literature on coordination games, this paper also suggests that limiting choices in each period, plus a well-designed path of allowed choices, may help reach social optimal outcomes. This follows the spirit of "libertarian paternalism" (Sunstein & Thaler, 2003): "designing institutions that help people make better decisions but do not impinge on their freedom to choose."[36] Questions about the optimal path to attain a long-run objective of social optimal deserve future studies (e.g., the best way of intra-organization training, the optimal process of political reform from an authoritarian regime towards a more open system).

---

[35] Fischbacher and Gächter (2010) study the dynamics of free riding in standard public good experiments, and find that both belief updating and social preferences are important in explaining the results.

[36] In the gradualism treatment of our experiment, we limit the number of choices in each period to 2 (i.e., each one can choose not to give or to give the threshold in each period). Since our experiment involves collective interactions and thus a more obvious externality of own action on others, it is less surprising that limiting individual choices may be good for social welfare. What is more interesting in our findings is that individuals still have the freedom to choose between two options in each period (i.e., they are not forced to cooperate), and there is no sanction, punishment, and social pressure.

To the best of our knowledge, our study is the first one that clearly tests the role of exogenous gradualism in coordination within a given group. It adopts an exogenous setting of threshold path from the perspective of social planner. The results have important implications for future research on concrete real world policies to promote cooperation and coordination among individuals, organizations, regions and countries.

References

Abbott, K. W., & Snidal, D. (2002). Filling in the folk theorem: The role of gradualism and legalization in international cooperation to combat corruption. Presented on the American Political Science Association Meeting, Boston, MA, August 29-September 1, 2002.

Arrow, K. J. (1974). *The limits of organization.* New York: W. W. Norton & Company.

Andreoni, J., & Samuelson, L. (2006). Building rational cooperation. *Journal of Economic Theory, 127(1),* 117-154.

Bagnoli, M., & Lipman, B. L. (1989). Provision of public goods: Fully implementing the core through private contributions. *The Review of Economic Studies, 56(4),* 583-601.

Bagnoli, M., & McKee, M. (1991). Voluntary contribution games: Efficient private provision of public goods. *Economic Inquiry, 29(2),* 351.

Berninghaus, S. K., & Ehrhart, K. M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of economic behavior & organization, 37(2),* 231-248.

Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior and Organization, 60(1),* 11-26.

Bohnet, I., & Huck, S. (2004). Repetition and reputation: Implications for trust and trustworthiness when institutions change. *American Economic Review, 94(2),* 362-366.

Bornstein, G., Gneezy, U., & Nagel, R. (2002). The effect of intergroup competition on group coordination: an experimental study. *Games and Economic Behavior, 41(1),* 1-25.

Brandts, J., & Cooper, D. J. (2006a). Observability and overcoming coordination failure in organizations. *Experimental Economics, 9(4),* 407–423.

Brandts, J., & Cooper, D. J. (2006b). A change would do you good.... An experimental study on how to overcome coordination failure in organizations. *American Economic Review, 96(3),* 669-693.

Cachon, G. P., & Camerer, C. F. (1996). Loss-avoidance and forward induction in experimental coordination games. *Quarterly Journal of Economics, 111(1),* 165-194.

Camerer, C., & Ho, T. H. (1999). Experience weighted attraction learning in normal form Games. *Econometrica, 67(4),* 827-874.

Carpenter, J. P. (2007). The demand for punishment. *Journal of Economic Behavior and Organization, 62(4),* 522-542.

Charness, G. (2000). Self-serving cheap talk: A test of Aumann's conjecture. *Games and Economic Behavior, 33(2),* 177-194.

Chaudhuri, A., Schotter, A., & Sopher, B. (2009). Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. *Economic Journal, 119(534),* 91-122.

Choi, S., Gale, D., & Kariv, S. (2008). Sequential equilibrium in monotone games: A theory-based analysis of experimental data. *Journal of Economic Theory, 143(1),* 302-330.

Clark, K., & Sefton, M. (2001). Repetition and signalling: Experimental evidence from games with efficient equilibria. *Economics Letters, 70(3),* 357–362.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, *69(2),* 318-328.

Clark, K., & Sefton, M. (2001). Repetition and signalling: Experimental evidence from games with efficient equilibria. *Economics Letters, 70(3),* 357-362.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1990). Selection criteria in coordination games: Some experimental results. *American Economic Review, 80(1),* 218–233.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1992). Communication in coordination games. *Quarterly Journal of Economics, 107(2),* 739–771.

Cooper, R., & John, A. (1988). Coordinating coordination failures in Keynesian models. *Quarterly Journal of Economics, 103(3),* 441-463.

Cornes, R., & Hartley, R. (2007). Weak links, good shots and other public good games: Building on BBV. *Journal of Public Economics, 91(9),* 1684-1707.

Costa-Gomes, M. A., & Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *Review of Economic Studies, 75(3),* 729-762.

Crawford, V. P. (1995). Adaptive dynamics in coordination games. *Econometrica, 63(1),* 103-143.

Croson, R. T. A., & Marks, M. B. (2000). Step returns in threshold public goods: A meta- and experimental analysis. *Experimental Economics, 2(3),* 239-259.

DellaVigna, S. (2009.) Psychology and economics: Evidence from the field. *Journal of Economic Literature, 47(2),* 315-372.

Duffy, J., & Feltovich, N. (2002). Do actions speak louder than words? Observation vs. cheap talk as coordination devices. *Games and Economic Behavior, 39(1),* 1–27.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90(4),* 980-994.

Fischbacher, U. (2007). z–Tree: Zurich toolbox for ready–made economic experiments. *Experimental Economics*, *10(2),* 171–178.

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review, 100(1),* 541–556.

Gächter, S., & Herrmann, B. Forthcoming. The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia. *European Economic Review*.

Gale, D. (1995). Dynamic coordination games. *Economic Theory*, *5(1),* 1-18.

Gale, D. (2001). Monotone games with positive spillovers. *Games and Economic Behavior, 37(2),* 295-320.

Gino, F., & Bazerman, M. H. (2009). When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. *Journal of Experimental Social Psychology, 45(4),* 708-719.

Hirshleifer, J. (1983). From weakest-link to best-shot: The voluntary provision of public goods. *Public Choice, 41(3), 371*-386.

Hyndman, K., Terracol, A., & Vaksmann, J. (2009). Learning and sophistication in coordination games. *Experimental Economics, 12(4),* 450-472.

Isaac, R. M., Schmidtz, D., & Walker, J. M. (1989). The assurance problem in a laboratory market. *Public Choice, 62(3),* 217-236.

Knez, M., & Camerer, C. F. (1994). Creating expectational assets in the laboratory: Coordination in weakest-link games. *Strategic Management Journal, 15(8),* 101-119.

Knez, M., & Camerer, C. F. (2000). Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. *Organizational Behavior and Human Decision Processes, 82(2),* 194-216.

Kurzban, R., Rigdon, M. L., & Wilson, B. J. (2008). Incremental approaches to establishing trust. *Experimental Economics, 11(4),* 370-389.

Ledyard, J. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *The handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Lockwood, B., & Thomas, J. P. (2002). Gradualism and irreversibility. *Review of Economic Studies, 69(2),* 339-356.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics, 116(4),* 1149-1187.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior, 10(1),* 6-38.

Marks, M., & Croson, R. (1998). Alternative rebate rules in the provision of a threshold public good: An experimental investigation. *Journal of Public Economics, 67(2),* 195-220.

Marks, M. B., & Croson, R. T. A. (1999). The effect of incomplete information in a threshold public goods experiment. *Public Choice, 99(1),* 103-118.

Masclet, D., Noussair, C., Tucker, S., & Villeval, M. C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review, 93*(1), 366-380.

Nyarko, Y., & Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica, 70(3),* 971-1005.

Offerman, T., Sonnemans, J., & Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *The Economic Journal, 106 (437), 817*-845.

Offerman, T., Sonnemans, J., & Schram, A. (2001). Expectation formation in step-level public good games. *Economic Inquiry, 39(2),* 250-269.

Offerman, T., & van der Veen, A. (2010). How to subsidize contributions to public goods: Does the frog jump out of the boiling water? *Working paper, University of Amsterdam.*

Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *American Economic Review, 100(3)*, 641-672.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review, 86(2),* 404-417.

Palfrey, T. R., & Rosenthal, H. (1984). Participation and the provision of discrete public goods: A strategic analysis. *Journal of Public Economics, 24(2),* 171-193.

Riechmann, T., & Weimann J. (2008). Competition as a coordination device: Experimental evidence from a minimum effort coordination game. *European Journal of Political Economy, 24(2)*, 437-454.

Rey-Biel, P. (2009). Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior, 65(2),* 572-585.

Pitchford, R., & Snyder, C. M. (2004). A solution to the hold-up problem involving gradual investment. *Journal of Economic Theory, 114(1),* 88-103.

Rutström, E. E., & Wilcox, N. T. (2004). Learning and belief elicitation: Observer effects. *Unpublished manuscript, University of Houston.*

Rutström, E. E., & Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior, 67(2),* 616-632.

Schram, A., Offerman, T., & Sonnemans, J. (2008). Explaining the comparative statics in step-level public good games. *Handbook of Experimental Economics Results, 1,* 817-824.

Schelling, T. C. (1960). *The strategy of conflict.* Cambridge: Harvard University Press.

Thaler, R., & Sunstein, C. (2003). Libertarian paternalism. *American Economic Review, 93(2),* 175-179.

Thaler, R., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112(1),* 164-187.

Tingley, D. H., & Wang, S. W. (2010). Belief updating in sequential games of two-sided incomplete information: An experimental study of a crisis bargaining model. *Quarterly Journal of Political Science, 5(3),* 243-255.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review, 80(1),* 234-248.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *Quarterly Journal of Economics, 106(3), 885-910.*

Weber, R., Camerer, C., Rottenstreich, Y., & Knez, M. (2001). The illusion of leadership: Misattribution of cause in coordination games. *Organization Science, 12(5),* 582-598.

Watson, J. (1999). Starting small and renegotiation. *Journal of Economic Theory, 85(1),* 52-90.

Watson, J. (2002). Starting small and commitment. *Games and Economic Behavior, 38(1),* 176-199.

Weber, R. A. (2005). Managing growth to achieve efficient coordination: Theory and experimental evidence. *Unpublished paper.*

Weber, R. A. (2006). Managing growth to achieve efficient coordination in large groups. *American Economic Review, 96(1),* 114-126.

Weingast, B. (1997). The political foundations of democracy and the rule of law. *American Political Science Review, 91(2),* 245-263.

Figure 1: Threshold Patterns of the Treatments in the First Stage

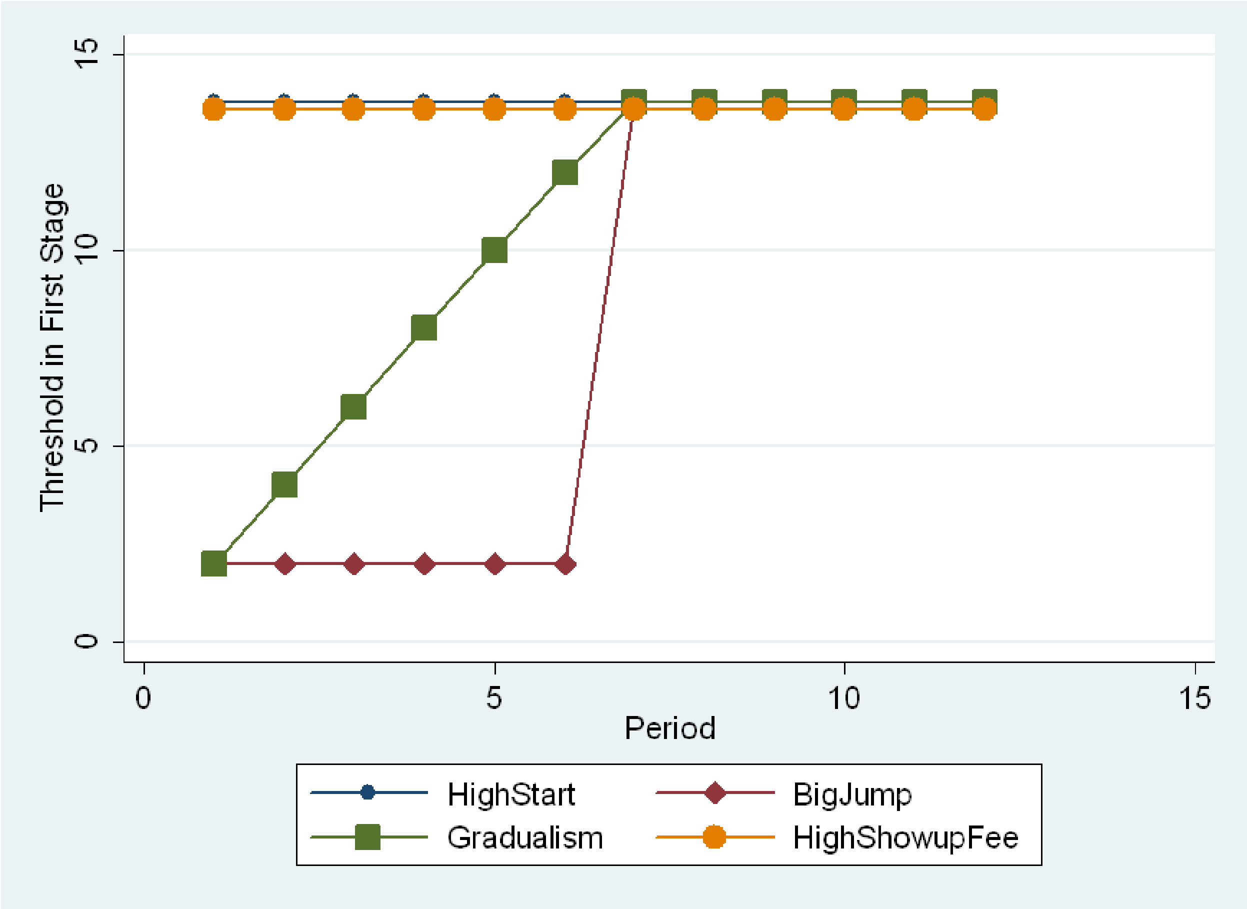
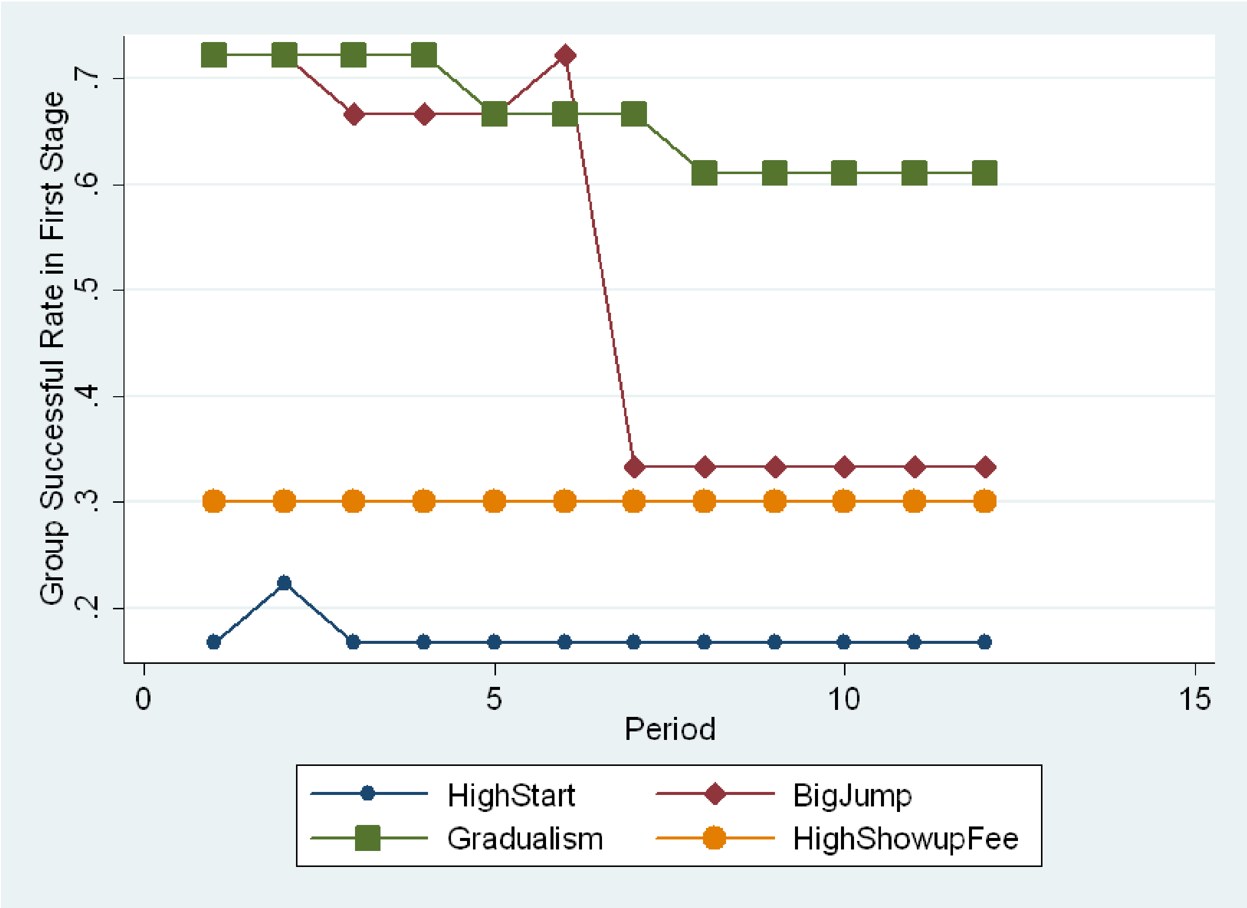
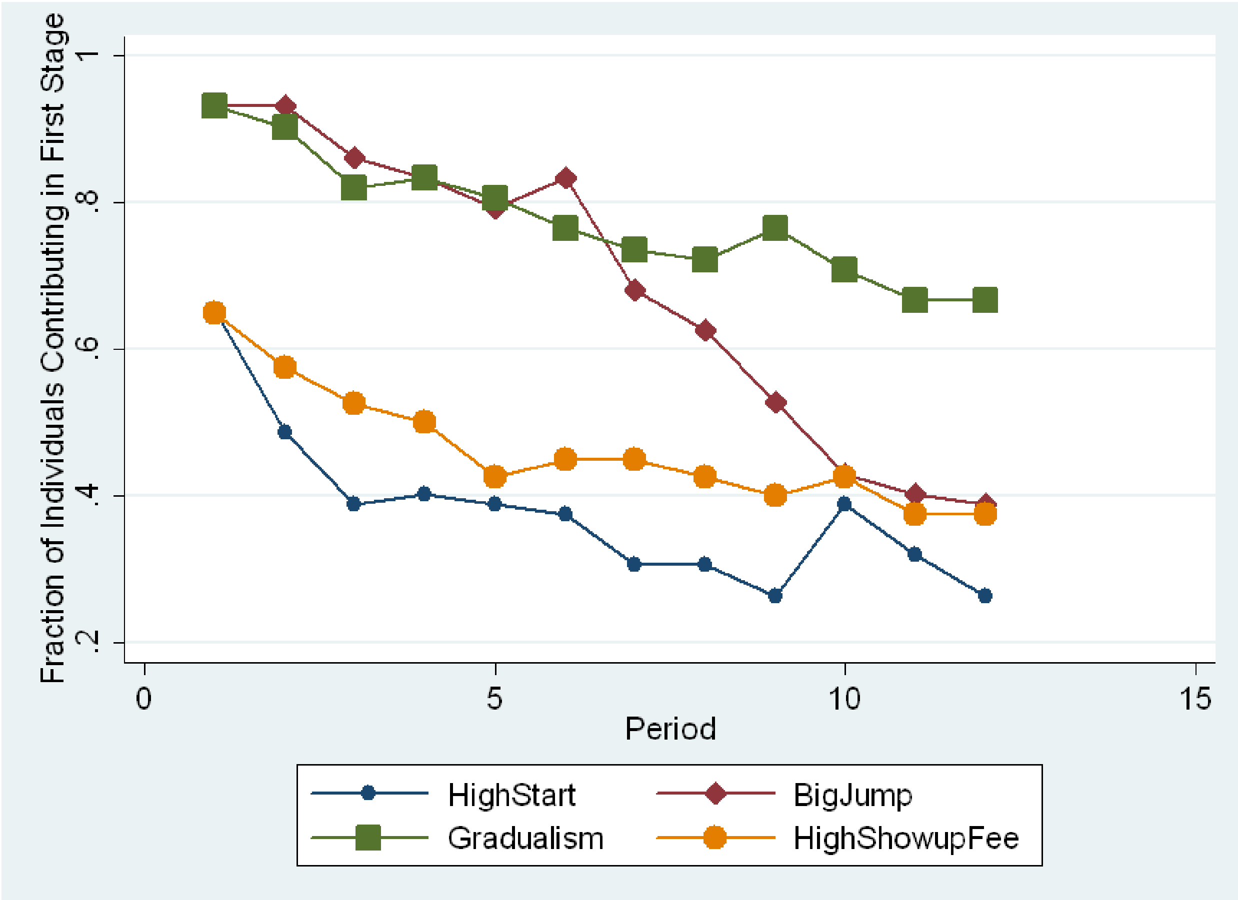


Figure 2: Success Rates of Groups by Treatment and Period in the First Stage



Note: A group is successful in coordination if all 4 members give the required amount (threshold) in that period.

Figure 3: Fraction of Individuals Contributing by Treatment and Period in the First Stage



Note: Contribution is a binary variable: 1 indicates that the individual gives the required amount (threshold) in that period, while 0 indicates not giving.

Figure 4: Average Individual Earning by Treatment and Period in the First Stage

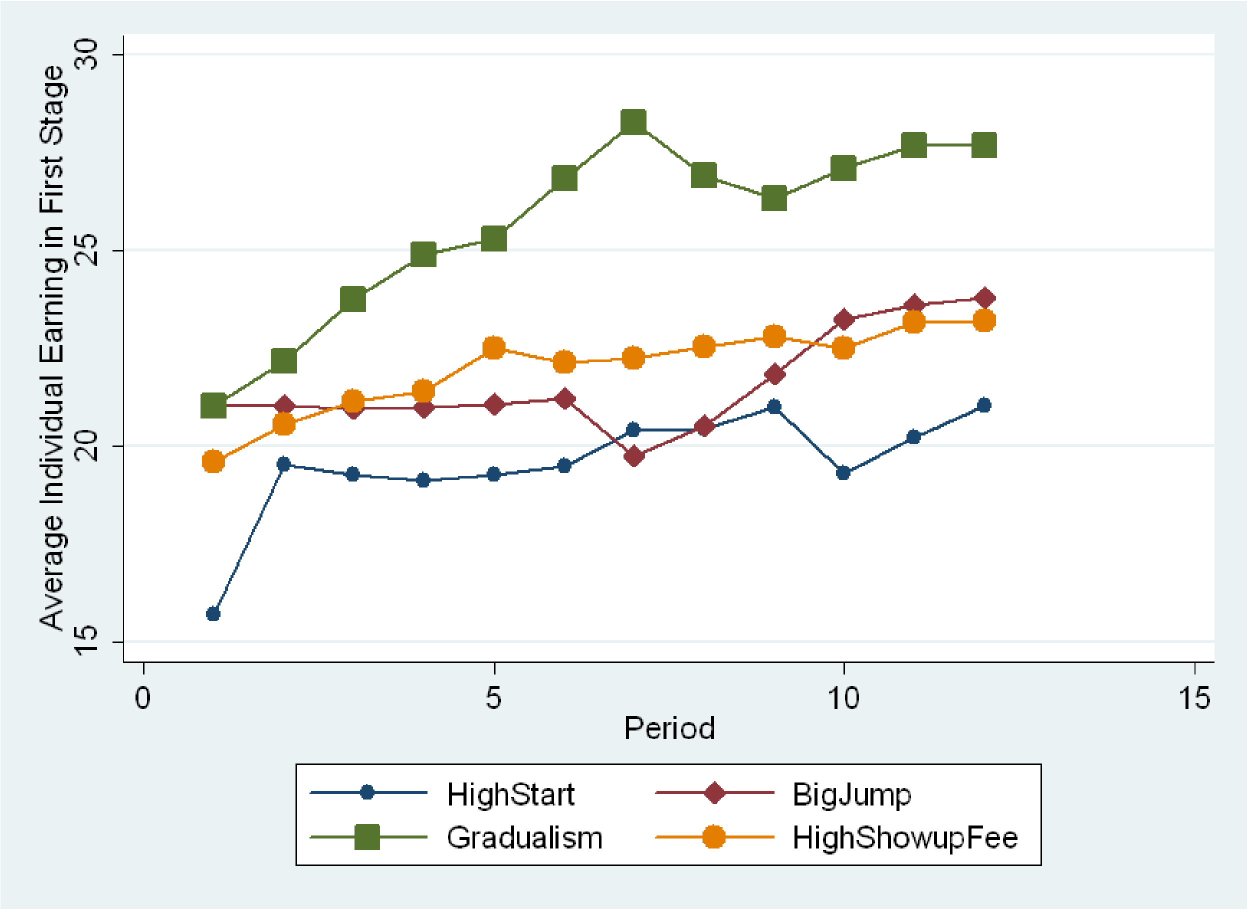


Table 1: Summary Statistics of Subjects' Survey Information

Variable	Mean and standard deviation	Observations
Age	22.05 (3.25)	255
Male	0.41 (0.49)	255
Income	1.32 (1.38)	255
Family Income	5.63 (2.69)	189
Family Economic Status	2.60 (0.74)	254
Risk Aversion Index	4.47 (1.80)	250
Han nationality	0.91 (0.29)	255
Student	0.91 (0.29)	255
Concentration:		
Economics	0.12 (0.33)	241
Other Social Sciences	0.16 (0.37)	241
Business	0.27 (0.45)	241
Humanity	0.12 (0.33)	241
Science	0.15 (0.35)	241
Engineering	0.17 (0.38)	241
Medical/Health	0.01 (0.09)	241

Note: Income is a scale variable from 0 to 13, with higher value indicating higher income (0: no income; 1: annual income<5000 *yuan*; 13: annual income>160,000 *yuan*). Family income is a scale variable from 1 to 12, with higher value indicating higher income (1: annual income<5000 *yuan*; 12: annual income>200,000 *yuan*). Family economic status is coded in the following way: 1 (lower), 2 (lower middle), 3 (middle), 4 (upper middle), 5 (upper). Risk aversion index is a scale from 0 to 10, with a higher value approximately indicating higher risk aversion, and is measured as the number of lottery A chosen by the subject in our questionnaire.

Table 2: Comparison of Participants' Characteristics by Treatment

	Dependent Variable									
	Age	Male	Income	Family Economic Status	Risk Aversion Index	Han Nationality	Student	Economics Major	Business Major
	(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	(10)
HIGH START	-0.306	-0.042	-0.125	-0.083	0.474*	0.028	0.083*	0.026	0.006
	(0.612)	(0.081)	(0.241)	(0.123)	(0.279)	(0.046)	(0.046)	(0.056)	(0.078)
BIG JUMP	-0.278	0.076	-0.360	0.217*	0.424	0.013	0.055	-0.000	0.030
	(0.627)	(0.083)	(0.236)	(0.128)	(0.295)	(0.048)	(0.050)	(0.053)	(0.080)
HIGH SHOWUP FEE	-0.111	0.086	0.056	-0.197	0.795**	-0.028	-0.050	0.053	-0.126
	(0.695)	(0.098)	(0.336)	(0.146)	(0.399)	(0.063)	(0.072)	(0.071)	(0.081)
Constant	22.236***	0.389***	1.444***	2.597***	4.097***	0.903***	0.875***	0.104***	0.284***
	(0.547)	(0.058)	(0.204)	(0.094)	(0.197)	(0.035)	(0.039)	(0.038)	(0.056)
Observations	255	255	255	254	250	255	255	241	241
R-squared	0.00	0.01	0.01	0.04	0.02	0.00	0.03	0.00	0.01

Note: The default treatment is "Gradualism." Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. For measures of variables, see note in Table 1.

Table 3: Summary of Treatments in the First Stage

Treatment	High Start	Big Jump	Gradualism	High Showup Fee
Endowment in each period	20	20	20	20
Show up Fee (points)	400	400	400	480
Exchange Rate (points/*yuan*)	40	40	40	40
Threshold in period 1	14*	2	2	14*
Threshold in period 6	14*	2	12	14*
Threshold in period 7-12	14*	14*	14*	14*
Number of groups	18	18	18	10
Number of subjects	72	72	72	40
Number of groups successful in period 1	3	13	13	3
Number of groups successful in period 7	3	6	12	3
Number of groups successful in period 12	3	6	11	3
Percent of groups successful in period 1	16.7%	72.2%	72.2%	30%
Percent of groups successful in period 7	16.7%	33.3%	66.7%	30%
Percent of groups successful in period 12	16.7%	33.3%	61.1%	30%

Table 4: Summary Statistics of Performance

Panel A: Performances over Periods in the First Stage

Variable	All Subjects	High Start	Big Jump	Gradualism	High Show-up Fee
All 12 periods:					
Fraction of group members contributing per period	0.59	0.38	0.69	0.77	0.46
	(0.49)	(0.49)	(0.46)	(0.42)	(0.50)
Number of group members contributing per period	2.36	1.51	2.75	3.11	1.86
	(1.62)	(1.42)	(1.52)	(0.41)	(1.63)
Group successful rate per period	0.43	0.17	0.51	0.66	0.3
	(0.49)	(0.38)	(0.50)	(0.47)	(0.46)
Individual earning by period	22.23	19.57	21.58	25.66	21.98
	(8.32)	(8.50)	(6.89)	(7.77)	(9.22)
Observations (subject*period)	3072	864	864	864	480
First 6 periods only:					
Individual earning up to Period 6	127.52	112.42	126.31	143.94	127.35
	(37.38)	(46.32)	(7.84)	(27.22)	(53.35)
Median of individual earning up to Period 6	120	106	130	162	106
Observations (subject)	256	72	72	72	40

Panel B: Overall Performance (Both Stages)

Variable	All Subjects
Individual total earning: points	856.82
	(116.14)
Individual total earning: *yuan*	21.42
	(2.90)
Observations	256

Note: The mean and standard deviation (in parentheses) are reported in the table. For variables in Panel A, there is one observation for each of the 256 subjects in each period over the 12 periods of the first stage. For the variable in Panel B, i.e., individual total earning (including show-up fee) of the whole game (including both stages), there are only 256 observations, since it is only attained at the end of the game. The actual earning in *yuan* from the experiment is the total earning in points divided by 40. The exchange rate of *yuan/USD* was about 6.7.

Table 5: Contribution, Earning and Success in Periods 7-12 of the First Stage by Treatments

	Individual Contribution		Individual Earning		Success (Group-level)		
	Period 7	Period 12	Period 7	Period 12	Period 7	Period 12	Period 7-12
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HIGH START	-0.431***	-0.403***	-7.861***	-6.639***	-0.500***	-0.444**	-0.454***
	(0.126)	(0.136)	(2.472)	(2.401)	(0.171)	(0.170)	(0.144)
BIG JUMP	-0.056	-0.278*	-8.528***	-3.889	-0.333*	-0.278	-0.287*
	(0.113)	(0.152)	(3.076)	(2.587)	(0.185)	(0.182)	(0.157)
HIGH SHOWUP FEE	-0.286*	-0.292	-6.028*	-4.467	-0.367*	-0.311	-0.320*
	(0.158)	(0.176)	(3.203)	(3.049)	(0.199)	(0.201)	(0.186)
Constant	0.736***	0.667***	28.278***	27.667***	0.667***	0.611***	0.620***
	(0.092)	(0.107)	(1.909)	(1.925)	(0.117)	(0.121)	(0.114)
Observations	256	256	256	256	64	64	384
R-squared	0.13	0.10	0.13	0.10	0.16	0.13	0.13

Note: The default treatment is "Gradualism." Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Individual contribution and group-level success are binary variables. Standard errors are clustered at session level for (5) and (6), and clustered at group level for all other regressions. The observations are at individual level for regression (1)-(4), and at group level for regression (5)-(7). There are 256 participants in 64 groups of 18 sessions. When we add survey controls for regression (1)-(4) where the observations are at individual level, the results (not reported) remain similar.

Table 6: Contribution and Earning in Each Period of the second Stage by Treatment

	Period 1		Period 2		Whole Stage	
	Contribution	Earning	Contribution	Earning	Contribution	Earning
	(1)	(2)	(3)	(4)	(5)	(6)
HIGH START	-0.125*	1.639	0	0.389	0.00174	0.514
	(0.0693)	(1.764)	(0.0727)	(1.524)	(0.0519)	(1.143)
BIG JUMP	-0.0972	3.111	0	1.833	0.0660	1.240
	(0.0720)	(1.890)	(0.0914)	(1.500)	(0.0658)	(1.297)
HIGH SHOWUP FEE	-0.161*	-1.467	0.0528	-5.100**	-0.0958	-2.447
	(0.0892)	(2.438)	(0.0913)	(2.230)	(0.0710)	(1.541)
Constant	0.861***	18.17***	0.597***	21.75***	0.505***	22.47***
	(0.0474)	(1.895)	(0.0639)	(1.493)	(0.0577)	(1.195)
Observations	256	256	256	256	2,048	2,048
R-squared	0.020	0.018	0.002	0.044	0.011	0.016

Note: The default treatment is "Gradualism." Robust standard errors in parentheses. Standard errors are all clustered at group level. * significant at 10%; ** significant at 5%; *** significant at 1%. Individual contribution is a binary variable.

Appendix A: Proofs

Lemma 1 (Belief Update) For any player *i*, there is $[LB_{t+1}^i, UB_{t+1}^i] \subseteq [LB_t^i, UB_t^i]$ for any period *t*.

Proof: From period *t* to period *t+1*, there are three cases regarding player *i*'s belief updating as follows.

Case 1: Player *i* has not given at period *t*. There is $[LB_{t+1}^i, UB_{t+1}^i] = [LB_t^i, UB_t^i]$.

Case 2: Player *i* has given at period *t* and the coordination succeeds. There is $[LB_{t+1}^i, UB_{t+1}^i] = [Th_t, UB_t^i]$ where $LB_t^i < Th_t < UB_t^i$.

Case 3: Player *i* has given at period *t* and the coordination fails. There is $[LB_{t+1}^i, UB_{t+1}^i] = [LB_t^i, Th_t]$ where $LB_t^i < Th_t < UB_t^i$.

Hence, in all cases there is $[LB_{t+1}^i, UB_{t+1}^i] \subseteq [LB_t^i, UB_t^i]$. *Q.E.D.*

Lemma 2 (Gradualism) There always exists a threshold $Th_t > \mathrm{Min}_i\, LB_t^i$ to guarantee successful coordination if $UB_t^i > LB_t^i$ for all *i* for any period *t*.

Proof: Consider the following equations for all *i*:

$$1 - F_t^i(Th_t) = \theta_i(Th_t)$$

As $F_t^i()$ and $\theta_i()$ are both increasing, $F_t^i(LB_t^i) = 0$, $F_t^i(UB_t^i) = 1$ and $0.5 < \theta_i(Th_t) < 1$, there must be a unique solution Th_t^{i*} for eath player *i*.

Since $UB_t^i > LB_t^i$ for all *i* for any period *t*, there must be $Th_t^{i*} > LB_t^i$ for all *i*. Hence, let $Th_t = \mathrm{Min}_i\, Th_t^{i*}$ as the threshold can guarantee successful coordination at any period *t*, and it is obvious that $Th_t = \mathrm{Min}_i\, Th_t^{i*} > \mathrm{Min}_i\, LB_t^i$. *Q.E.D.*

Proposition 1 (Gradualism) There exists an infinite path of threshold $\{Th_t \mid Th_{t+1} \geq Th_t\}$ to guarantee successful coordination from period 1.

Proof: We prove this corollary by mathematical induction.

At period 1, the supports of $F_1^i()$ for all *i* are $[0, UB]$, and thus $UB_t^i > LB_t^i$ is satisfied for all *i*. Hence Lemma 2 applies and there exists a Th_1 to guarantee successful coordination.

Note that, if a coordination succeeds at period *t*, there is $UB_{t+1}^i = UB_t^i$ for all *i*.

Therefore, in period *t* given coordination succeed at all periods prior to *t*, the upper bound of supports of $F_t^i()$ for all *i* are still UB, and thus $UB_t^i > LB_t^i$ is still satisfied for all *i*. Hence Lemma 2 still applies and there exists a Th_t to guarantee successful coordination.

This process will never stop, and $Th_{t+1} \geq Th_t$ is ensured by Lemma 1. Furthermore, we can see that there is $Th_{t+1} > Th_t$ as long as $Th_t < UB$ from Lemma 2. *Q.E.D.*

Proposition 2 (Memory of Failure) In the path of the game, if there are and only are fewer than *I* players who have given in any period *t* with a threshold Th_t , the future threshold that can make coordination successful must be $Th_\tau < Th_t$ for all $\tau > t$.

Proof: We prove this proposition by contradiction. At period *t*, there are and only are fewer than *I* players who have given. Hence, the coordination fails at period *t*, and for any player *i* who has given, her support of $F^i_{t+1}()$ at time *t+1* is $[LB^i_t, Th^i_t]$. Suppose there is a coordination with $Th_\tau \geq Th_t$ at period $\tau > t$. This requires player *i*'s support of $F^i_\tau()$ to be $[LB^i_\tau, UB^i_\tau]$ where $LB^i_\tau < Th_\tau < UB^i_\tau$. However, this violates Lemma 1, which is a contradiction. *Q.E.D.*

Appendix B: Experimental Instruction

The study is conducted anonymously. Participants will be identified only by code numbers. There is no communication among the participants. The experiment will last from 30 minutes to one hour. Please raise your hand if anything is unclear to you.

Experiment Structure

This experiment will consist of 2 independent stages. You will receive instructions for each stage on the screen before that stage begins. In each stage, you are playing in a group of 4 members (including yourself). For each stage, the group members are randomly selected and would NOT change during that stage. However, the groups would be reshuffled in new groups after the first stage.

Rule of Each Period

Please note that the experiment consists of 2 stages, and each stage has some periods. The following rule applies to each period.

In each period, you are assigned an endowment of 20 points, and asked to give a stated amount to a group pool. The stated amount may change across periods. In each period you can decide whether not to give, or to give exactly that amount, but can not give other points. You cannot know others' choices when you make your own decision.

If all 4 members of your group (including yourself) give the stated amount, you will get twice that amount back (thus having a net return equaling that amount). But if not all of your group members give, you will NOT get any of your given points back and will thus end the period with only the points you do not give.

So in each period, your earning will depend on the following cases:

Case 1: If all 4 members give the stated amount, then you earn: 20+that amount

Case 2: If you give, but not all other 3 members give, then you earn: 20-that amount

Case 3: If you do not give, no matte whether other members give or not, then you earn: 20

A special case of Case 3 is as follows:

- Case 4: If all 4 members do not give, then you earn 20 (each of 4 members earns 20)

Examples

Example 1: You are asked to give 10 points. You give, and all other group members also give. Then in this period each of you earns 20+10=30 points.

Example 2: You are asked to give 10 points. You give, two other group members also give, but the last one does not. Then in this period, each of you and the other two members earns 20-10=10 points, while the last member earns 20 points.

Example 3: You are asked to give 10 points. You do not give, but all three other group members give. Then in this period you earn 20 points, and each of the three other members earns 20-10=10 points.

We will have examinations on the computer to make sure you understand the rule. You can start the experiment only after you answer all questions correctly.

Payment

Your final payment for this experiment is the sum of two parts. The first is a show-up fee of about 400 points.[1] The second is a performance payment, i.e., the sum of your earnings from all periods in two stages. The conversion rate is 40 points = ¥1.00. All payments will be in cash.

At the end of the experiment, you will be asked to fill out a simple questionnaire. Then you can collect your earnings by presenting your code number to the supervisor. Your earnings will be in an envelope marked with your code number.

[1] For the eight sessions without the "High Show-up Fee" treatment, it is stated as "a show-up fee of 400 points."

Appendix C: Risk Aversion Questions

Your code___ ______

In the table below, you are presented with a choice between two lotteries, A or B, along with the payoff matrix for each lottery.

For example, the first row shows that lottery A offers a 10% chance of receiving ¥20.00 and a 90% chance of receiving ¥16.00. Similarly, lottery B offers a 10% chance of receiving ¥38.50 and a 90% chance of ¥1.00.

In the third table column, simply indicate given the two lotteries in each row, which one would you prefer if you are given the choice? A or B for each row?

Lottery A				**Lottery B**				**Your lottery choice**
prob(¥20.00)		prob(¥16.00)		prob(¥38.50)		prob(¥1.00)		
0.1	¥20.00	0.9	¥16.00	0.1	¥38.50	0.9	¥1.00	
0.2	¥20.00	0.8	¥16.00	0.2	¥38.50	0.8	¥1.00	
0.3	¥20.00	0.7	¥16.00	0.3	¥38.50	0.7	¥1.00	
0.4	¥20.00	0.6	¥16.00	0.4	¥38.50	0.6	¥1.00	
0.5	¥20.00	0.5	¥16.00	0.5	¥38.50	0.5	¥1.00	
0.6	¥20.00	0.4	¥16.00	0.6	¥38.50	0.4	¥1.00	
0.7	¥20.00	0.3	¥16.00	0.7	¥38.50	0.3	¥1.00	
0.8	¥20.00	0.2	¥16.00	0.8	¥38.50	0.2	¥1.00	
0.9	¥20.00	0.1	¥16.00	0.9	¥38.50	0.1	¥1.00	
1	¥20.00	0	¥16.00	1	¥38.50	0	¥1.00	