Computable and Experimental Economics Laboratory

Stefania Bortolotti, Giovanna Devetag and Andreas Ortmann

Exploring the effects of real effort in a weak-link experiment

CEEL Working Paper 1-09

Exploring the effects of real effort in a weak-link experiment

Stefania Bortolotti

CIFREM, University of Trento

Giovanna Devetag

University of Perugia

Andreas Ortmann

CERGE-EI,
Charles University and
Academy of Sciences of the Czech Republic, Prague

January 23, 2009

Abstract

We report results from a weak-link – often also called minimum-effort – game experiment with multiple Pareto-ranked strict pure-strategy Nash equilibria, using a real-effort rather than a chosen-effort task: subjects have to sort and count coins and their payoff depends on the worst performance in the group. While in the initial rounds our subjects typically coordinate on inefficient outcomes, almost 80 percent of the groups are able to overcome coordination failure in the later rounds. Our results are in stark contrast to results typically reported in the literature.

JEL classification: C72, C92.

Keywords: real effort, weak-link game, coordination, laboratory experiments.

1. Introduction

Coordinating the activities of employees is a key problem of all organizations (March and Simon, 1958; Schelling, 1960; Lawrence and Lorsch, 1967; Malone and Crowston, 1990; Malone and Crowston, 1991). An important experimental literature seems to have demonstrated convincingly that, even when incentive problems are not present, employees often fail to match the actions of "co-workers", causing "organizations" to drift into, or stay locked in, inefficient

equilibria for a wide range of coordination games (see, for reviews, Camerer, 2003 and Devetag and Ortmann, 2007b). Arguably the most prominent workhorse in this literature has been the weak-link, or minimum-effort, game which features multiple Pareto-ranked pure-strategy Nash equilibria and the toughest task interdependence possible: the minimum effort supplied by a member of a group of employees determines the outcome of the organization and, consequently, everybody's payoff.[1] Van Huyck et al. (1990) were the first to demonstrate the speedy downward drift to the minimum effort for this game (a process commonly called "coordination failure", even if subjects manage to coordinate on the worst equilibrium), a result which – ceteris paribus – has been replicated consistently.[2]

As also suggested by Weber (2006), the evidence on coordination failure in the lab seems, however, partly inconsistent with what we observe outside the lab. Relatedly, Devetag and Ortmann (2007b) have shown that there are a number of ways, which typically seem to increase external validity, to engineer coordination successes in the lab. However, all weak-link game experiments conducted so far rely on the critical assumption that the choice of a nominal effort level from a given range, or a "chosen effort", is a reliable proxy for real effort. This assumption is problematic since in actual organizations work "involves effort, fatigue, boredom, excitement and other affections not present in the abstract experiments" (van Dijk et al., 2001: 189).

The concern about this assumption stems from important recent questions about the external validity of laboratory data (e.g., List, 2006; Levitt and List, 2007). Indeed, experiments in other classes of games have demonstrated that the choice of real effort can make a significant difference (Gneezy and List, 2006; Hennig-Schmidt et. al., 2005; García-Gallego et. al., 2008). To the best of our knowledge, no evidence exists yet about how real effort affects coordination in the weak-link

---

[1] The take-off of airplanes is the prototypical example of a coordination problem of the weak-link type (e.g., Knez and Simester, 2001); the airplane cannot depart before all operations (e.g., fueling, security checks, loading of luggage, boarding of passengers, etc.) have been completed. On-time departure depends upon the slowest operation: a unilateral increase in the level of effort is likely to be in vain if it is not matched by an increase in effort of complementary activities. For example, speeding up the boarding of passengers does not make sense if it is not matched by speedy loading of luggage. Other examples of interest are the writing of a grant proposal involving several participants or an edited volume involving several authors.

[2] Engelmann and Normann (2007) is the only exception that we are aware of.

game or, for that matter, any other coordination game often studied in the lab. Thus, in this paper we explore whether coordination failure in weak-link game experiments survives real-effort implementations.

We designed an experiment where effort had to be exerted in performing the following task: our subjects (employees, hereafter) had to sort and count coins worth 1, 2, 5, and 10 Euro cents within a given time interval. Through pilot experiments, we calibrated the time interval so that employees had to exert some effort to accomplish the task; moreover, since the time allotted was likely to be, at least initially, not sufficient for some of the participants, employees were offered the opportunity to buy extra time to complete the task (i.e., employees of our laboratory firm could decide to spend extra time at their workplace to complete their task).[3] In *Group* treatments, employees were randomly assigned to groups of four (firms, hereafter) and were paid according to a weak-link earnings table: specifically, their earnings were a function of the worst counting performance of a firm member. In *Individual* control treatments, strategic interaction and strategic uncertainty were stripped away by having employees "work" alone and paying them according to their individual performance. Motivated by results of previous chosen-effort weak-link game experiments, we also explored the impact of temporarily increased incentives, different cost parameters, and different information treatments.

Our results are in sharp contrast to the speedy downward drift to the minimum effort commonly observed in previous experiments. Indeed, after initially failing, almost 80 percent of our laboratory firms succeeded in coordinating on the efficient equilibrium. A substantial part of the firms that coordinated on the inefficient outcome in the initial rounds were able to overcome coordination failure well before incentives were modified through increased bonuses, rendering this treatment of lesser importance than in other contexts (e.g., Brandts and Cooper, 2006b).

Our results contribute in at least three ways to the existing literature. First, we provide a laboratory testbed that allows the study of coordination games with real effort rather than chosen

---

[3] In real life the cost of staying longer at work is the opportunity cost of leisure, which is difficult to implement in the lab, especially in a repeated game. Therefore, extra time is bought here for money.

effort tasks. Second, we show that in our testbed a history of coordination failure in performing a real task can be overcome in a large majority of cases. We thus show that chosen effort might not be a reliable measure of real effort. Our findings fundamentally question the results reported in the large literature on coordination games and raise new questions. Third, we provide evidence on *how* and *why* coordination failure might be overcome in a real-effort setting and hence lay out a promising research agenda.

The paper is organized as follows: in Section 2 we introduce our real-effort weak-link game and in Section 3 we describe experimental design and implementation. In Section 4 we briefly review the literature. Section 5 presents the results and Section 6 concludes.

2. The real-effort weak-link game

2.1. The chosen-effort weak-link game

Purely theoretical considerations fall short of predicting which of the Pareto-ranked equilibria employees will select in a coordination game since none of the known selection criteria has empirically strong support. Van Huyck et al. (1990), for example, demonstrated the speedy downward drift to the worst equilibrium for a weak-link game with multiple Pareto-ranked strict pure-strategy Nash equilibria. Their result attracted considerable attention for its apparent support of selection principles such as security or risk dominance over efficiency.

The closely related laboratory set-up of Brandts and Cooper (2006b) is our point of departure. In their study, each group of employees represents a firm whose productivity is determined by the minimum effort provided by one of the firm's employees. The employees perform the task repeatedly with fixed matching. In each round, all employees decide their individual levels of effort.[4] The employees are guaranteed a fixed wage, F, and a bonus rate, B, which is announced by the experimenter at the beginning of each round. Employees independently select an effort level

---

[4] In chosen-effort experiments performance is solely determined by effort choice, as innate ability and experience clearly do not apply, contrary to our real-effort task.

4

$e_i \in \{0,10,20,30,40\}$ and incur a cost Ce_i, which increases linearly in the level of effort chosen. The per-round payoff to employee i is given by the following function:

$$\pi_i = F + B\left[min_{j \in \{1,2,...,N\}}(e_j)\right] - C(e_i) \qquad (1)$$

where $\left[min_{j \in \{1,2,...,N\}}(e_j)\right]$ is the minimum effort provided by one of the firm's employees.

[Table 1 about here]

The parameters are set in such a way that coordinating on any effort level is a pure-strategy Nash equilibrium that is strict. If everyone picks the highest possible action ("effort"), then the Pareto-dominant ("efficient") equilibrium is selected (see Table 1). Choosing the highest effort is risky, however, since one may end up with zero if the minimum effort in the firm turns out to be zero. A similar argument applies in a repeated game, although for fixed personnel constellations the analysis gets complicated because across-round reasoning might be applied by employees. Consider, for example, an employee that is paid according to Table 1 and has experienced a history of coordination failure in previous periods (minimum effort = 0).[5] The employee has an incentive to provide a higher effort only if he/she expects all co-workers to raise their effort with a fairly high probability.[6] Clearly, if the cost of effort, C, is decreased, or the bonus rate, B, is increased, attempting to overcome coordination failure by raising one's own effort becomes less risky *ceteris paribus*.

2.2. The real-effort task

Our real-effort task required employees to sort and count, within a given time interval, a variable number of coins worth 1, 2, 5, and 10 Euro cents.[7] The task was, quite intentionally, highly

---

[5] The payoff matrix in Table 1 is extremely unforgiving and, since the purpose of the studies was to study how an initial history of coordination failure can be overcome, was designed to be so in order to induce coordination failure. Indeed, speedy unraveling towards inefficient outcomes was observed in all previous experiments which employed the same or similar tables (Brandts and Cooper, 2006b; Cooper, 2006; Hamman et al., 2007).

[6] For picking the higher effort level of 10 in order not to have a negative expected value, the player should expect that the probability that all others raise the effort is at least 5/6 (for more details see Brandts and Cooper, 2006b).

[7] The total number of coins contained in each bag varied from 32 to 35. Coins worth 1 and 2 cents were present in amounts that varied from 8 to 14, coins worth 5 cents varied from 6 to 10, and coins worth 10 cents varied from 2 to 4. The composition of each bag was made in such a way that overall difficulty and total number of coins were roughly the

repetitive and unexciting, so as to minimize the possibility that our employees would derive any positive utility from performing the task. Although highly stylized, our task shares many features with a variety of blue and white collar jobs. For instance, assembly line workers are required to perform highly repetitive tasks under tight time constraints and low quality in one step of the production process may result in low quality outcomes. Our real-lab task also captures the main features of data entry and payroll clerks types of task. Since the time for completion was short, the task required both physical and cognitive effort. Indeed, drawing on a series of pilot experiments, we calibrated our experiment so that the regular time employees were endowed with would, at least initially, not be sufficient for some participants, which would induce coordination failure. However, employees were given the opportunity to buy extra time to complete the task whenever needed; notably, the additional time, if asked for, gave less skilled employees a chance to complete the task correctly. This opportunity, however, came at a cost. Think of this set-up as capturing the essence of organizational scenarios in which workers are willing to voluntarily spend extra time at their workplace, or at home, to complete their work.[8] The introduction of extra time is a crucial ingredient of our design and serves two main purposes: first, it makes coordination possible but not trivial even in the case of highly heterogeneous skills and poorly performing employees; second, it is risky since it entails a net loss, as compared to the safe strategy of doing nothing, if coordination fails.

2.3. The real-effort weak-link game

Following the advice of Davis and Holt (1993: 520) to not change too many things at once, we draw on the payoff function used in Brandts and Cooper (2006b), but depart by necessity from (1) in three important aspects. First, since the costs of effort are not known to us we do not explicitly include them in the payoff function. Second, the minimum effort in the firm is defined as the

---

same for all subjects and rounds. Pilot experiments confirmed that there was no difference in the time needed for sorting and counting coins included in different bags. The sequence of bags was the same in all sessions.

[8] Of course, in real life overtime represents the opportunity cost of reduced leisure, while in this experiment extra time entails an explicit cost.

highest number of errors in the counting task (worst performance, henceforth), truncated at four. We define the number of errors as the sum of the differences between the correct and reported number of coins for each coin type. Finally, the cost of buying extra time is included in the function. In each round, up to four slots of extra time could be bought. Thus, the per-round payoff received by any of the employees in our experiment is represented by the following payoff function:

$$\pi_i = F + B\left[4 - max_{j \in \{1,2,...,N\}}\left(e_j\right)\right] - C * (t_i) \qquad (2)$$

Where $\left[max_{j \in \{1,2,...,N\}}\left(e_j\right)\right] \in \{0,1,2,3,4\}$ is the worst performance (i.e., the highest number of counting errors in the firm), C is the cost of one slot of extra time, and t_i is the number of slots bought by employee i.

[Table 2 about here]

Table 2 presents the earnings tables for different parameterizations of B and C. A proper analysis of the equilibria in the game is not possible since, contrary to chosen-effort experiments, there is no one-to-one correspondence between effort and performance; indeed, "subjective payoffs" include the cost – possibly heterogeneous – of effort and necessarily differ from the nominal payoffs reported in the earnings table. Given the impossibility of knowing employees' subjective cost of effort in our setting, Tables 1 and 2 are not directly comparable.

That said, our earnings table shares many features with the one used by Brandts and Cooper (2006b). First, the bonus is computed on the basis of the worst outcome in the firm. Second, in each round employees have the possibility of getting a sure fixed wage of 200 by exerting zero effort as in Brandts and Cooper (2006b). (We explicitly stated in the instructions that employees would receive their fixed wage in each round even if they chose to do nothing.) Third, it seems reasonable to assume that the subjective cost of effort increases as the number of errors decreases. Moreover, in case one or more slots of extra time are needed to lower counting errors, employees also face an explicit cost and achieving higher coordination levels is, indeed, risky. That is, if the worst performance is equal to four errors, by buying extra time employees incur a net – and explicit – loss with respect to the fixed wage. At the same time, payoffs are such that, even buying four slots of

7

extra time may be worth the explicit cost if worst performance is equal to zero errors. (Compare that scenario to one where no extra time is bought but four or more errors are committed).[9] Finally, to overcome a history of failure might be difficult and risky. For instance, an employee that is willing to reduce his errors from 4 to 3, but needs to buy a slot of extra time to do so, incurs a cost of 50 while gaining at most 10.[10]

3. Experimental design

3.1. Design

The design of our experiment is meant to test whether real effort affects coordination in the weak-link game for selected parameterizations of the payoff function and different informational feedback modes. Moreover, in order to better understand our results, we compare conditions in which employees are paid according to either worst (*Group* condition) or individual performance (*Individual* condition).

Sixteen rounds divided in three Parts were played in each session. Informed by our pilot experiments and what is typically observed in auctions, we expected a task-learning effect for the initial rounds. Part 1, therefore, lasted eight rounds rather than the four rounds each that constituted Parts 2 and 3. We expected a decrease in the number of errors in the initial rounds and a marginal role of task learning in the subsequent rounds.

Previous evidence on the weak-link game suggests low deviation (from equilibrium) costs and the higher relative attractiveness of the payoff-dominant equilibria to be efficiency-enhancing (Devetag and Ortmann, 2007b for a review; see also Devetag and Ortmann, 2007a), while the evidence on the effects of full feedback is mixed (Van Huyck et. al., 1990; Devetag, 2005; Brandts and Cooper, 2006a). Hence, we tested whether these results also hold for real-effort tasks and varied our treatments accordingly along three dimensions. First, we changed the bonus rate *B* to test

---

[9] Clearly, it pays to buy extra time to bring errors (and worst performance) down only if the perceived cost of effort is not too high. In this case, and for this parameterization, the perceived cost of effort has to be no higher than 40.
[10] This example makes reference to *B*=60 and *C*=50 as reported in the first panel of Table 2.

how the relative attractiveness of the payoff-dominant outcome affects coordination. In *Variable* treatments, the bonus rate B was temporarily increased in Part 2, while in *Fixed* treatments it remained constant (see Table 3). Our sequencing of the bonus was designed to reproduce in the lab plausible real-world scenarios,[11] following Brandts and Cooper (2006b) and Hamman et al. (2007).

[Table 3 about here]

Second, in order to investigate how deviation costs affect coordination, the cost of buying each slot of extra time was varied by having high and low deviation-cost treatments.[12] Finally, two modes of informational feedback were implemented: in *Partial Feedback* treatments employees were informed only about the worst performance in the firm, while in *Full Feedback* treatments they were informed about the number of errors made (errors being truncated at four) and the extra time bought by each employee of their firm.

In Group treatments, four employees were randomly and anonymously selected to constitute a firm and this assignment did not change throughout the experiment (and was common knowledge). For the Group conditions we implemented only the variable bonus scheme in a within-subject treatment, whereas the between-subject treatment was composed of a 2 [Full vs. Partial Feedback] x 2 [High vs. Low deviation costs] design (see Table 4).

In Group conditions, performance (measured as number of errors) is the result of effort and ability, which may go in the same or the opposite direction. While ability tends to increase with task repetition and hence, ceteris paribus, causes the number of errors to decline over time (what we call *task learning*), the level of effort exerted may be a function of strategic considerations, and may increase or decrease over time as a function of others' behavior in the game and of expectations about the outcomes of future games (to which we refer as *strategic learning* below). It seems plausible to assume that the presence of uncertainty about others' effort and performance ("strategic

---

[11] The case of Continental Airlines, presented in Knez and Simester (2001) as a prototypical example of weak-link in the field, presents a similar evolution of incentives over time (see Knez and Simester, 2001 for detailed case discussion).

[12] In our experiment, contrary to previous chosen-effort experiments, the deviation costs neither refers to the cost of effort nor to the deviation from the equilibria since both effort and equilibria cannot be explicitly determined. In the remainder of the paper we use "deviation costs" to make reference to the cost of buying extra time.

uncertainty") tends to decrease effort with respect to a situation in which an employee is only rewarded according to his/her own performance. Hence, in order to disentangle the relative contribution of strategic learning and task learning, we ran baseline treatments without strategic uncertainty, for which the assumption that effort of employees is constant over time was therefore easier to rationalize.

In the Individual conditions, the same stage game was repeated by having employees working alone (i.e., group size equal to 1), with their per-round payoff still depending upon the number of errors made and extra time bought. The payoff tables were the same as the ones used in the Group treatments, only the label "Worst performance in the group (number of errors)" was substituted with "My performance (number of errors)". In the Individual treatments, we implemented a 2 [High vs. Low deviation costs] x 2 [Variable vs. Fixed bonus] design.[13] Table 4 summarizes the experimental design for both Individual and Group conditions, with the number of employees (subjects) reported for each treatment.

[Table 4 about here]

To sum up, our experimental design is meant to address several research questions. First of all, we want to explore the effect of real effort in the weak-link game as compared to the chosen effort used in previous lab experiments. In particular, we use our testbed to investigate whether the initial failure to coordinate on the efficient equilibrium, and the subsequent unraveling typically observed in similar weak-link experiments towards the worst of the equilibria, survives our real-effort setting, Specifically, we would like to answer these questions:

> *Question I: Is failure to coordinate on the efficient equilibrium observed in the initial round(s) in our real-effort setting?*

> *Question II: Is speedy unraveling towards the inefficient equilibrium observed in later rounds, or is it possible to overcome coordination failure by introducing real effort in the lab? How and why exactly does coordination evolve over time?*

---

[13] Note that in the Individual condition, the distinction between Full and Partial feedback treatment does not make sense.

Moreover, we investigate how strategic uncertainty and task learning interact, by comparing Group and Individual conditions, and how this interaction may affect the ability for employees to coordinate:

Question III: Does strategic uncertainty undermine the learning process in the counting task?

Finally, we explore the effect of different parameterizations and informational modes:

Question IV: Does an increased bonus enhance coordination?

Question V: Do modes of information feedback and size of deviation costs affect errors and coordination?

The Results section is organized around these five questions.

3.2. Implementation

A total of 144 subjects participated in the experiment, which was conducted at the Computable and Experimental Economics Lab (CEEL) at the University of Trento. A total of 16 sessions were conducted. All subjects were students from the University of Trento and were recruited through ads and posters.

Upon arrival, subjects were randomly assigned to a cubicle and no form of communication was allowed from that moment on. A paper copy of the instructions (see Appendix) and of the relevant earnings table was distributed at the beginning of each Part and instructions were read aloud to assure common knowledge. Following Brandts and Cooper (2006b), we used a corporate context instead of neutral terminology. After reading the instructions, a questionnaire was distributed to ensure correct understanding; in case of incorrect answers, the instructions were read again. Before the experiment proper started, in order to familiarize subjects with the task, they were given a small bag containing coins, similar to the one that would be used in the experiment, and were asked to report the result of their counting task on a paper. Before each round started, and before they received a small bag containing coins worth 1, 2, 5, and 10 cents, subjects were asked to wear

11

headphones. The total number of coins contained in the bag varied from round to round and from firm to firm, while the types of coins were the same throughout the experiment.

[Figure 1 about here]

In each round, subjects had 45 seconds of regular time to perform the task; a sound informed them of the beginning of the round and a clock visualized the passing of time on each participant's monitor. In each round, subjects had the option of buying at most four slots of extra time, lasting 12 seconds each, simply by clicking a button (see Figure 1). The result of the counting activity had to be reported in four cells in the bottom part of the screen before time expired; a sound always alerted subjects when the time (regular or extra) was about to expire. The number of errors made by each participant was computed as the sum of the errors in each cell; in case one or more cells were left blank or the numbers were not confirmed by clicking the "OK" button, the number of errors translated automatically into the worst possible outcome of 4 or more.

The sessions averaged 1 hour and 30 minutes. All values were expressed in Experimental Currency Units (ECUs) and were converted at the end of the experiment at the rate of 1 Euro for 333 ECUs. Subjects knew the conversion rate in advance and were paid their earnings plus a fixed show-up fee of two Euros (see instructions in the Appendix) privately. The average total earnings for the Group condition were 16.81 Euros. For the Individual condition the average total earnings were 20.98 Euros. Before we present and discuss our results, we briefly review related literature. A more detailed discussion may be found in Devetag and Ortmann (2007b).

4. Related literature

In their path-breaking work, Van Huyck et al. (1990) have documented that initial failure to coordinate on the efficient equilibrium and subsequent unraveling to Pareto-dominated equilibria seems to occur almost invariably in large groups of 14 to 16 subjects. The speedy unraveling toward Pareto-dominated equilibria (and in fact the lowest ranked of the Pareto-ranked equilibria) has been – ceteris paribus – reproduced without fail (Knez and Camerer, 1994; Berninghaus and Ehrhart,

1998; Cachon and Camerer, 1996; Chaudhuri et al. 2005; Blume and Ortmann, 2007) and in several of these cases with lower numbers of subjects.

Various ways to improve coordination, relative to these baseline cases, have been recently identified in the lab. For instance, several researchers have demonstrated that the lower the attractiveness of – and the deviation costs from – the secure strategy, the more likely the payoff-dominant equilibrium will be reached. Goeree and Holt (1999) have provided evidence that reducing the cost of deviation from equilibrium in one-shot games with two or three players improves coordination. Berninghaus and Ehrhart (1998) have obtained similar results by increasing the number of rounds. Van Huyck et al. (2007) have enhanced coordination by using a finer grid of actions. Furthermore, a change in the attractiveness of the risky actions has been investigated by manipulating the bonus rate. Brandts and Cooper (2006b) have shown that an increase in financial incentives – namely, an increase in the bonus from coordination on higher-payoff equilibria – is effective in overcoming the coordination failure that is almost invariably reached in initial rounds, even with relatively small groups (four subjects). They also find that subsequent reductions in the bonus have only marginal effects on coordination and small incentives are as effective as large ones. Hamman et al. (2007), on the contrary, have found that substantive incentives are far more effective than nominal ones, and once incentives are removed a speedy unraveling toward the inefficient outcome occurs. [14]

Moreover, there is evidence that communication, competition, pattern of growth, and subject-pool composition affect coordination. Costless pre-play communication has been shown to reduce coordination failure in case of both minimal informational content (Blume and Ortmann, 2007) and open-ended messages (Cooper, 2006; Brandts and Cooper, 2007; Chaudhuri et al. 2005). Bornstein et al. (2002) have observed more efficient coordination by introducing inter-group competition. Weber (2006) has shown that large groups of 12 players can coordinate successfully if the groups

---

[14] The difference in these results might be due to the different financial incentives used; in Brandts and Cooper (2006b) incentives applied to all equilibria, while in Hamman et al. (2007) players obtained the bonus only if they reached a given threshold.

start out small and new players are gradually added in subsequent periods. Information about a group's history is crucial for efficient coordination to be achieved. However, Knez and Camerer (1994) have found that merging two groups of three players significantly lowers the minimum effort even when information about group history is publicly available. Engelmann and Normann (2007) have shown that the subject-pool composition seems to affect coordination in the weak-link game[15].

Informational feedback and precedent transfer, however, do not have an unambiguous effect on coordination. Van Huyck et al. (1990) and Devetag (2005) have found a negative or no effect of full feedback (i.e., information about the full distribution of efforts in previous periods) on coordination, while Berninghaus and Ehrhart (2001) and Brandts and Cooper (2006a) have found a positive effect. Devetag (2005) has found that precedent transfer from a similar game enhances coordination in the minimum game, while no evidence of transfer has been found by Knez and Camerer (2000).

5. Results

It might be useful to recall at this point some terminology and idiosyncrasies of our experiment. We cannot directly observe the cost of effort exerted by our experimental subjects; thus, we use the number of errors as an approximation for effort, the assumption being that, controlling for task learning, lower numbers of errors correspond to higher effort. The term errors refers to the number of mistakes made at the employee-level, while worst performance indicates the highest number of errors made at the firm level by any of the employees (a variable corresponding to the minimum effort in chosen effort experiments). Notice that in our experiment efficient coordination is achieved when all employees make the smallest number of errors possible (i.e., worst performance equals zero), while in previous studies successful coordination was achieved when all employees chose the highest action value.

---

[15] The experiment was run in Denmark and two-thirds of the participants were Danes; the minimum effort increased as the share of Danish subjects in the group increased.

We are mainly interested in coordination failure/success in a repeated game setting: for this reason, and since in the first round learning and precedent effects do not apply, we analyze behavior in the first (and initial) rounds separately from behavior in later rounds.

Repeated observations at the employee level are not independent, and, furthermore, employee-level data also show correlation for employees of the same firm. Therefore, when testing our hypotheses (and if not otherwise specified), our independent unit of observation is the mean value of the variable for a single firm. All tests are two-sided.

5.1. Initial coordination

Is failure to coordinate on the efficient equilibrium observed in the initial round(s) in our real-effort setting? At the firm level, coordination failure was pervasive in the first round: 18 firms out of 24 coordinated on the worst outcome (i.e., worst performance = 4), although an average of 1.8 slots of extra time per firm were bought. Of the remaining 6 firms, 5 achieved a worst performance of 1 or 0 in the first round; however, this good performance was the result of several slots of extra time bought (2.25 on average per firm). Initial behavior, and possibly beliefs, seem not affected by treatments: a Mann-Whitney ranked-sum test finds no statistically significant differences at any conventional level between treatments in round 1 for both worst performance and extra time. A clear tendency toward coordination failure at the firm level is also shown in subsequent rounds. Figure 2 presents the distribution of worst performances over the first four periods, by pooling all Group treatments. In each of the first four rounds, modal worst performance was equal to 4 errors, while average worst performance over rounds 1 to 4 was equal to 2.7 and it was greater or equal to 3 in 14 out of 24 firms.

[Figure 2 about here]

At the employee level, the data are more varied than at the firm level. Performances are heterogeneously distributed in the first four periods in both Group and Individual conditions. Table 5a and 5b report the distributions of errors and extra time, in round 1 and in rounds 1 through 4

respectively, divided by "types" of performers. In the first round, the distribution of errors is concentrated on the extremes of zero and four errors. Although for about 40 percent of the employees, in both the Group and Individual treatments, the number of errors was zero in the first round, the task was calibrated so that high performance was feasible but not trivial. In fact, more than 34 percent of the employees who made zero errors (and more than 66 percent in the Individual conditions) could do so only by buying extra time – a quite risky strategy, considering that it entails a certain loss if not matched by a sufficiently higher performance at the firm level. The Individual condition, as compared to the Group condition, does not affect errors in the first round ($p = 0.821$, two-tailed Mann-Whitney ranked-sum test); however, it does have a significant effect on the number of slots of extra time bought ($p = 0.043$);[16] hence, to the extent that employees are more likely to buy extra time in the Individual than in the Group conditions, behavior in the first round is affected by strategic uncertainty. We note that more extra time does not translate into an improvement in average performance.

In subsequent rounds, the frequency of employees that perform very poorly tends to decrease. However, in rounds 1 through 4, only 31 percent of the employees in the Group conditions managed to make, on average, between 1 and 0 errors without buying extra time. Table 5b also reveals that a non-negligible fraction of the participants (almost 17 percent for both the Group and Individual conditions) performed poorly over the first four periods by making, on average, more than 2 errors. A Mann-Whitney ranked-sum test reveals a statistically significant difference between Individual and Group for both average errors ($p = 0.097$) and extra time ($p = 0.006$) in the first four rounds; specifically, errors are fewer and extra time is higher in the Individual relative to the Group conditions. This suggests that the lower efficiency in Group conditions stems primarily from the externalities imposed on Group members through weakest performance rather than from differences in behavior.

---

[16] In the first round, the single employee has been considered as the independent unit of observation since there was no previous interaction among members of the same firm. We pooled all Group treatments (and all Individual treatments) since we did not find any statistical difference in the number of errors in the first round.

[Table 5 a and b about here]

More generally, our data suggest a positive answer to our first research question: coordination failure is observed in the initial round(s) when real effort is implemented. This is, to some extent, a necessary condition for our real-effort experiment to be of interest, since it demonstrates that the task was not too easy and, hence, coordination was not trivial to achieve.

Result I: We observe both a high dispersion of individual performances and a high frequency of coordination failure over the first four periods of our real-effort experiment.

5.2. Dynamic Analysis: Group Conditions

Is speedy unraveling toward the inefficient equilibrium observed in later rounds, or is it possible to overcome coordination failure by introducing real effort in the lab? How and why exactly does coordination evolve over time? We now turn to a dynamic analysis of firm-level data to answer our second set of research questions, and test whether coordination failure is overcome in our real-effort experiment. Table 6, which summarizes errors, worst performance, and extra time pooled across groups and divided by treatment, provides evidence for a positive answer to the first question. Worst performance in rounds 5 through 8 was lower than in the first four rounds for all conditions (1.85 as compared to 2.70, all treatments pooled). The average worst performance in Parts 2 (rounds 9-12) and 3 (rounds 13-16) was less than 2.1 for all treatments (see bottom panel in Table 6). Importantly, the steady decline of errors and worst performance started well before the bonus rate was increased. That is, initial inefficiency is overcome without any change in the structure of the incentives in 16 out of 24 firms.

[Table 6 about here]

Even though the majority of the firms manage to overcome initial coordination failure and to reach the efficient outcome, coordination is not easy. In fact, 5 firms (unsuccessful firms, hereafter) never reached the payoff-dominant outcome in any of the 16 periods.[17] A Mann-Whitney ranked-

---

[17] Furthermore, 4 out of 5 unsuccessful firms performed very poorly with an average worst performance of 3 or more in Part 2 or 3, or in both.

sum test reveals that the distribution of firms' average worst performance differs for successful and unsuccessful firms (p = 0.001, two-sided).[18] Figure 3 shows worst performance over time both for successful and unsuccessful firms. A non-negligible number of firms that were not able to coordinate successfully provides important evidence that task-learning is not strong enough to render successful coordination trivial for all firms. Indeed, task-learning was heterogeneous across both employees and firms, and such heterogeneity presumably created diversity of beliefs across firms, and heterogeneity in the level of effort exerted.

[Figure 3 about here]

Result IIa: After an initial history of failure, the majority of the firms (19 out of 24) successfully coordinate on the most efficient outcome, and many of them (more than 84 percent) do so before any change in the structure of payoffs is introduced through increased bonuses.

In our real-effort experiment, the majority of the firms overcome initial coordination failure, whereas no similar dynamics have been commonly observed in previous experiments with chosen effort. Towards a possible explanation of how and why coordination failure is overcome, we consider two factors: first, differences between successful and unsuccessful firms in initial rounds; second, individual adjustment dynamics.

[Table 7 about here]

Table 7 reports average worst performance, errors, and extra time in rounds 1 through 4, conditional on firms being successful or unsuccessful. Relative to successful firms, unsuccessful firms performed poorly in initial rounds. However, if only best performers (i.e., the employee with the smallest number of errors within a firm in a round) are considered, there is no difference between successful and unsuccessful firms in rounds 1 through 4. Hence, unsuccessful firms are more heterogeneous, in that the observed gap between best and worst performers is significantly higher (albeit weakly so) from that observed in successful firms (p = 0.085, Mann-Whitney ranked-sum test, two-tailed).

---

[18] Successful firms achieved a worst performance of zero in at least one round, while unsuccessful firms did not.

To gain some insight into the individual adjustment dynamics, we apply a simple adaptation rule derived from Learning Direction Theory (Selten and Stoecker, 1986; Berninghaus and Ehrhart, 1998). According to this qualitative theory, players adjust their behavior, on the basis of ex-post reasoning, in the direction of the action that would have been a best reply to others' behavior in the previous round. We apply the following rules adapted to our game from Berninghaus and Ehrhart (1998):

– Rule 1: when an employee "is the minimum" in a period, i.e. when his number of errors is the highest in the firm, he does not increase his errors in the next period;

– Rule 2: when an employee "is the non-minimum" in a period, i.e., when her number of errors is lower than the worst performance, she does not decrease her errors in the next period. [19]

Previous research on chosen effort weak-link games has shown that non-minimum employees (i.e., employees who overshot relative to the firm minimum) tend to retreat to lower action choices after a few rounds out of equilibrium, while minimum employees often do not increase their action choices. This robust mechanism draws both individual effort and minimum effort down toward the worst outcome and, subsequently undermines the possibility of escaping from the coordination failure trap. Our data reveals an interesting behavioral regularity in the opposite direction: in successful firms, non-minimum employees held steady for several periods eventually drawing firms out of initial inefficiency.

[Table 8 about here]

Indeed, as shown in Table 8, Rule 1 captures well the behavior of minimum employees, as it explains more than 80 percent of the observed choices, while Rule 2 falls short at capturing non-minimum employees behavior: in fact, almost 66 percent of the observed behavior is in the opposite direction. Interestingly, minimum and non-minimum employees do not differ substantially in their likelihood to reduce or increase errors.

---

[19] Since changes from the two extremes of zero and four errors can only occur in one direction, the results for these extreme cases are reported separately.

Learning Direction Theory only states the direction of change, and not the reasons for, or the magnitude of such change. However, average errors made and extra time bought by both types of employees (minimum and non-minimum) can partly overcome this shortcoming and shed light on the differences occurring between successful and unsuccessful firms. Specifically, in Table 9 we report the average number of errors and the average extra time bought in round *r* separately for minimum and non-minimum employees in round *r-1*.

[Table 9 about here]

Both minimum and non-minimum employees in unsuccessful firms perform worse than their counterparts in successful firms. Furthermore, non-minimum employees in successful firms buy extra time more than twice as often as compared to those in unsuccessful firms. The implications are twofold: on the one hand, in unsuccessful firms, strategic considerations induced non-minimum employees not to ask for extra time and possibly not to perform the task at their best; on the other, in successful firms, non-minimum employees were highly motivated not to give up, given that they bought a higher, although not statistically significant, amount of extra time.

Learning Direction Theory fares poorly at explaining how and why coordination failure is overcome, while the presence of "leaders" that tend to hold steady better explains the observed behavior in successful firms.

Result IIb: Best performers in successful firms, apparently in an attempt to overcome a history of coordination failure, tend to hold steady by committing few errors and buying a fairly high amount of extra time. In contrast, unsuccessful firms are characterized by a number of employees performing very poorly and, as a consequence, by non-minimum employees giving up after some periods.

5.3. Dynamic Analysis: Group vs. Individual Conditions

Does strategic uncertainty undermine the learning process in the counting task? We now turn to the comparison between Group and Individual conditions to answer our third set of research

20

questions on the interaction between strategic and task learning. Indeed, since strategic learning is not present in Individual conditions, task learning alone is responsible for variations in the number of errors, Individual treatments indicate to what extent strategic uncertainty affects employees' performance in Group treatments. In particular, we expect the decline in the number of errors to be slower in Group as compared to Individual if negative expectations develop within a firm. Figure 4 shows average extra time and errors over time for both Individual and Group conditions, pooled by treatments. An overall downward tendency is evident in both the Group and Individual conditions and suggests that task learning was similar across conditions. In particular, errors are not significantly different between Individual and Group conditions; the parallel decline in the number of errors and behavior over time is interesting given the extremely unforgiving nature of the weak-link game. Moreover, extra time is lower in Group as compared to Individual conditions, hence suggesting that employees in the Group conditions may have worked even harder than those in the Individual conditions, since they performed the same task (with the same level of precision) in a shorter time.

[Figure 4 about here]

Furthermore, successful firms performed better than employees playing alone, although the difference is not significant at any conventional level (p=0.889, two-sided Mann-Whitney ranked-sum test), while the opposite was true, and statistically in a significant manner (p=0.006, two-sided), for unsuccessful firms. In particular, the number of errors over time declined more slowly in unsuccessful firms as compared to Individual conditions. Consequently, repeated failure to coordinate on higher-payoff outcomes may have lead employees to exert lower levels of effort relative to Individual conditions and thus may have interfered with task learning.

Result III: Individual and Group conditions follow similar patterns of learning and both initial and overall performance are largely indistinguishable between the two conditions; that is, at the aggregate level, the process of task learning is not undermined or "slowed down" by strategic uncertainty.

In the previous section we presented results from the Learning Direction Theory for Group data; the right column of Table 8 presents a similar analysis for the Individual conditions. Since no distinction between minimum and non-minimum employees is possible (i.e., each firm is composed of only one employee), changes in behavior are classified only according to the number of errors (Errors=0, Errors=4, and 0<Errors<4).[20] A comparison between Individual and Group behavior reveals that the direction of change was almost the same in the two conditions, although it is worth noticing a difference regarding the behavior of the worst performers, i.e., of employees making 4 errors in a given round. In the Individual conditions, only 52 percent of the employees who made 4 errors in a round made fewer errors in the following round, while more than 80 percent of the employees in the Group conditions did so. It appears that a form of peer pressure (or maybe competition) induced very poor performers (i.e., employees committing four errors or more) to try to improve their performance in the following round more often than the corresponding employees who were playing alone. While behavior in successful firms and Individual treatments follows similar adaptation dynamics, employees in unsuccessful firms are less likely to adjust their behavior in the direction of efficiency than employees playing alone.[21]

5.4. Effectiveness of the Bonus

Does an increased bonus enhance coordination? In the period immediately before the bonus increase, few firms were still trapped in coordination failure; as a consequence, the increase in the bonus rate did not translate into an abrupt fall in the number of errors[22]. As expected, only a few

---

[20] Observations from the latter range of errors have to be compared with both minimum and non-minimum players.

[21] The differences between Individual and successful firms are not significant at any conventional level. The differences in the percentages of employees that decreased their errors between Individual and unsuccessful firms are statistically significant at the 5% level according to a Mann-Whitney ranked-sum test. Only the difference for non-minimum employees that did 0 errors at time *t* is statistically significant at the 10% level.

[22] Both the worst performance and the number of errors slightly increased. However, the worsening in firm performance was due to a small number of subjects (14 out of 96) and a Wilcoxon signed-rank test failed to reject the null hypothesis (p=0.5984, two-tailed) that the number of errors were equal in round 8 and 9. Even though our result seems to contradict previous findings, two factors not present in other experiments render the comparison difficult. First, subjects were not trapped in the worst possible outcome. Actually, 55 out of 96 subjects made zero errors immediately before

firms showed an increase in the number of errors at the firm level (5 out of 24) or at the employee level (1 out of 24). Moreover, in 11 out of 24 firms, errors (both at employee and firm level) increased or remained constant after the bonus was decreased; in all the other groups the worst performance and the errors declined even when the bonus was decreased. Extra time slightly increased from the last four rounds of Part 1 to Part 2 (from 0.20 to 0.22), after the high bonus was introduced,[23] and decreased to 0.14 in the last Part of the experiment.

In order to disentangle the relative contribution of bonus and experience, we use the Individual conditions in which the bonus was kept fixed (I50Fixed and I20Fixed) as a benchmark. Figure 7 shows no effect of the increase in the bonus rate on errors.

However, unsuccessful firms show a tendency to decrease their errors in Part 2, even though there is some variability in the trend.

[Figure 5 about here]

Result IV: in 11 out of 24 firms the bonus had the expected effect (i.e., errors declined in Part 2 and slightly increased in Part 3), while in 12 of the remaining firms the number of errors declined even when the bonus rate was decreased. Given the individual behavior in the fixed-bonus condition, the evidence suggests that the bonus increase did not significantly enhance coordination.

5.5. Effectiveness of Informational Feedback and Importance of Deviation Costs

Do modes of information feedback and size of deviation costs affect errors and coordination? Different information feedback did not result in statistically significant differences in errors and worst performance in the high-deviation cost treatments (F50 and P50), whereas, per Mann-Whitney ranked-sum test, the low-deviation cost treatments (F20 and P20) are significantly

_____

and after the bonus was introduced. Second, in our setting, a sharp decline in the number of errors was subject to individual abilities constraints.

[23] The increase in the number of extra-time in Part 2 was for the major part due to the Partial20 treatment: the average extra-time bought in round 9-12 was 0.44 while in the four rounds before it was 0.28; interestingly, the number of errors and the worst performance in round 9-12 was the same of the one in round 5-8.

different in terms of both errors (p= 0.036, two-tailed) and worst performance (p= 0.076, two-tailed), suggesting that partial feedback is efficiency-enhancing for low-deviation costs.

As predicted, different deviation costs have an impact on the number of slots of extra time acquired; a pairwise comparison reveals that the amount of extra time bought differs significantly between the low- and high-deviation cost treatments (p=0.0684, Mann-Whitney two-tailed ranked-sum test). However, the higher amount of extra time bought in the low-deviation cost treatments did not translate into any statistically significant difference in either the average number of errors or the average worst performance according to a Mann-Whitney two-tailed ranked sum test.

Result V: Partial feedback was efficiency-enhancing (in term of errors and worst performance) only in treatments with low deviation costs. Low deviation costs resulted in a higher amount of extra time bought.

6. Discussion and conclusion

We have explored whether (1) initial failure to coordinate and (2) subsequent unraveling in weak-link experiments survives the use of real effort. While we observe frequent coordination failure in the initial round(s), the answer to (2) seems to be largely negative: our results are, indeed, in stark contrast to previous chosen-effort experiments in that the majority of our laboratory firms were able to reduce their errors and worst performance over time, thus achieving high levels of coordination. Also in contrast to previous lab evidence on the weak-link game, many firms overcame initial inefficiency without payoff changes that are known to be efficiency-enhancing.[24]

We have investigated how task and strategic learning interact, and we have found that both task learning (which pertains to practice) and strategic interaction (which pertains to employees' expectations) are the two forces that seem to increase performance over time. The comparison between our Group and Individual treatments leads us to conclude that successful coordination was

---

[24] In previous experiments, a similar pattern has only been observed in Engelmann and Normann (2007).

not determined solely by task learning, but also by the effort exerted, which did not seem to decline over time in contrast to what has been observed in previous chosen- effort experiments.

We have analyzed how and why successful coordination was achieved; coordination failure was overcome mainly because of employees who, in an apparent attempt to help their firm overcome coordination failure, held steady even when it was costly to them because other members of their firm exhibited a poorer performance. This behavior is inconsistent with adaptation but recalls "signaling" and "strategic teaching" dynamics (e.g., Camerer et al., 2002; 2003), which had been observed in previous chosen effort weak-link experiments where the combination of strong leaders and responsive followers helped overcome initial coordination failure in some cases (e.g., Brandts and Cooper, 2006b; Brandts et. al., 2007). However, in the previous experiments "strategic teaching" was only observed under specific conditions (i.e., increased bonuses) and with full information feedback, which allowed players to directly signal better strategies to others through their choices of actions. In our case, this behavior is much more pervasive, suggesting that other forces might be at work. It may be that subjects tried to perform at their best instead of choosing effort strategically out of some sense of competitiveness that may exist in groups engaged in real-effort tasks (but not in subjects playing alone). A related explanation would be the existence of a social norm – again, triggered by the real-effort context – of "working" (rather than "shirking") when the whole group's outcome depends on each member (Akerlof 1982). Both factors may account for the fact that extremely poor performers in the Group conditions (i.e., subjects making four or more errors) improved their performance in the next round with a significantly higher frequency than in the Individual condition, and this finding holds for both successful and unsuccessful firms. A further explanation is that the behavior of other players in a real-effort task is more predictable than in a chosen-effort task: as a consequence, the strategic uncertainty that Van Huyck et al. (1990) identified as the major reason for the speedy unraveling toward the worst equilibrium was reduced in our experiment. We are not able to disentangle the relative contributions of these three explanations through our experiment,.

Our results do confirm that expectations and strategic considerations play also a role in real-effort experiments. Indeed, the behavior of participants in unsuccessful groups shows that negative expectations caused by repeated failure to coordinate affected task learning, slowing it down relative to the Individual condition and hence presumably causing a decrease in the level of effort exerted, although we note that even in unsuccessful groups most employees kept exerting a positive effort level throughout the experiment. Hence, we conclude that, despite the extreme task interdependence and highly unforgiving payoff structure of the weak-link game, "virtuous" dynamics emerging at the group level can produce an average performance that is not lower (and in some cases even higher) than the average performance that the group members could produce if they were paid individually. To what extent these virtuous dynamics are triggered by strategic teaching, competition, social norms, or other motivations will be investigated in further experiments.

Finally, we have explored the effect of increased bonuses, deviation costs, and informational feedback. Successful coordination is reached by the majority of firms well before any change in the payoffs structure (i.e., increased bonuses), rendering the variation in the bonus rate of lesser importance than in other contexts (e.g., Brandts and Cooper, 2006b). Notably, a temporary increase in the bonuses has no overall effect, whilst it has a positive effect when only unsuccessful firms are considered. That is, an increased incentive may have a positive, even though transient, effect on poorly performing groups. The effect of deviation costs goes in the predicted direction, in that less extra time is asked when deviation costs are higher; errors and worst performance, however, are not positively affected by higher levels of extra time. This behavior suggests, on the one hand, that employees understood well that the action of buying extra time is risky and especially so when deviation costs are high and, on the other hand, that the challenging nature of the task served as an incentive to work harder. Contrary to what Berninghaus and Ehrhart (2001) and Brandts and Cooper (2206a) have found, partial feedback was efficiency-enhancing, but only in treatments with low deviation costs.

Our results are in line with some empirical evidence on actual organizations in which group-oriented payment schemes work as good as, and in some cases outperform, individual-based payment (Pfeffer, 1998). Further experiments employing real effort might help shed some light on the reasons why this happens.

References

Akerlof, G.*(1982)*. Labor Contracts as Partial Gift Exchange, *The Quarterly Journal of Economics*,, 97, 4, 543-69.

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

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

Blume, A. and Ortmann, A. (2007). The Effects of Costless Pre-Play Communication: Experimental Evidence from Games with Pareto-Ranked Equilibria. *Journal of Economic Theory*, 132, 1, 274-290.

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

Brandts, J. and Cooper, D. J. (2006a). Observability and overcoming coordination failure in organizations: An experimental study. *Experimental Economics*, 9, 407-423.

Brandts, J. and 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, 669-693.

Brandts, J. and Cooper, D. J. (2007). It's what you say not what you pay. An experimental study of manager-employee relationship in overcoming coordination failure. *Journal of the European Economic Association*, 4.

Brandts, J., Cooper, D. J., and Fatas, E. (2007). Leadership and Overcoming Coordination Failure with Asymmetric Costs. *Experimental Economics*, 10, 269-284.

Cachon, G.P. and Camerer, C. (1996). Loss-Avoidance and Forward Induction in Experimental Coordination Games *The Quarterly Journal of Economics*, 111, 1, 165-194.

Camerer, C. (2003). *Behavioral Game Theory. Experiments in Strategic Interaction.* Princeton: Princeton University Press.

Camerer, C., Ho, T., and Chong, K. (2002). Sophisticated EWA learning and strategic teaching in repeated games. *Journal of Economic Theory*, 104, 1, 137-188.

Camerer, C., Ho, T., and Chong, K. (2003). Models of thinking, learning and teaching in games. *American Economic Review*, 93, 2, 192-195.

Chaudhuri, A., Schotter, A. and Sopher, B. (2005). Talking ourselves to efficiency: Coordination in intergenerational minimum effort game with private, almost common and common knowledge of advice. Mimeo.

Cooper, D. J. (2006). Are Experienced Managers Expert at Overcoming Coordination Failure? *Advances in Economic Analysis & Policy*, 6, 2, 1-30.

Cooper, R.W. (1999). *Coordination Games: Complementarities and Macroeconomics*, Cambridge University Press.

Davis, D.D. and Holt, C.A. (1993), *Experimental Economics*. Princeton: Princeton University Press.

Devetag, G. (2005). Precedent transfer in coordination games: An experiment. *Experimental Economics*, 6, 53-73.

Devetag, G. and Ortmann, A. (2007a). Classic Coordination Failures Revisited: The Effects of Deviation Costs and Loss Avoidance. CERGE-EI Working Paper, n.327.

Devetag, G. and Ortmann, A. (2007b). When and Why? A Critical Survey on Coordination Failure in the Laboratory. *Experimental Economics*, 10, 331-344.

Engelmann, D. and Normann, H-T., (2007). Maximum Effort in the Minimum-Effort Game. Mimeo.

García-Gallego, A., Georgantzís, N., and Jaramillo-Gutiérrez, A. (2008). Ultimatum salary bargaining with real effort. *Economics Letters*, 98, 78-83.

Gneezy, U. and List, J. A. (2006). Putting behavioral economics to work testing for gift exchange in labor markets using field experiments. *Econometrica*. 74, 5, 1365-1384.

Goeree, J.K., and Holt, C.A. (2005). An experimental study of costly coordination. *Games and Economic Behavior*, 51, 349-364.

Hamman, J., Rick, S., and Weber, R. (2007). Solving Coordination Failure With "All-Or-None" Group-Level Incentives. *Experimental Economics*, 10, 285-303.

Hennig-Schmidt, H., Rockenbach, B. and Sadrieh, A. (2005). In search of workers real effort reciprocity: A field and a laboratory experiment. Working Paper, GESY Discussion Paper, No 55.

Knez, M. and Camerer, C. (1994a). Creating Expectational Assets In The Laboratory: Coordination In 'Weakest-Link' Games. *Strategic Management Journal*, 15, 101-119.

Knez, M. and Camerer, C. (1994b). Increasing cooperation in prisoner dilemma by establishing a precedent of efficiency in coordination games. *Organizational Behavior and Humane Decision Process*, 82, 194-216

Knez, M. and Simester, D. (2001). Form-Wide Incentives and Mutual Monitoring at Continental Airlines. *Journal of Labor Economics*, 19, 4, 743-772.

Lawrence, P. R. and Lorsch, J. W. (1967). *Organization and environment: managing differentiation and integration.* Homewood, Ill.: Irwin.

Levitt, S. D. and List, J. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? .*Journal of Economic Perspectives*, 21, 2, 153-174.

List, J. (2006). The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions. *Journal of Political Economy*, 114, 1, 1-37.

Malone, T.W. and Crowston, K.G. (1991). Toward an interdisciplinary theory of coordination. Tech. Rep. no. 120. Massachusetts Institute of Technology, Center for Coordination Science, Cambridge, Mass.

Malone, T.W. and Crowston, K.G. (1990). What is coordination theory and how can it help design cooperative work systems? *Proceedings of the 1990 ACM conference on Computer-supported cooperative work.* , p.357-370, October 07-10, 1990, Los Angeles, California, United States.

March, J. G. and Simon, H. A. (1958). *Organizations*. New York: Wiley.

Pfeffer, J. (1998). Six dangerous myths about pay. *Harvard Business Review*, 76(3), 109-119.

Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge, Massachusetts: Harvard University Press.

Selten, R., and Stoecker, R. (1986). End Behavior in Sequences of Finite Prisoner's Dilemma Supergames. *Journal of Economic Behavior and Organization,* 7, 47-70.

van Dijk, F., Sonnemans, J., and van Winden, F. (2001). Incentive system in a real effort experiment. *European Economic Review*, 45, 187-214.

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

Van Huyck, J. B., Battalio, R. C., and Rankin, F.W. (2007). Evidence on learning in coordination games. *Experimental Economics*, 10, 205-220

Weber, R. (2006). Managing Growth to Achieve Efficient Coordination in Large Groups. *American Economic Review*, 96, 114-126.

Appendix

Instructions (Part 1)[25]

Instructions for Full Information and High Deviation Costs
Group condition: *(italic in brackets)*
Individual condition: [<mark>highlighted in yellow and in squared brackets</mark>]

<u>General information</u>: The purpose of this experiment is to study how people make decisions. [<mark>The experiment is strictly individual</mark>] For having shown up on time you have earned 2 Euros. From now on, and until the end of the experiment, any communication with other participants is not allowed. If you have a question, please raise your hand and one of the experimenters will come to your desk to answer it.

You will be able to earn money in the experiment. All the money you earn during the experiment is expressed in Experimental Currency Units, or ECUs, that will be converted in Euros at an exchange rate of one Euro for 333 ECUs. Upon completion of the experiment the amount that you earned will be paid to you in cash. Payments are confidential; no other participant will be told the amount you earned.

<u>Parts, rounds (*and groups*)</u>: This experiment will have 3 parts. In Part 1 there will be 8 rounds. After these 8 rounds have finished, we will give you instructions for Part 2 of the experiment. *(Before the experiment proper starts, the software will randomly match you with three other participants in this room. The composition of the groups, made of 4 participants each, will be the same in all rounds.)*

The following instructions are about Part 1 only.

<u>Description of the task in Part 1 of the Experiment</u>: You *(and the other members of your group)* are [<mark>an</mark>] employee*(s)* of a firm, and you can think of a round of the experiment as being a workweek.

In each round:
- *(each of the employees is)* [<mark>you are</mark>] assigned a task to be completed. *(The task, that will be described later on, is the same for all four employees of the firm.)* Specifically, you will receive a small bag containing 4 different types of coins worth 1, 2, 5, and 10 cents, respectively; your task consists of correctly counting how many coins of each type are contained in the bag you were given;
- you will be endowed with a fixed amount of regular time (45 seconds) to complete the task. This regular time will pay you a fixed wage of 200 ECUs regardless of your performance;
- you may choose to buy extra time to complete the task if necessary. Extra time gives you more time to complete the task but entails a cost that you will be charged for. Since you can earn additional money depending on [<mark>your performance</mark>] *(the performance of your group (including yourself)),* it might be profitable for you to pay for having extra time. When exactly that will be the case is explained next.

<u>Performance measurement</u>: *(For each employee, the number of mistakes is the sum of mistakes that he or she has made in counting each type of coins. The firm's performance and your earnings depend however, upon the highest number of mistakes a member of your group (including you) has made. From here on we refer to this as **the worst performance in the group**.)* [<mark>Your performance and your earnings depend upon the number of your counting mistakes, that is, the sum of mistakes that you have made in counting each type of coins.</mark>]

---

[25] This is a translation of the original instructions which were in Italian.

Payoffs: The payoff you will receive in a round depends on your performance (*and on the performance of the other members of your group (and in particular the worst performance in the group)*) according to the following payoff table (all values are expressed in ECUs):

		[My performance (number of errors)] *(Highest number of mistakes in my firm (number of errors))*				
		0	**1**	**2**	**3**	**4 or >**
Slots of Extra Time I have bought	**0**	440	*380*	*320*	*260*	*200*
	1	390	*330*	*270*	*210*	*150*
	2	340	280	*220*	*160*	*100*
	3	290	230	170	*110*	*50*
	4	240	180	120	60	*0*

In each round, you are guaranteed a fixed wage of 200 ECUs (unless you buy slots of extra time about which more below). You will also be paid a bonus that depends on [your] *(the worst)* performance *(in the group)*. The maximum bonus for Part 1 of the experiment is 240 ECUs. Thus, if *(none of the employees of the firm including)* you [do not] make any mistakes you can earn a maximum of 440 ECUs .

Focus initially on the first row of the payoff table (i.e., Slot of Extra Time I have bought = 0): in that row the *(highest)* number of [your] mistakes *(in your firm)* is reported in the columns of the payoff table: if the *(highest)* number of mistakes is greater or equal to 4, the bonus will be zero. Please, note that you are guaranteed a fixed pay of 200 ECU in each round (unless you buy slots of extra time about which more below), no matter how bad your performance is, or even if you do nothing. *(For example, if you and two other employees of the firm have made 0 mistakes and the fourth employee of your firm has made 4 (either because s/he has actually committed four or more counting mistakes, or because s/he has chosen not to perform the task), the highest number of mistakes in your firm is 4 (the last column of the matrix).)*

Your payoff also depends on the extra time that you buy. You can buy up to 4 slots of extra time of 12 seconds in each round; each slot of extra time will cost you 50 ECUs. You will incur a cost if and only if you ask for extra time (*(extra time asked for by the other employees of your firm will be paid by them))*. Rows 2 – 5 break out the payoffs for the various contingencies. Note that your payoff depends on how many slots of extra time you buy but also, IMPORTANTLY, on *(the worst)* [your] performance *(in the group (the highest number of mistakes may be yours or that of another member of the group.).)*

For example, if you buy 2 slots of extra time, the third row gives you the relevant payoffs. For instance, [if your number of mistakes] (*if the highest number of mistakes in your group*) is 4 (*(which is not necessarily the number of your errors))* and you have bought 2 slots of extra time, your earnings will be shown in the third row and the last column, yielding a payoff of 100 ECUs.

On your desk you can find a paper copy of the payoff table in order to have the possibility to easily check it whenever you want. Please, make sure you understand how to read the payoff matrix since your earnings are based on it. If you have doubts that you understand how to read the payoff table please raise your hand now.

Practical implementation:
Before each round starts you have to wear the headphones you have on your desk.
Before each round starts, you will receive a small bag containing coins of four different types: 1, 2, 5, and 10 cents. In general, the total number of coins contained in the bag varies from round to round while the types of coins (1, 2, 5, and 10 cents) will be the same throughout the experiment Your counting mistakes are given by the sum of the difference between the correct number of coins of each type and the number that you report.

For each round of the experiment, a screen like the one shown below will be displayed on your monitor. A sound will inform you that the round has started and you can start performing the task.
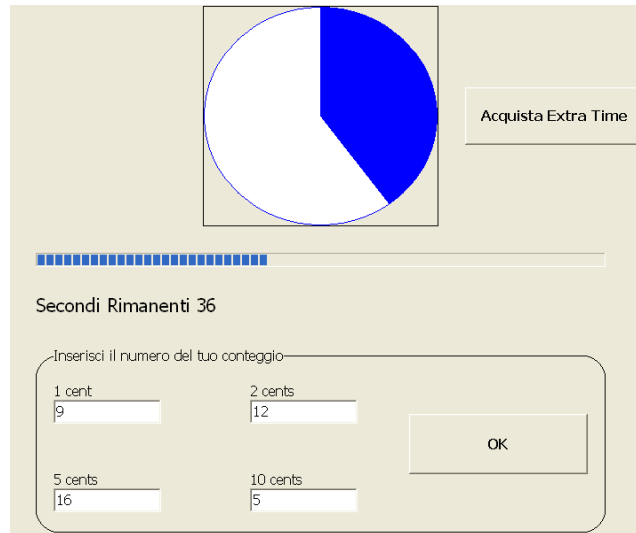
You are allowed to open the bag with the coins only after the round has started. Once you have opened your bag you can pour the coins in the tray on your desk (make sure to remove all objects from the bag).

Recall that the task consists of counting how many coins of each type are contained in the bag you were given for that round.

Please, pay attention to the screenshot:
the four cells in the bottom part of the screen have to be filled with the number of coins you have counted. Each cell corresponds to one of the types of coins you have in your bag. You may change the entered numbers as many times as you like, but once you click on *'OK'* your choice for that round is final. WHEN THE TIME (REGULAR AND EXTRA TIME YOU MIGHT BUY) IS OVER, THE SCREEN WILL DISAPPEAR AND YOU WILL NOT BE ALLOWED TO ENTER THE NUMBERS. In the case that you have not filled the cells with numbers, or you have NOT CLICKED ON 'OK', all cells will be automatically filled with zero. This will translate automatically into *(the firm)* [you] having 4 or more mistakes. *(Please recall that your earnings do not depend on your performance only, but they depend on* **the worst performance in the group***.)*

During the round you can, at any time, buy extra time by clicking on *'Acquista Extra Time'* and the time will be automatically added to the timer. Each time you ask for extra time you will incur a cost. In each round you can buy up to four slots of extra time. IF YOU WANT TO BUY EXTRA TIME MAKE SURE TO BUY IT BEFORE THE TIME HAS EXPIRED because once the time is over the program will NOT ALLOW you to buy extra time. A sound will alert you that the time is

about to expire. Once the time is expired the round is finished and you can NEITHER BUY EXTRA TIME NOR INSERT THE NUMBER OF OBJECTS YOU HAVE COUNTED.

Once the round is finished please fill up the small bag with all the objects you have counted. After the end of each round the experimenter will collect your bag and will give you a new one for the next round.

Information that you will receive: After each round the computer will display a screen like the one shown below. The actual number of mistakes you have made, and, in brackets, the extra time you asked for are reported in the second column (Num Errori (sec extratime)). In the third column the (*highest*) number of mistakes *(made by all of the employees in your firm (you included))* [and extra time] is displayed, [but], if the *(highest)* number of mistakes is greater than four, the screen will display 4 and not the actual number of errors, *(and the extra time asked by each of the members of your group will be displayed in brackets. Please recall that your bonus depends on the worst performance in your group (= highest number of mistakes in your firm, as reported in the second column).)* Your payoff for the latest round, and your cumulated payoff is shown in the last two columns.

At the end of the 8 rounds, information about that round will be displayed and a message on the screen will alert you that the first part is over.

Payment: At the end of the experiment you will be paid, in cash, the sum of the payoffs that you will have earned during the experiment plus the show up fee of 2 Euros. The conversion rate is one Euro for 333 ECUs.

To make sure that you have correctly understood how your earnings are computed, please answer to the questions in the anonymous quiz you have on your desk. Once you have completed the questionnaire, please raise your hand and the experimenter will collect it. If mistakes will be detected in the correction of the questionnaire, the experimenter will repeat aloud the instructions and he will answer whatever questions you might have. During the experiment, if any type of communication is detected by the experimenter the session will be immediately concluded and nobody will be paid.
If you have questions, please raise your hand.

Questionnaire

Questionnaire for High Deviation Costs and Full Feedback
Individual condition: [in squared brackets]
Group condition: *(italic in brackets)* and highlighted in grey, when different

Before beginning the experiment, please answer the following questions. This short anonymous questionnaire is to make sure that everybody has correctly understood the instructions and it does not affect your earnings. If you are having trouble answering the questions, please raise your hand and one of the experimenter will come to your desk. After having answered all questions, please raise your hand and one of us will collect the questionnaire.

For all of these questions, make reference to the following payoff matrix:

		[My performance (number of errors)] *(Highest number of mistakes in my firm (number of errors))*				
		0	**1**	**2**	**3**	**4 or >**
Slots of Extra Time I have bought	**0**	440	*380*	*320*	*260*	*200*
	1	390	*330*	*270*	*210*	*150*
	2	340	280	*220*	*160*	*100*
	3	290	230	170	*110*	*50*
	4	240	180	120	60	*0*

The first question is an example and will be illustrated by the experimenter:

1. Suppose you have made 2 errors and have bought 1 slot of extra time:

Your payoff is _____

Suppose a firm is composed of employees A, B, C, and D. The following table summarizes their choices in one round:

Player	**Number of mistakes**	**Slots of extra time bought**
A	0	4
B	2	0
C	2	3
D	3	0

The highest number of errors is _____

Player A payoff is _____
Player B payoff is _____
Player C payoff is _____
Player D payoff is _____

Please, answer individually to the following questions:

2. Suppose that your bag contains 10 coins of 1 cent, 12 coins of 2 cents, 13 coins of 5 cents and 5 coins of 10 cents. The result of your counting is the following

The number of your mistakes is   _____

3. Please complete the following table with the missing values

Number of errors	Number of Extra Time	Payoff
0	4	
4	1	
2		270
1		280
1	3	
0	0	

Suppose a firm is composed of employees A, B, C, and D. The following table summarizes their choices in one round:

Player	Number of mistakes	Slots of extra time
A	0	4
B	0	0
C	1	3
D	4	1

Player A payoff is   _____
Player B payoff is   _____
Player C payoff is   _____
Player D payoff is   _____

Suppose you make 0 errors and buy 2 slots of extra time. All the other employees of your firm make 2 errors and buy 0 slots of extra time.

Your payoff is   _____

The other employees' payoff is   _____

Suppose you make 2 errors and buy 3 slots of extra time. The other employees of your firm make 0, 1, and 4 errors and buy, respectively, 3, 0, and 1 slots of extra time.

Your payoff is   _____

4. Suppose the following screen to be displayed on your computer after having played a round.

Round	Num Errori (sec extratime)	Errori (extra time)	Payoff	Payoff cumulato
1	0 (0)	0 (0)	440.00	440.00
2	13 (12)	4 (12)	150.00	590.00

Your number of errors in round 2 is   _____
The number of slots of extra time you bought in round 2 is   _____
Your payoff in round 2 is   _____
Your cumulated payoff is   _____

Suppose the following screen to be displayed on your computer after having played a round.

Round	Num Errori (sec extratime)	Gruppo Errori (extra time)	Payoff	Payoff cumulato
1	6 (0)	1 1 4 4 (0 36 12 0)	200.00	200.00

Your number of mistakes is   _____
The highest number of mistakes in your firm is   _____
The number of slots of extra time you asked for is   _____
The highest number of slots of extra time asked in your group is   _____

5. On your desk you have a bag similar to the ones you will receive during the experiment. Please count how many coins of each type are contained in the bag and fill the blank cell in the following screenshot according to the result of your counting.

Table 1: Brands and Cooper (2006) Payoff Table, *B*=6 and *C*=5

		Minimum Effort by Employees				
		40	30	20	10	0
	40	240	180	120	60	0
	30		230	170	110	50
Effort by Employee *i*	20			220	160	100
	10				210	150
	0					200

Table 2: Payoff Tables

Employee payoff, *B*=60 and *C*=50

		Worst performance in the group (number of errors)				
		0	1	2	3	4 or >
Slots of Extra Time I have bought	0	440	380	320	260	200
	1	390	330	270	210	150
	2	340	280	220	160	100
	3	290	230	170	110	50
	4	240	180	120	60	0

Employee payoff, *B*=100 and *C*=50

		Worst performance in the group (number of errors)				
		0	1	2	3	4 or >
Slots of Extra Time I have bought	0	600	500	400	300	200
	1	550	450	350	250	150
	2	500	400	300	200	100
	3	450	350	250	150	50
	4	400	300	200	100	0

Employee payoff, *B*=60 and *C*=20

		Worst performance in the group (number of errors)				
		0	1	2	3	4 or >
Slots of Extra Time I have bought	0	440	380	320	260	200
	1	420	360	300	240	180
	2	400	340	280	220	160
	3	380	320	260	200	140
	4	360	300	240	180	120

Employee payoff, *B*=100 and *C*=20

		Worst performance in the group (number of errors)				
		0	1	2	3	4 or >
Slots of Extra Time I have bought	0	600	500	400	300	200
	1	580	480	380	280	180
	2	560	460	360	260	160
	3	540	440	340	240	140
	4	520	420	320	220	120

Table 3: Summary of the three parts of the experiment for the two bonus rate schemes

	Bonus Rate	
	Variable	Fixed
Part 1 (1-8)	60	60
Part 2 (9-12)	100	60
Part 3 (13-16)	60	60

Table 4: List of Treatments

	F50	P50	F20	P20	I50Var	I20Var	I50Fixed	I20Fixed
Firm Dimension	Group (4)	Group (4)	Group (4)	Group (4)	Individual (1)	Individual (1)	Individual (1)	Individual (1)
Feedback	Full	Partial	Full	Partial	/	/	/	/
Deviation Costs	High (50)	High (50)	Low (20)	Low (20)	High (50)	Low (20)	High (50)	Low (20)
Bonus	Variable	Variable	Variable	Variable	Variable	Variable	Fixed	Fixed
# of subjects	24	24	24	24	12	12	12	12

Note: Treatments have been varied along the four dimensions reported in the first column; observations for each cell are reported in the bottom line.

Table 5a: Distribution of Errors and Extra Time in Round 1, by Group and Individual Conditions

		Round 1: Distribution of Errors (e_1)					
		$e_1 = 0$	$e_1 = 1$	$e_1 = 2$	$e_1 = 3$	$e_1 = 4$	Total/ Mean
Employees %	Group	39.58	12.50	8.33	2.08	37.50	100
	(Individual)	*(43.75)*	*(8.33)*	*(8.33)*	*(2.08)*	*(37.50)*	*(100)*
Extra Time Slots	Group	0.42	0.42	0.63	1.00	0.42	0.45
	(Individual)	*(0.95)*	*(1.25)*	*(0.75)*	*(1.00)*	*(0.17)*	*(0.67)*

Table 5b: Distribution of Errors and Extra Time Round 1 through 4, by Group and Individual Conditions

		Round 1 through 4: Distribution of Average Errors $(\bar{e}_{1-4})$				
		$\bar{e}_{1-4} \in [0,1]$ Very Good Performers	$\bar{e}_{1-4} \in (1,2]$ Good Performers	$\bar{e}_{1-4} \in (2,3]$ Bad Performers	$\bar{e}_{1-4} \in (3,4]$ Very Bad Performers	Total
Employees %	Group	57.29	26.04	9.38	7.29	100
	(Individual)	*(66.67)*	*(16.67)*	*(10.42)*	*(6.25)*	*(100)*
Extra Time Slots	Group	0.23	0.48	0.25	0.43	0.31
	(Individual)	*(0.70)*	*(0.28)*	*(0.45)*	*(0.75)*	*(0.61)*

Table 6: Errors, Extra Time and Worst Performance by Treatment, all Session Pooled

Errors		Full50	Partial50	Full20	Partial20	Ind50Var	Ind20Var	Ind50Fixed	Ind20Fixed
Round	1-4	1.0	1.7	1.4	0.7	1.0	1.1	1.2	1.0
	5-8	0.6	0.9	0.8	0.4	0.6	0.9	0.8	0.9
	9-12	0.5	0.7	0.5	0.4	0.6	0.3	0.6	0.6
	13-16	0.3	0.5	0.6	0.4	0.7	0.3	0.7	0.8

Extra Time		Full50	Partial50	Full20	Partial20	Ind50Var	Ind20Var	Ind50Fixed	Ind20Fixed
Round	1-4	0.16	0.24	0.29	0.55	0.46	0.56	0.71	0.71
	5-8	0.16	0.11	0.23	0.28	0.04	0.38	0.44	0.42
	9-12	0.17	0.06	0.22	0.44	0.04	0.40	0.25	0.38
	13-16	0.13	0.06	0.14	0.25	0.04	0.19	0.33	0.46

Worst Performance		Full50	Partial50	Full20	Partial20
Round	1-4	2.5	3.2	3.3	1.8
	5-8	1.7	2.1	2.3	1.3
	9-12	1.4	1.8	1.8	1.3
	13-16	0.9	1.3	2.1	1.3

Table 7: Worst Performance, Errors, Extra Time, and Gap (WP-BP) in Rounds 1 through 4, by Successful and Unsuccessful Firms

Rounds 1-4	Successful Firms	Unsuccessful Firms	M-W ranked-sum
Worst Performance	2.42	3.80	p=0.003
Errors	1.01	2.04	p=0.004
Extra Time	1.14	1.60	p=0.107
Gap (WP-BP)	2.28	3.15	p=0.085

Note: all values are computed as the average over rounds 1 through 4. The variable Gap (WP-BP) is computed as the difference between the worst performance (WP) and the best performance (BP) for each firm and for each round.

Table 8: Learning Direction Theory: Individual Adjustments

	Group Data	Successful Firms		% decreased errors Successful – Unsuccessful (Mann-Whitney)	Individual Data
MINIMUM EMPLOYEES		No	Yes		
Errors = 4 (Obs)	(150)	(68)	(82)		(88)
% unchanged errors	24.00%	38.24%	12.20%	26.04%	47.73%
% decreased errors	76.00%	61.76%	87.80%	(p=0.243)	52.27%
Errors <4 & >0 (Obs)	(184)	(31)	(153)		(136)
% increased errors	17.39%	35.48%	13.73%	26.13%	17.42%
% unchanged errors	12.50%	16.13%	11.76%	(p=0.020)	11.36%
% decreased errors	70.11%	48.38%	74.51%		71.21%
NON-MINIMUM EMPLOYEES					
Errors <4 & >0 (Obs)	(139)	(61)	(78)		(136)
% increased errors	16.54%	27.87%	7.44%	22.82%	17.42%
% unchanged errors	17.99%	18.03%	17.95%	(p=0.095)	11.36%
% decreased errors	65.47%	54.10%	76.92%		71.21%
Errors = 0 (Obs)	(591)	(140)	(451)		(500)
% increased errors	24.70%	31.43%	22.61%	8.81%	20.00%
% unchanged errors	75.30%	68.57%	77.38%	(p=0.303)	80.00%

Note: frequencies of changes in the number of errors from round t to round $t+1$ by errors and type of employee (minimum vs. non-minimum). Last column reports frequencies of changes in the number of errors from round t to round $r+1$ in the Individual treatments, all session pooled.

Table 9: Learning Direction Theory, Average Errors and
Extra Time by Type of Player

	Successful Firms		Successful vs. Unsuccessful
	No	Yes	(Mann-Whitney)
Minimum Players in round *r-1*			
Average Error in round *r*	1.86	0.75	p=0.001
Average extra time In round *r*	0.26	0.20	p=0.441
Non-Minimum Players In round *r-1*			
Average Error in round *r*	0.89	0.43	p=0.046
Average Extra Time in round *r*	0.05	0.12	p=0.403

Average errors and extra time in round *r* by type of employees
(minimum/non-minimum) at time *r-1*

Figure 1: Sample Screen for Counting Task
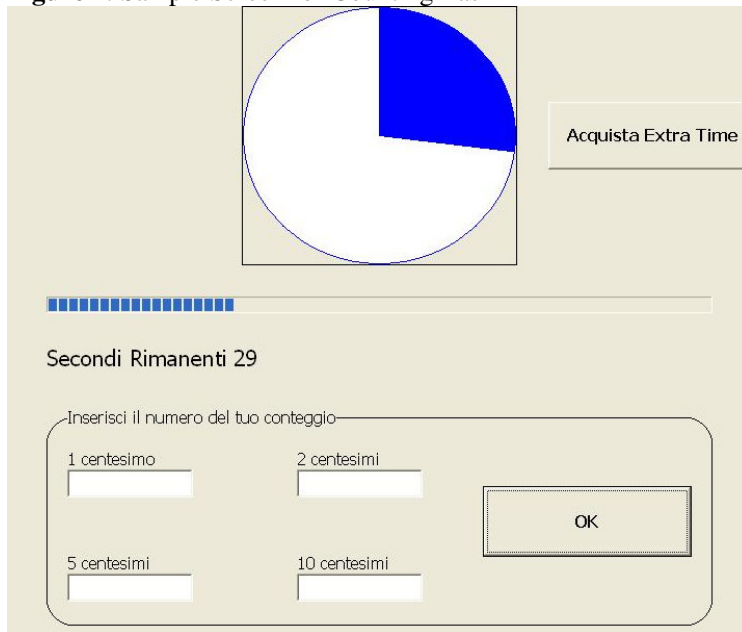
Figure 2: Distribution of the Worst Performance in Rounds 1 tthrough 4, all sessions pooled
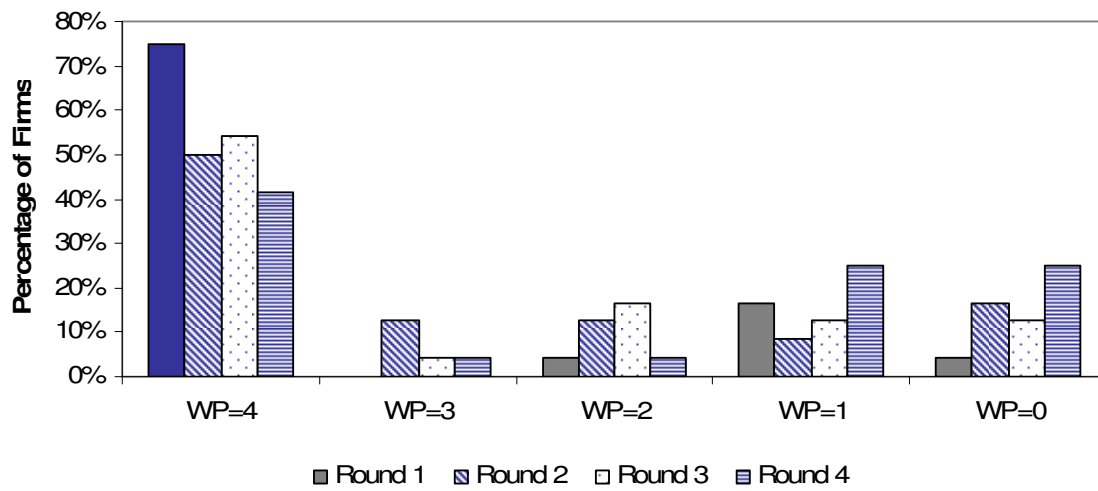
Figure 3: Worst Performance over Time, by Successful and Unsuccessful Firms
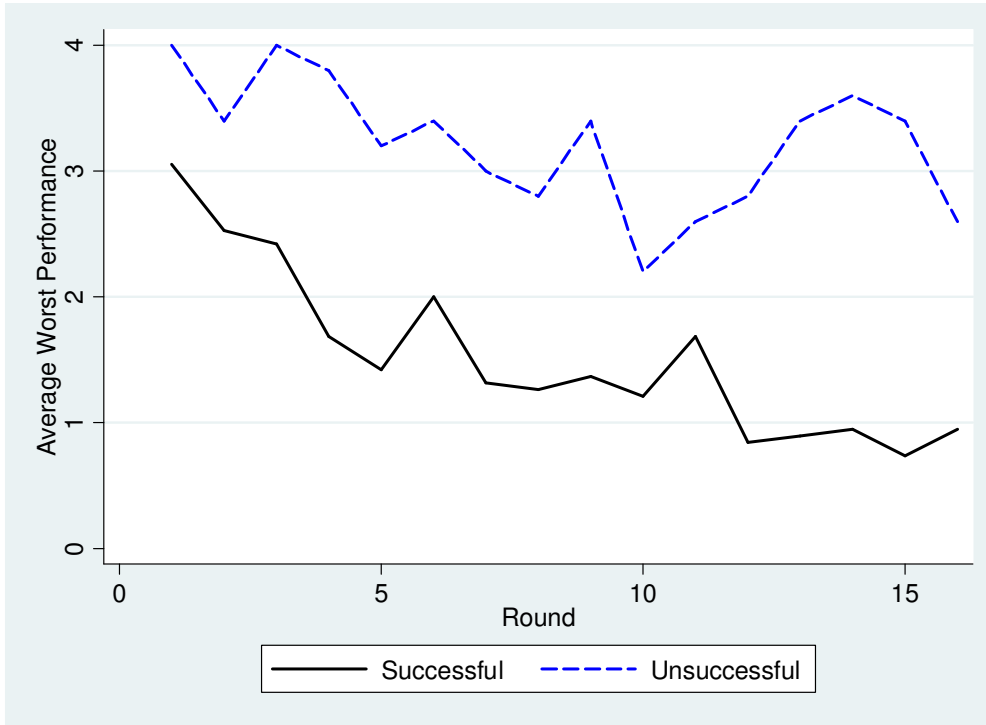

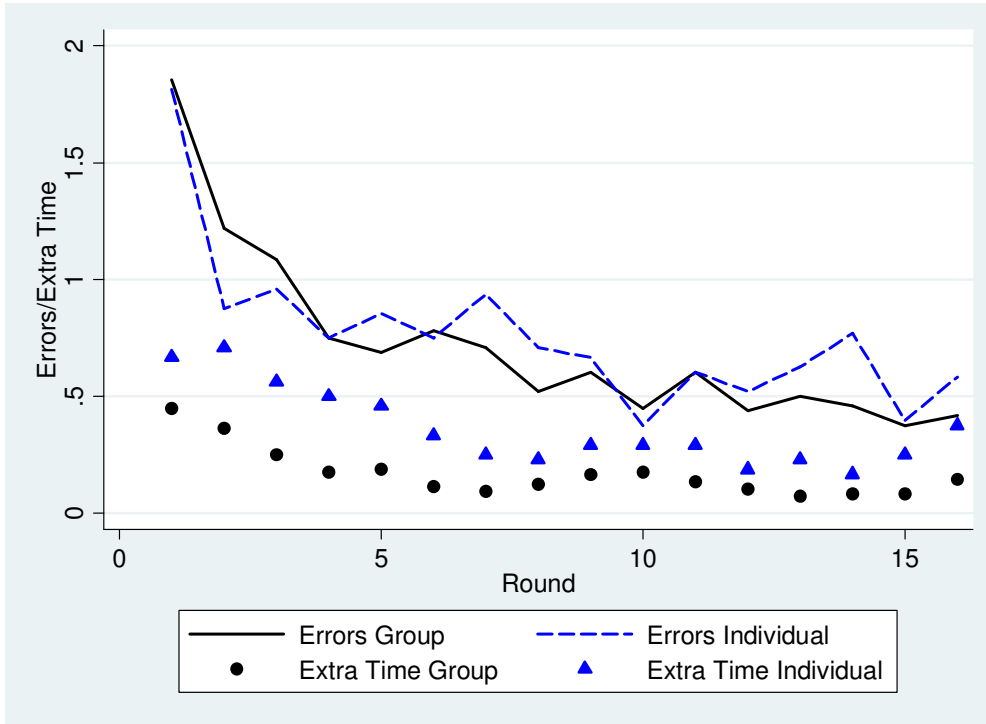**Figure 4**: Extra Time and Errors over Time, by Group and Individual conditions

Figure 5: Comparison between Fixed (Individual) and Variable (Group) Bonus, all treatment pooled