UNIVERSITY OF TRENTO
CIFREM
INTERDEPARTMENTAL CENTRE FOR RESEARCH TRAINING
IN ECONOMICS AND MANAGEMENT

DOCTORAL SCHOOL IN ECONOMICS AND MANAGEMENT

INCENTIVES, GROUP PRIDE, AND REAL EFFORT IN THE WEAK-LINK GAME

AN EXPERIMENTAL ANALYSIS

A DISSERTATION
SUBMITTED TO THE DOCTORAL SCHOOL OF ECONOMICS AND MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DOCTORAL DEGREE
(PH.D.)
IN ECONOMICS AND MANAGEMENT

Stefania Bortolotti

April, 2010

Supervisors

Prof. Giovanna Devetag

Prof. Andreas Ortmann

Referee Committee

Prof. Enrique Fatas

Prof. Luigi Mittone

Doctoral Committee

Prof. Gabriella Berloffa

Prof. Massimo Warglien

Prof. Luca Zarri

Acknowledgments

I would like to express my gratitude to a number of people that supported and influenced my work and made this thesis possible.

First of all, I would like to thank Prof. Andreas Ortmann for introducing me to experimental economics and guiding me through this fascinating field. I owe a great debt of gratitude to Prof. Giovanna Devetag who, together with Andreas, has been a great source of ideas and encouragement over the years. It was a great privilege to have Andreas and Giovanna as supervisors.

This thesis would have never been possible without my dear fellows Andrea, Dominique, and Ivan; thanks for the long nights spent at Cifrem studying and writing, for cheering me up in difficult times, and for being good friends.

A special thanks goes to the Social and Decision Sciences department at CMU. In particular, I wish to express my deepest gratitude to Prof. Roberto Weber for invaluable insights and stimulating conversations and to Prof. George Loewenstein for the excellent course and thoughtful conversations. I owe sincere gratitude to Filippos, Jonathan, Selen, and Rahul for being my family in Pittsburgh.

I am very grateful to Matteo Ploner, Prof. Luigi Mittone, and Marco Tecilla for patiently listening to my ideas and for letting me use the facilities of the lab. I am also grateful to Matteo Bonifacio and Prof. Paolo Collini for introducing me to economic research during my undergraduate studies and for supporting my decision to embark on a Ph.D.

Finally, I would like to thank my parents and my sisters for their love, patience, and unconditional support over the years.

Contents

Introduction

Coordination problems are ubiquitous in actual organizations and other real-world contexts (March and Simon, 1958; Schelling, 1960; Lawrence and Lorsch, 1967; Malone and Crowston, 1990; Cooper, 1999) and are characterized by problems of equilibrium selection. Driving on the left or the right side of the road, choosing between two different technologies characterized by network externalities (e.g., Windows vs. Apple; VHS vs. Betamax) or getting workers of an assembly line to accomplish their task in a timely manner are just some examples. Coordination commonly arises from the need to integrate different interdependent activities. Given the increase in complexity and interdependence of actual organizations, successful coordination among team members, or among different teams is pivotal to sustain competitiveness (e.g., March and Simon, 1958; Schelling, 1960). However, employees often fail to match the actions of co-workers, causing coordination failure (e.g., Heath and Staudenmeyer, 2000; Camerer, 2003; Devetag and Ortmann, 2007), hence calling for the search of effective mechanisms aimed at enhancing efficient integration of specialized activities.

In recent decades, coordination problems have hence been a subject of major concern in various fields of study. Already back in the 1980s, Malone described coordination as an inherently interdisciplinary field drowning upon “a variety of different disciplines including computer science, organization theory, management science, economics, and psychology” (1988: 2). Nevertheless, while a great deal of attention has been paid by organizational scholars to solutions proposed by computer scientists, far less attention has been paid to experimental findings despite the evidence accumulated by economists on the issue. The lack of dialogue between disciplines probably stems from methodological differences. Economists have focused on simple formal theoretical models with high internal consistency but often not empirically grounded; in contrast, organizational scholars have tried to address organizational complexity but have mainly based their research on data collected in the field (e.g., case studies and surveys). Experimentally-based evidence seems to emerge as an important methodological tool to help bridging the gap between insights on organizational design from economics and management since it combines the simplicity of economic theory and the richness of empirical evidence. Despite the potential advantages of controlled experiments (e.g., control, replicability, possibility to establish causal relationships), organizational scholars have mostly focused on experimental shortcomings (e.g., non-representative subject pool and reduced generalizability).

This dissertation illustrates the potential of an experimentally-grounded approach to organizational coordination problems in which the effectiveness of different organizationally-relevant variables is first tested in a controlled setting. Even though an increasing number of economic experiments have studied a host of organizational aspects (for a review, see Camerer and Weber, 2007), an experimentally-grounded approach to organizational design has been neglected for a long time.[1] This is quite surprising given that back in 1969 a special issue of a leading organizational journal, *The Administrative Science* Quarterly, was devoted to organizational experiments. Despite some re-appraisals, the reluctance of organizational scholars toward this methodology marginalized experiments in the organizational literature. Griffin and Kacmar (1991) reported that in a selected sample of leading organizational journals for the time interval between 1986 and 1988 the number

[1] This is especially true for economic experiments, while psychological experiments had been more successful in informing organizational scholars.

of published field studies was seven times as large as the number of published laboratory studies.[2] Research in the area of coordination problems is by no means an exception.

Several reasons may account for the lack of communication between organizational research and research in (experimental) economics. First, the same class of problems is often referred to with different names in the two disciplines (e.g., Weber, 2000), and organizational definitions often leave room for several interpretations which render difficult for economists to formalize the problem at hand. Second, it is difficult, if not impossible, to map the complexity of organizational phenomena into the simple games used in the lab. Mintzberg (1977), among others, claimed that "Management Policy researchers cannot use laboratory effectively, for at the policy level, the very complexity of phenomena determines the organization's behavior" (93). Third, lab experiments are commonly believed to provide for far less generalizability to actual organizations than field studies do. As posited by Chapanis (1976), laboratory experiments "have only very limited relevance for the solution of practical problems" (730, in Dipboye and Flanagan, 1979).

A crucial question for an experimentally-grounded approach to organizations is, thus, if and to what extent results obtained in the lab generalize to other domains, in general, and to actual organizations, in particular. The difference in generalizability (or external validity) between field and experimental data has long been taken for granted. In particular, experiments have been criticized to sacrifice relevance for rigor (e.g., Gordon et al., 1986), and this critique toward laboratory experiments has long been undisputed among organizational scholars, which have left unexplored the potential of organizational experiments. In the following, I discuss external validity of both field and laboratory data, and show that the gap between the two approaches is not necessarily as large as commonly

[2] Virtually none of the experiments was an economic experiment. In the present work we are mostly interested in economic, rather than psychological, experiments; for a discussion of the main differences between the two approaches, see Hertwig and Ortmann (2001).

believed. Drawing on both organizational and experimental literature, I discuss three arguments that can help understand the potential of an experimentally-grounded approach to organizational design. First, organizational field studies are not necessarily more generalizable than lab experiments. Second, some aspects of lab research often criticized share many commonalities with actual organizations. Finally, several strategies to increase the external validity of economics experiments have recently been explored.

Organizational research is currently mostly conducted in actual organizations by observing actual behavior, administrating surveys or collecting data (e.g., number of outputs, turnover, absenteeism). Although it might be tempting to have no reservations about the external validity of these methodologies just because the data are collected “in the wild”, it should be noted that both subjects and situations can vary significantly across different organizations (Dipboye and Flanagan, 1979). For instance, behavioral regularities found in one culture do not need to generalize to other cultures as convincingly demonstrated by an important organizational literature (e.g., Hofstede, 1980). In addition, subjects and situations used in organizational studies are not as representative of the organizational reality as commonly asserted. The assumption that field studies are inherently more generalizable were tested by Dipboye and Flanagan (1979) in a survey of the literature; the authors concluded that “this assumption is oversimplified at best and in several respects is erroneous” (146). Indeed, field studies included in the review used a very biased sample of the workforce: participants in the studies were mostly male professionals or managers. Moreover, field studies mostly relied on self-reported measures and not on performance or actual outcomes which are usually considered the dimension of most interest to organizations.

In a note about organizational experiments, Weick (1977) proposed an interesting way to reinterpret what are commonly intended as limitations and flaws of lab studies. In

Weick's words, "(o)ne of the ironies of laboratory experimentation is that presumed liabilities turn out to become conceptual assets for organizational researchers" (1977: 124). Among the other points, the author noticed that: (1) the desire of subjects to look smart and knowledgeable to the experimenter is no different from the desire of employees to be considered valuable by the employer; (2) only a small portion of subjects' skills are relevant to experiments and so is for actual organizations; (3) the relationship between experimenter and participants closely resemble the relationship between employer and employees; (4) participants are a highly selected and biased sample of the population, as are employees in actual organizations. Although experimental demand effect or unrepresentativeness of the sample are not per se desirable and should be carefully avoided in the experimental practice, it should be noted that they may not necessarily represent an obstacle to generalizability of lab results to the organizational domain.

Given their training as theorists, experimental economists have primarily seen experiments as a clean test of theories. According to Plott (1982), theory-testing experiments need to 'mirror' the theory being tested —which is commonly abstract and simple— rather than the corresponding real-world situation. Smith (1985)'s advice that "(…) experimental and other investigations should not be *confined* to testing formal theories" has recently gained new momentum among experimentalists, and a steady increase in the number of experiments aimed at "searching for facts" or "whispering to the ears of princes" (Roth, 1995) has been observed.[3] The surge of experiments not devoted to test theories has brought about a debate on the generalizability of lab results; indeed, the real world rather than the theory has to be the guidance for designing experiments when searching for empirical regularities or policy advices (Schram, 2005). Probably prompted by this focus shift, a growing number of experimental economists have run experiments with

[3] It should be noted that Charles Plott was among the first advocated a data-first experimental approach.

augmented realism. More realism has been injected both on the subjects and the situation sides by using representative subjects, framed instructions, clerical tasks, and moving out of the lab (for an interesting experiment including all the aforementioned elements, see List, 2006; for a discussion of the literature, see Harrison and List, 2004; Levitt and List, 2007).

To sum up, the experimental methodology has often been discarded among organizational scholar based on the lack of generalizability, a claim that, as discussed in previous paragraphs, is not necessarily as robust as commonly perceived; this is especially true if the new trend among experimentalists to augment realism is taken into account. However, the contribution and relevance of the experimental research should not be judged based on realism and generalizability only (e.g., Dobbins et al., 1988: 282). Half a century ago, Weick proposed a multi-method approach to organizational design and noticed that experiments "single out some issues that have gone unnoticed in the complexity of actual organizations, and the results generate questions that take the aura of causes to be pursued" (1969: 294). Although a productive two-way communication between experimentalists and organizational researchers has by and far not occurred, I believe more integration would prove productive for both disciplines. Indeed, notable examples of successful integration between experiments and other methodologies can be found in other disciplines. For instance, as regards the design of economic institutions, such as markets and auctions, experiments have recently been used with remarkable results (Kagel and Roth, 2000; Roth, 2002; Goeree et al., 2005).

In this dissertation I investigate organizationally-relevant aspects experimentally and by way of a coordination game commonly understood to well describe a host of organizational situations. Both financial incentives and psychological elements of team work are investigated "in vitro". I contribute to the body of evidence and the debate on "experimental organizational economics" by proposing a new experimental testbed with augmented realism thus getting closer to the actual organizations I intend to describe. Specifically, I challenge the critically maintained assumption that “chosen-effort” is a reliable proxy for “real-effort” by introducing a clerical task in the lab. The evidence from my studies also adds to the methodological debate on generalizability of lab results, in general, and reliability of chosen effort, in particular.

Outline of the dissertation

In this dissertation organizationally relevant aspects are tested in controlled coordination experiments with strategic uncertainty. Specifically, I focus on the most prominent workhorse in this literature, the minimum-effort game, and introduce real-effort tasks to help bridge the gap in realism and abstraction between lab experiments and actual organizations. The material in this dissertation is presented in six chapters.

Chapter 1 provides an overview of the coordination problem. The first part of this chapter is devoted to the definition of the problem according to three different perspectives: organizational theory, game theory, and experimental economics. Although an universally agreed-on definition does not yet exists in organizational theory, the need for coordination usually stems from the need to integrate interdependent activities. Game theory and experimental economics share a common definition for this class of problems: coordination games are games with at least two equilibria for which coordinated actions yield higher earnings as compared to non-coordinated actions. Even though organizational theory and economics differ in scope and abstraction, I show, drawing on Weber (2000), that organizational situations are related closely with simple games.

The second part of Chapter 1 reviews the experimental evidence on the minimum-effort or weak-link game. This game features multiple Pareto-ranked pure-strategy Nash equilibria and the toughest task interdependence possible: the minimum effort supplied by a member of a group. The minimum-effort game is commonly believed to capture well organizational situations in which "failure of anyone can threaten the whole and thus the other parts" (Thompson, 1967: 54). Van Huyck et al. (1990) were the first to demonstrate the speedy unraveling toward the minimum effort for this game (i.e., coordination failure), a result which –ceteris paribus– has been replicated consistently. I focus on three organizationally-relevant situations —avoiding coordination failure, overcoming organizational failure, and precedent transfer— and discuss the relative experimental evidence. Specifically, I concentrate on aspects such as communication, competition, group commitment, and financial incentives, of relevance for actual organizations and often studied in the organizational literature. I notice that experiments in organizational economics have grown steadily in recent years, and most of the organizational mechanisms tested in the lab have shown to be efficiency enhancing. In spite of the recent surge of organizational experiments, several topics are still largely unexplored and a list, although necessarily tentative and incomplete, of these topics is provided in the final part of the chapter.

Chapter 2 discusses potential benefits and shortcoming of introducing real-effort tasks in both lab and field controlled experiments. External validity and generalizability of lab results has recently been a subject of major concern; this debate favored the increased use of more realistic experiments, such as framed, artefactual, and field experiments (Harrison and List, 2004). I review experiments introducing more realism by using real-effort, rather than chosen-effort, tasks. Although real and chosen effort have been implicitly assumed to lead to similar results, several factors might lead to differences between the two methods (e.g., intrinsic utility, heterogeneity in task-relevant abilities, task learning, and heuristics triggered by the task). Since disutility associated with effort is not observable to the experimenter, the introduction of real effort poses several important methodological questions: (a) how can disutility from effort be measured?; (b) how do subjects differ in terms of disutility from effort?; (c) are there any utility gains attached to the task at hand?. I consider the aforementioned problems from three different points of view: choice of the clerical task; proxies for effort; task-relevant abilities and attitudes toward the task. Although complete control over the cost of effort is difficult to achieve, I found in the literature a number of interesting designs that seem to do a good job in minimizing possible confounds (e.g., Brüggen and Strobel, 2007; Tomer et al., 2007).

Testing whether chosen effort is a reliable proxy for real effort is central to better understanding to what extent lab results generalize to the real world. Brüggen and Strobel (2007) claimed that “chosen effort is an appropriate way of operationalizing effort in experiments” (233). I reviewed the existing real-effort literature in three classes of games —gift-exchange game, intrinsic motivation, and tournaments— and found that Brüggen and Strobel’ s claim is not robust at least; thus, more evidence is needed to understand whether chosen effort is a reliable proxy for real effort.

Chapter 3 explores whether coordination failure in the minimum-effort game survives real-effort implementations. As discussed in Chapter 1, a host of institutions and mechanisms to enhance coordination have been implemented in the game at hand; however, all experiments conducted so far rely on the critical assumption that the choice of a nominal effort level, or chosen effort, is a reliable proxy for real effort. The main novelty of the present experiment is the introduction of a clerical task in the minimum-effort game. The experimental set-up with augmented realism also represents an important attempt to get closer to the organizational reality we sought to describe; indeed, 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 experimental study reported in this chapter also provides evidence on relative effectiveness of financial incentives; specifically, individual-based and group-based incentives are confronted. Whereas individual-based incentives are by and large considered more effective by practitioners, Pfeffer (1998) classifies this belief as one of the most dangerous myths about pay. Several minimum-effort experiments have studied financial incentives in the lab: group bonuses (e.g., Brandts and Cooper, 2006; Hamman et al., 2007), competition for bonuses both at the group (e.g., Bornstein et al., 2002) and individual level (Fatas et al., 2006). The novelty of our study resides in the use of individual, or piece-rate, incentives; while in a chosen-effort experiment such a comparison would be meaningless, the introduction of a certain degree of uncertainty about the performance and task learning render the treatment of great importance for a better understanding of organizational coordination.

Chapter 4 reports an experimental study on how the motivation derived from being part of an elite group, or "group pride", affects equilibrium selection. I hypothesize that strategic uncertainty hampering coordination can be reduced in elite groups through two psychological elements: first, being selected as a member of an elite group has a strong impact on self-esteem which has long been thought as a strong motivation to act (Salancik, 1977; James, 1890); second, identification with a cohesive group in general, and an elite group in particular, induces high levels of group trust, as suggested by a myriad of studies in social psychology (for a review of Social Identity Theory, see Brown, 2000).

This experiment contributes in several ways to the existing literature. Indeed, despite the surge of evidence on both self-esteem and group trust, to our knowledge none of the previous studies considered coordination games. The study most akin to the present set-up is the one by Chen and Chen (2009) who studied the role of group identity in the minimum-effort game. Unlike Chen and Chen (2009), I consider asymmetric groups (high vs. low status) so to induce different levels of self-esteem and not only group identity. The present study also presents two important methodological novelties. First, self-esteem and group trust are induced by using a real-effort task. Contrary to previous experiments on segregation and Social Identity Theory (SIT, hereafter), subjects' classification is based on the same task used in the stage game, rather than an unrelated and trivial task, hence increasing the external validity of our study. Second, the testbed allows for the study of coordination games with real effort rather than chosen effort, and presents tighter control over the payoff structure and equilibria of the game as compared to our first study (Chapter 3).

Chapter 5 follows up on the study reported in the previous chapter and confronts diverse ways to induce group pride. I extend the experimental testbed by using different clerical tasks to induce group pride; specifically, I vary the degree of similarity between the tasks used for inducing group pride and for the coordination game. The possibility to control for the similarity between different tasks and contexts (segregation and coordination phase) is one of the main contributions of this study. This experimental study adds to the economic literature that found that the magnitude of the in-group bias highly depends upon the salience of the categorization procedure (e.g., Charness et al., 2007; Chen and Chen, 2009). Along this line, I test to what extent pride based on one task can be transferred to another clerical task; indeed, group pride may apply to all those groups that have some (real or imaginary) reasons to be proud of themselves even if these reasons are only loosely related to the task at hand.

Another element of novelty of this study resides in the fact that I control for an additional element of uncertainty induced by real effort; even though actions and strategy space are common knowledge in our study, there is uncertainty about the feasibility of some actions and strategies (that is, some subjects might not be able to correctly complete the task due to low task-relevant abilities). The introduction of different tasks allows to control the relative contribution of this additional uncertainty component on coordination.

Chapter 6 provides a summary of the results reported in this dissertation and outlines the importance of integrating organizational insights into experiments, and vice-versa. I also briefly discuss promising lines of future research.

Chapter 1

Organizational coordination and experimental evidence

1.1. Introduction

In spite of a surge of theoretical models and both lab and field evidence, the coordination problem seems alive and well (Heath and Staudenmeyer, 2000), with differences in methodologies and findings between economics and organizational theories continuing to exist (e.g., Weber, 2006). Although a clear gap in scope and abstraction exists between the economic and organizational approach, several scholars have highlighted the potential advantages of blending together different perspectives and methodologies (see, for instance, Weick, 1969; Camerer and Weber, 2007 for a more recent discussion).

Picking up on this exhortation, I consider how controlled experiments can productively inform organizational theories on coordination, and vice versa (e.g., Knez and Camerer, 1994; Camerer and Weber, 2007). Crossing the boundaries of different disciplines often requires an additional effort to establish common bases and shared definitions of the problems. Coordination problems are no exception: mapping organizational theories into economic definitions is not always straightforward. In the first part of this chapter, I present an overview of the organizational and the economic definitions of coordination; differences and possible commonalities are discussed. In particular, I map a variety of organizational situations characterized by pooled interdependence (Thompson, 1967) into a game theoretical framework; namely, I relate pooled interdependence and what is known as the minimum-effort game, or stag-hunt game, to each other (Weber, 2000). The definition of a common ground would also prove useful to better understand which aspects of the coordination problem are better captured by the experimental, rather than the organizational, approach, and vice versa.

In the second part of the chapter, I survey the experimental evidence on the minimum-effort game (Van Huyck et al., 1990); particular attention is devoted to experiments that studied organizational variables and managerial hypothesis in the lab. Even though organizational insights tend to not to be the focus of economics experiments, a growing number of lab experiments have considered and implemented variable of interests to organizations, such as communication, competition, incentives, and leadership, in coordination games.

1.2. Coordination: organizational, game theoretic, and experimental literature

In this section I discuss different work definitions of coordination, and provide some evidence on how the gap between the different approaches can be bridged.

1.2.1. Coordination in the organizational literature

The need for organizational coordination stems from the specialization process, by and large, undertaken by actual organizations of virtually any dimension (March and Simon, 1958; Schelling, 1960; Thompson, 1967; Lawrence and Lorsch, 1967; Malone and Crowston, 1990; Malone and Crowston, 1991); indeed, task and environment complexity call for the need to decompose activities, on the one hand, and to integrate specialized activities, on the other hand. In spite of the relevance of the issue from both a managerial and academic standpoint, a clear cut and commonly agreed-on definition of organizational coordination does not exist (for a discussion, see Okhuysen and Bechky, 2009). For instance, Simon (1945) defined coordination as "the adoption of all members of a group of the same decision" (8), according to Argote (1982) "coordination involves fitting together the activities of organization members, and the need for it arises from the interdependent nature of the activities that organization members perform" (385), while for Faraj and Xiao (2006) coordination is "a temporal unfolding and contextualized process of input regulation and interaction articulation to realize a collective performance" (1157). However, there seems to be a common understanding that coordination is about *integrating interdependent activities*. Another shared element, although often not explicitly mentioned, is the idea that coordination problems do not bring about conflict of interests, in the sense that everyone prefer to coordinate rather than not.

Despite a surge of evidence on the coordination problem, a general theory of organizational coordination has not yet been developed. Moreover, evidence is commonly based on data from small samples of actual firms, and the lack of a precise work definition of coordination render difficult to understand similarity and differences between the analyzed context and new domains to which we may want to extend the insights.

1.2.2. Coordination in the game theoretic and experimental literature

Unlike organizational literature, game theory provides a clear-cut definition for coordination: coordination games are games which have at least two equilibria and where coordinated actions pay higher payoffs than non-coordinated actions. Since multiple equilibria are present in the game, players do not know the action others will choose and, consequently, are not sure about which equilibrium will be selected; this phenomenon is commonly known in the literature as *strategic uncertainty*. The second distinguishing feature, shared with the organizational literature, of a coordination game is that *no conflicts of interests* among players are present, in that coordination on any of the equilibria is always preferred to miscoordination (e.g., there are no free-riding problems).[1]

Coordination games can be divided into two categories: shared and mixed interests. While no conflicts of interest are present in both shared and mixed motives, in the former class, players' preferences over equilibria coincide, while in the latter players' preferences over equilibria are ranked differently. Consider game *a* (aka matching game) in Figure 1.1, the two equilibria (A,A) and (B,B) are ranked the same way by the two players; in contrast, in game *b* (aka Battle of the Sexes) the column player prefers equilibrium (B,B) over (A,A), while the row player prefers (A,A) over (B,B).

FIGURE 1.1: Examples of different classes of coordination games

Game a

	A	B
A	1,1	0,0
B	0,0	1,1

Game b

	A	B
A	2,1	0,0
B	0,0	1,2

Game c

	A	B
A	2,2	0,1
B	1,0	1,1

[1] It should be noted that subjects can have different preferences over the set of equilibria (i.e., the common interest is only partial).

Shared interests, or pure, coordination games can be further divided into two sub-categories: symmetrical games in which all equilibria have the same payoff (e.g., game *a* in Figure 1.1) and non-symmetrical games that exhibit Pareto-ranked equilibria. Game *c* in Figure 1.1 is a non-symmetrical game in that (A,A) yields higher payoffs as compared to (B,B). Even though coordination on the Pareto-dominant equilibrium might seem trivial at a first glance, strategic uncertainty can hamper efficient coordination; notice that, a sure gain of 1 can be earned by playing B, no matter what the other player will play, while 0 might be earned if you play A and your opponent chooses B.

While standard game theory provides numerous selection criteria for choosing among different equilibria, it does not provide an agreed-on ranking of these criteria, and the equilibrium selection problem "will probably be solved with a healthy dose of observation" (Camerer, 2003: 336). Experimental economists have widely borrowed from game theoretic concepts to construct clean experiment and test theoretical predictions under different institutions. Simple games have the clear advantage that are quite easy to be understood and leave little room for subjective interpretations of the problem at hand. As other classes of games, coordination games of several types have been extensively studied in the lab.

1.2.3. Coordination: bridging the gap between different approaches

Organizational and game theoretical definitions of coordination reported in the previous sections may appear completely unrelated to each other, hence suggesting that little can be done to blend insights from the two disciplines. However, I assert that some organizational theories can be nicely represented by using games that can capture, although in a stylized way, phenomena relevant "in the wild".[2] An interesting attempt to bridge the gap between organizational models and game theory, and a demonstration that games are not a farfetched way to describe organizational phenomena, was provided by Weber (2000; for a similar exercise see also Nanda, 1999). In particular, Weber (2000) posited that the three forms of interdependence (i.e., pooled, sequential, and reciprocal interdependence) proposed in a highly influential work by Thompson (1967) can be well described by three coordination games. For brevity, I only consider one type of integration problem, namely pooled interdependence; for expositional purposes I first provide some organizational examples and then map this examples into organizational and economic definitions.

The take-off of airplanes is a prototypical example of coordination problems (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 relationships between different branches of a bank (Weber, 2000); the bank's reputation is beneficial to each branch, but the maintenance of a good reputation requires each branch to exert some effort. In addition, the bank's reputation critically depend on the worst branch; an unpleasant experience with one branch will spoil the reputation of the bank as a whole, not just of the specific branch. A similar problem of reputation maintenance is common to many businesses under franchise. Other examples are the writing of a grant proposal

[2] That is, the phenomena under study are relevant in real occurring contexts such as actual organizations and not only in controlled lab settings or theoretical models

involving several participants or an edited volume involving several authors (Knez and Camerer, 1994; Weber, 2000).

I now consider the definition of pooled interdependence and link this definition to the previous examples. Pooled interdependence describes situations in which "unless each performs adequately, the total organization is jeopardized: failure of anyone can threaten the whole and thus the other parts" (54). As noted by Weber (2000), pooled interdependence and the aforementioned examples share three main characteristics. First, individual performance by each single actor does not depend on other actors; for instance, the speed in loading luggage is not critically affected by the speed in boarding passengers in the take-off example. The quality of the service in a restaurant does not depend on the quality of the service in other restaurants of the same chain. Second, actions are interdependent at the organizational level: the worst performance will determine the overall outcome. In the take-off example, on-time departure depends upon the slowest operation, no matter how diligently all the other operations have been implemented. Third, the operations do not need to be performed in a certain sequence (as, for instance, in an assembly, line). For instance, loading of luggage, boarding of passengers, and security check can be done simultaneously.

The aforementioned examples and, more in general, pooled interdependence can be captured by non-symmetric coordination games with Pareto-rankable equilibria, like game *c* in Figure 1.1. A game from the same class of problems have been firstly tested in the lab by Van Huyck et al. (1990), and is commonly referred to as minimum-effort or weak-link game. The normal form of this game is defined as follows: *N* players have to simultaneously choose an effort level $e \in \{1, \dots, \overline{e}\}$ and each period is paid according to the following earnings function:

$$\pi\ (e_i, e_{-i}) = a\ [min(e_i, e_{-i})] - \ be_i, \qquad a > b > 0 \tag{1.1}$$

and earnings and strategy space are common knowledge.[3] The parameters are set in such a way that coordinating on any effort level is a pure-strategy Nash equilibrium: in the version used by Van Huyck et al. (1990) the seven equilibria of the game are on the main diagonal of the earnings table (see Figure 1.2).

FIGURE 1.2: Earnings table for the minimum-effort game (Van Huyck et al., 1990: Table A)

		Smallest value of effort in the group						
		7	6	5	4	3	2	1
	7	1.30	1.10	0.90	0.70	0.50	0.30	0.10
	6	_	1.20	1.00	0.80	0.60	0.40	0.20
Your	5	_	_	1.10	0.90	0.70	0.50	0.30
Effort	4	_	_	_	1.00	0.80	0.60	0.40
choice	3	_	_	_	_	0.90	0.70	0.50
	2	_	_	_	_	_	0.80	0.60
	1	_	_	_	_	_	_	0.70

A real world analogue for each of these elements can be found in the take-off example. Each of the employee in charge of a specific activity (e.g., hostess, captain, etc) has to simultaneously decide how much he/she wants to exert him/herself to accomplish the assigned task in a timely manner. Although each employee has to bear his/her own personal costs (i.e., disutility from work), on-time departure depends only on the last activity to be completed; that is, the minimum effort in the group. Moreover, possible effort levels and outcomes are common knowledge among employees (everyone knows that the others know that I know that they know, and so forth). In case reputation of the air carrier and wages are affected by on-time departure (see, Knez and Simester, 2001) and this effect looms larger than subjective cost of effort, everyone has an incentive to do his part of the job. However, if someone think that someone else in the group will delay the

[3] Van Huyck et al. (1991) proposed a similar order-statistic game named 'median game' in which the earnings depends on the individual effort level and the distance between the individual effort level and the median effort in the group. The minimum-effort game is an extension to more than two players and actions of the stag hunt game(Cooper et al., 1992)

take-off he/she should not perform at his best: for instance, if the person responsible for loading luggage think that boarding will be delayed, he/she should not exert the maximum effort to complete his/her task on time. The same tension is present in the minimum-effort game. Notice that if everyone picks the highest possible action ("effort"), then the Pareto-dominant ("efficient") equilibrium in the upper-left corner is selected, while if any of the players chooses the lower possible action, then the Pareto-dominated ("secure") equilibrium in the lower-right corner is selected (see Figure 1.2). Thus, a tension between efficiency and security is present in the game and, as noted by Van Huyck et al. (1990), deductive equilibrium analysis does not help solving the aforementioned tension. Indeed, players may have inconsistent beliefs about others' choices, hence failing to correctly forecast the minimum within the group; in addition, best responding can lead to coordination on Pareto-dominated equilibria. Although the complexity of real-world pooled interdependence cannot be fully capture by this simple game, it can represent an useful tool to clearly define organizational concepts, such as pooled interdependence, that are too often vague and ambiguous[4] so to build a common ground for scholars from different disciplines.

1.3. The minimum-effort game: experimental evidence

Given the link between the minimum-effort game and situations characterized by pooled interdependence, it is interesting to understand if and to what extent this framework can inform organizational scholars. Although a number of selection principles, or equilibrium refinements, such as efficiency, security, risk-dominance (Harsanyi and

[4] For instance, the term coordination and cooperation are often confused in the organizational literature, hence leading to confusion and misunderstanding (for a discussion, see Weber, 2000: 10-14)

Selten, 1988), and quantal response (Kelvey and Palfrey, 1995), were proposed in the literature for the minimum-effort game and other coordination games, it is a priori unclear which principle is used in practice, and understanding how players coordinate is mainly an empirical, and thus experimental, question (Camerer, 2003).

Initial failure to coordinate on the Pareto-dominant equilibrium and subsequent speedy unraveling to inefficient equilibria (i.e., coordination failure) in the minimum-effort game was firstly documented in the lab by Van Huyck et al. (1990);[5] result which —ceteris paribus— has been consistently replicated by a large number of subsequent studies (see, for example, Knez and Camerer, 1994; Berninghaus and Ehrhart, 1998; Cachon and Camerer, 1996; Chaudhuri et al. 2009; Blume and Ortmann, 2007). The downward drift toward the secure strategy was shown to be robust across cultures, the only exception I am aware of is reported by Engelmann and Normann (2007),[6] and gender (Dufwenberg and Gneezy, 2004).

The severe coordination failure and the striking robustness of the results have attracted considerable attention in the experimental literature and have prompted a surge of studies on this game. A number of researchers have focused on the structure of the game and showed that loss avoidance and deviation costs (that is, the difference between individual effort and group minimum) greatly affect the likelihood of successful coordination (for a review of the literature, see Devetag and Ortmann, 2007b; for experimental evidence on the median game and the relationship between risk aversion and coordination, see Devetag and Ortmann, 2007a, and Heinemann et al., 2004, respectively). For instance, improved coordination was obtained by reducing the cost of deviation from

[5] A number of learning models have modeled the behavior commonly observed in this class of games (eg., Selten and Stoecker 1986; Crawford, 1995; Erev and Rapaport, 1998; Camerer, Ho, and Chong, 2002; 2003).
[6] The minimum effort increased as the share of Danish subjects in the group increased.

equilibrium in one-shot games with a small number of subjects (Goeree and Holt, 2005); increasing the number of rounds (Berninghaus and Ehrhart, 1998); and using a finer grid of actions (Van Huyck et al., 2007). In Cachon and Camerer (1996), the opportunity for subjects to choose whether to opt out or pay a fee to enter the game —which implicitly communicates the expectation of a high minimum— has resulted in improved coordination.[7]

Several other ways to enable successful coordination in the minimum-effort game have been implemented in the lab, and “(m)ost of these ways also seem to increase external validity (e.g., various forms of communication or repetition of slightly payoff perturbed games), at least for organizational contexts broadly construed” (Devetag and Ortmann, 2007b: 342). In the present chapter, I focus on minimum-effort experiments that investigated variables of relevance for actual organizations, such as communication, leadership, mergers, financial incentives, and group membership. While the focus is on the minimum-effort game, I also discuss evidence from other types of coordination games but only when relevant to organizations and not in a systematic manner. The review is organized around three main topics —preventing coordination failure, overcoming coordination failure, and transferring precedents across games— and discusses the impact of a host of organizational-relevant dimensions on behavior in the game at hand. Even though a review of the organizational and psychological literature for each of the dimensions considered in the lab is far beyond the purposes of this survey, I try to briefly highlight and discuss some differences between the organizational and the experimental approach to the problem. Unfortunately, I could not offer a proper comparison between the two literature since they have analyzed the problem from very different perspectives that will be discussed later on in the chapter.

[7] Van Huyck et al. (1993) reached similar conclusions for a median-effort game in which the right to play was auctioned.

1.3.1. Preventing coordination failure

In this section I discuss five different mechanisms, namely communication, group growth and mergers, group salience, leadership, and competition, that have been tested in the lab.

Communication: According to standard game theory, costless and non binding communication (i.e., cheap talk or pre-play communication) do not affect the structure of the game and thus behavior. However, mounting experimental evidence has shown that cheap talk can have a dramatic impact on behavior, and, as noted by Farrell and Rabin (1996), cheap talk does not affect *directly* payoffs, but "given people respond to it, talk definitely affects payoffs." (104). Cooper et al. (1992) were among the first to introduce pre-play communication in a repeated one-shot stag hunt game; interestingly, one-way communication was found to be more effective in enhancing coordination as compared to two-way communication. A dramatic improvement in coordination was also achieved in both minimum and median game when subjects had the possibility to send costless pre-play messages with minimal informational content about the action they intended to play (Blume and Ortmann, 2007). Manzini et al. (2009) reported evidence that communication cues, such as smiles, are a reliable means to produce efficient coordination in the minimum-effort game and costly signals are more effective than non-costly signals. Not only pre-play but also post-play communication was studied in lab minimum-effort games; indeed, the results in Dugar (2008) suggest that disapproval ratings may be more effective than approval messages. Credible assignments —that is, non-binding advices coming from a third party— first studied by van Huyck et al. (1992) and Brandts and MacLeaod (1995), were shown to render efficient equilibria more salient in a stag-hunt game (Bangun et al., 2006); that was the case even when the assignment was only 'almost common knowledge'.

Some other studies, however, have shown that the benefits of communication are by no means universal and the introduction of communication costs, even if trivial, can have a detrimental effect. Chaudhuri and Paichayontvijit (2007) showed that credible assignments in a minimum-effort game are not effective in case of randomly-rematched groups unless associated with performance bonuses. Even though considerable evidence has shown the beneficial effect of communication on coordination, there are situations in which communication can be detrimental. Cason et al. (2009) reported evidence on efficiency losses when within-group communication is allowed in a minimum-effort contest. A similar efficiency loss was also observed by Blume et al. (in prep) when a trivial communication cost was introduced.

It is worth to note that the aforementioned experiments implemented extremely simple form of communication with very limited room for misunderstandings and ambiguity. In contrast, organizational scholars, influenced by authors such as Ludwig Wittgenstein and Noam Chomsky, have prompted a new understanding of language, not intended as mere codified information, but as a complex, subjective, and domain-specific entity (e.g., Dougherty, 1992; Boland e Tenkasi, 1995; Carlile, 2002; Becky, 2003). Nevertheless, the aforementioned complexity of communication have been highly neglected in laboratory experiments. A notable exception is the study by Weber and Camerer (2003) in which two groups that have independently developed an idiosyncratic and 'home-made' language are merged together and asked to coordinate. The authors found that idiosyncratic languages hamper coordination in merged groups and subjects tend to systematically underestimate the communication problem.[8]

Group growth and mergers: Several studies have shown that the number of players within a group has a strong and negative impact on the minimum effort; that is, large groups tend to fall prey of coordination failure more frequently than small groups (for a

[8] The same testbed was also used by Feiler and Camerer (forth.) and Rick et al. (2007).

review, see Devetag and Ortmann, 2007b; Engelmann and Normann, 2007). Since most of the real world groups are rather large, at least as compared to the dimension of laboratory groups, the observed regularity poses an important question: how large and successfully coordinated groups can be grown? Thus, a series of studies have investigated in the lab whether efficient coordination obtained in small groups can be sustained when adding entrants or merging groups. Knez and Camerer (1994) found that initial coordination is spoiled when two groups are merged together; information about the other group history produced only limited spillovers and did not result in significantly higher coordination levels. Gurguc (2008) reported evidence that when a merger is the result of a democratic decision —subjects are asked to vote for or against the merger— efficient coordination in the minimum-effort game can be achieved even in large groups. Weber (2006) obtained rather efficient coordination in large groups by gradually adding entrants to an initially small group; however, the growth processes has to be slow enough and new entrants need to be aware of the group's history. It is worth to notice that, when the pattern of growth is decided by an experimental subject acting as a manager, too many entrants are typically added by the manager in the first place (Weber, 2005). Salmon and Weber (2009) noticed that in many real-world situations new entrants are incorporated in an efficient group with a high performance level; for instance, leading organizations hire new employees to expand their business and developed countries need immigrants coming from poorer countries to sustain their growth. The authors report evidence on how different entry rules affect coordination in the minimum-effort game when subjects coming from a low performance group are added to a high performance group: both entry quota and entry exams led to higher levels of coordination as compared to an unrestricted entry rule. The accumulated body of evidence seems to suggest that a more participative and gradual process of integration can mitigate the negative effects of a merger.

A great deal of attention was paid to mergers and acquisitions (M&As, hereafter) by organizational scholars after the wave of unsuccessful M&As that took place in the 1980s. Communication and coordination have been claimed to lay at the core of successful M&As, and several obstacles to the integration process have been highlighted: culture clash, difference in the incentives schemes, difference in the career advancement systems, and psychological issues related to downsizing are just a few of the obstacles reported in the literature (see Larsson and Finkelstein, 1999 for a discussion). The experimental literature has addressed some of this issues[9] but aspects such as downsizing and career concerns in a M&A are still largely unexplored. The threat of downsizing could interact with the democratic and participated processes that have been shown to significantly reduce the negative effects of M&As. That is, would group members vote in favor of a merger even though they are aware of the threats it brings about?

Group salience: Organizations commonly invest money to foster commitment to the firm and to identify employees with their team and the firm itself. The process of identification within a group has a long tradition in social psychology, and, more recently, has also been acknowledge by economists. The power, and the possible drawbacks, of the identification process within a group have been investigate in the minimum-effort game by Chen and Chen (2009) that found that social identity (induced by using a near-minimal and an enhanced near-minimal group paradigm) enables coordination among in-group members, while it has no effect on coordination between members belonging to different groups. Feri et al. (forth.) noticed that, in actual organizations, groups (e.g., units, working teams, or divisions) have to coordinate with other groups; this observation shifts the focus from individual to group decisions in situations of strategic uncertainty. In line with mounting evidence from both psychology and economics that group decisions substantially

[9] For an interesting artefactual experiment on M&As and incentives in a different class of games, see Montmarquette et al. (2004).

differ from individual decisions (see, for instance, Bornstein and Yaniv, 1998; Cooper and Kagel, 2005; Sutter, 2009), the authors reported evidence that groups better coordinate, as compared to individuals, in a series of coordination games including the weak-link game. It is interesting to notice that both group membership and group-based decisions lead to very similar patterns of behavior (for a discussion of the link between the two literatures in a variety of settings, see Sutter, 2009).

Leadership: Charismatic and capable leaders are one of the most valuable organizational resources; large amounts of money are thus invested to hire and train managers and leaders. Weber et al. (2001) showed in a minimum-effort game that the role of leadership is often overrated and leaders are mistakenly blamed for coordination failure while failure is due to games characteristics, namely group size. Randomly selected subjects acting as leaders had to give a brief speech before the beginning of the game and the effectiveness of their leadership style was assessed at the end of the stage game; larger groups, in which mis-coordination was more likely independently from the leader's skills, evaluated leaders as less able as compared to small groups. While Weber et al. (2001) studied misattribution of leadership, Cojuharenco and Gozluklu (2008) investigated leaders' misattribution of responsibilities and found that leaders are by and large unable to correctly attribute the size effect. In a pure coordination game with no tension between efficiency and security, Dana et al. (2007) found that advices from self-interested leaders are far less effective as compared to advices from a neutral party.

The aforementioned experimental studies present an interesting test of organizational and psychological theories of leadership. However, experiments can provide a valid tool to test many other theories developed in the managerial field; for instance, lab experiments could help shed light on whether leadership style is an innate characteristic or can be learned. In addition, the well know managers' bias toward short-run objectives could be studied in the lab (Dana et al. 2007 provide a first interesting step in this direction by introducing self-interested managers). Another highly debated issue that could be undertaken in the lab is to test which mechanisms better help reducing both misattribution of leadership and leaders' misattribution.

Competition: Firms and employees commonly operate, and need to coordinate, in very competitive environments: firms compete among each other for resources and market shares, while employees compete for promotions and bonuses. Competition has long been considered to be an important driver of efficiency and performance. A growing number of scholars investigated the role of competition in the minimum-effort game. Bornstein et al. (2002) observed more efficient coordination by introducing inter-group competition, while Myung (2008) showed that competition for external investments enhances coordination. Kogan et al. (2008) reported evidence that the introduction of an asset market based on the results of a minimum-effort game increases efficiency, but, at the same time, harms coordination. Cason et al. (2009) provided evidence that within-group communication can produce inefficiency in a competitive context. Finally, Fatas et al. (2006) studied within-group, rather than between-group, competition in a modified version of the minimum-effort game: specifically, subjects contributing the minimum effort within the group do not benefit from group profits, unless the group coordinated on any of the equilibria. Within-group competition increased coordination with respect to the baseline condition, although a strong polarization on either the efficient or the secure equilibrium was observed.

While competition at the group level has received quite some attention in lab coordination games, competition among group members is still largely understudied in coordination games. In contrast, organizational literature has always considered competition at the individual level, such as promotions and individual bonuses, as an extremely important element for managing human resources.

1.3.2. Overcoming coordination failure

We have so far discussed a number of mechanisms and institutions aimed at enhancing coordination in the minimum-effort game. A series of recent studies have tackled the problem from a different perspective and have investigated how to overcome, rather than avoid, coordination failure. The aforementioned shift is of primary importance for actual organizations for at least two main reasons. First, organizations are often trapped into inefficient equilibria;[10] as noted by Nanda (1999), organizational inertia is commonly observed because employees and managers are often concerned with regard to skepticism toward change from other employees and managers (for a survey beliefs' misalignment among managers, see Collins, Ross, and Ross, 1989). Second, it might well be the case that what helps avoiding coordination failure is not necessarily well suited for overcoming coordination failure, and vice versa. A systematic study of different devices aimed to pull a group out of coordination failure has been carried out by Brandts, Cooper, and Fatas by way of the so called 'corporate turnaround game'. The typical experiment was divided in three parts and subjects played in fixed groups of four; in the first part of the game, the parameters were set in such a way that the Pareto-dominant equilibrium was extremely risky so to induce, almost invariantly, coordination failure. Diverse variables were then manipulated in the two subsequent parts. In the following, I report evidence on how and to what extent monetary incentives, observability and leadership, and communication may help escaping the coordination trap.

Monetary incentives: Brands and Cooper (2006b) implemented an increase in the bonus after having trapped subjects into the Pareto-dominated equilibrium; the bonus increase was found to trigger the process of change, and, interestingly enough, the magnitude of the bonus had no, if not negative, effect. The bonus rise is thus important for

[10] See Knez and Simester (2001) for an interesting case study on Continental Airlines.

the signal it conveys, rather than for the reduction of the riskiness of the efficient strategy; indeed, reducing the bonus to the initial level had little effect on coordination. As suggested by the dynamic analysis of the data, the presence of both strong leaders (i.e., subjects that significantly increase their effort level as a response to the bonus rise) and responsive followers are necessary ingredients for escaping the coordination trap. Hamman et al (2007), building on Brands and Cooper (2006b), have varied the magnitude, the valence, and the type of the bonus. Specifically, both untargeted and targeted bonuses were used; while the first type of bonus was exactly the same as the one used in Brandts and Cooper and applied to all equilibria, the targeted bonus only applied to the Pareto-dominant equilibrium. Targeted incentives had a stronger effect on effort immediately after their introduction and were, overall, more effective as compared to untargeted incentives. However, targeted incentives only had a temporary effect and did not sustain coordination when not in place.

Observability and leadership: Brandts and Cooper (2006a) compared monetary and non-monetary incentives, namely bonus and feedback, in the corporate turnaround game. The role of leaders, which can emerge only in case information about the full distribution of effort is give to subjects, was proven crucial in overcoming coordination failure and more effective than the magnitude of the monetary bonus. Notwithstanding, observability had little effect in preventing groups to slipper into coordination failure when the high bonus was removed. Given the importance of leaders in pulling a group out of a dominated equilibrium, Brandts et al. (2007) explored the relationship between cost heterogeneity and emergence of leaders; the authors found that subjects with the most common, rather than the smallest, cost type are more likely to become leaders and systematically overshoot the minimum. Brandts et al. (2009) considered the emergence of leadership within a group with heterogeneous members. Specifically, more able subjects (i.e., subjects with a lower effort cost) can help out disadvantaged subjects within the group (i.e., subjects with a higher effort cost). Commitment to helping others had a strong effect in overcoming coordination failure, while help offered on a temporary base (that is, the decision was taken at the beginning of each period) was only weakly effective.

Communication: the role of non-binding messages, open-ended text messages in particular, was studied by Brandts and Cooper (2007) and communication was found to be more effective than financial incentives in overcoming coordination failure. In an artefactual experiment, Cooper (2006) had both students and experienced managers acting as managers in the turnaround game and found that the latter are able to lead the group to coordinate on higher effort levels significantly faster than the former. Chaudhuri et al. (2009) used inter-generational advices to overcome an history of initial coordination failure; the effectiveness of the advices depended on both the content itself and the way the message was conveyed to subjects.

To sum up, overcoming an history of coordination failure is more about sending a credible signal to all subjects rather than reducing the strategic uncertainty through a change in the riskiness of the efficient strategy. Both leadership and communication have indeed proved more effective than financial incentives in enhancing coordination. These results seem to suggest that coordination can be accomplished even without a vast expense of money if organizations can send a clear signal on how and when to change behavior. However, not all mechanisms have a long term effect on coordination; for instance, targeted incentives have a transient effect which vanish when they are removed.

1.3.3. Precedent transfer and learning across games

In previous experiments, the focus was on a single game played repeatedly and learning dynamics within the same game were typically considered. In real life and actual organizations, however, it seems very unlikely that the exact same game is repeated over and over; indeed, subjects are often confronted with different games, and lessons learned in one domain can generalize to other domains. For instance, outdoor experiential training is often used to build trust among team workers in the hope that what has been learned in the outdoor camp will transfer to the workplace (e.g., Knez and Camerer, 2000: 196). Even though there is considerable psychological evidence on cross-games transfer and related difficulties, economists have neglected the problem for a long time. Knez and Camerer (2000) experimentally studied whether a precedent of successful coordination in the minimum-effort game produces efficient cooperation in the Prisoner's Dilemma. The authors found a strong and positive effect of precedents transfer; however, the positive effect largely depended upon descriptive similarity rather than trust formation or shared experience of playing the efficient strategy. Rankin et al. (2000) used similar but not identical strategic situations to prevent subjects to base their decisions on descriptive characteristics of the game and found that the efficiency principle predicts choices in a perturbed stag hunt game played repeatedly. Unlike Knez and Camerer (2000), Devetag (2005) tested how efficiency in the weak-link game can be enhanced by means of a precedent transfer; the author showed that efficient coordination in the critical mass game (Devetag, 2003) creates positive spillovers for coordination in the minimum-effort game.

Cross-game learning to play strategically was observed in a related class of coordination problems called signaling games (e.g., Cooper and Kagel, 2008). Rick and Weber (forth.) found that "meaningful" learning generalizable across diverse domains can be produced in repeated games when no feedback are provided. The absence of feedback, in fact, led subjects to think more deeply about the structure of the game and prevented them from dwelling upon descriptive elements only.

The paucity of evidence on precedent transfer and cross-games learning contrast starkly with the dynamic environment faced by organizations and the need for transfer of knowledge across domains. Thus, how to derive meaningful lessons from previous experience is of great importance for practical purposes. In addition, a better understanding of far too long unexplored issues, such as the perceived similarity between different situations and the perceived structure of a game (for a notable exception, see Devetag and Warglien, 2008), would prove extremely useful for organizations broadly construed as well as for academic purposes.

1.4. Organizational vs. experimental literature

The success of the advocated interaction between organizational and experimental approaches to coordination problems critically relies on at least two factors: first, possibility to generalize results obtained in the lab to organizational-relevant situations (see the Introduction and Chapter for a discussion of the external validity issue); second, possibility to establish shared definitions, or at least possibility to link the different formalizations used in the two disciplines (see §1.2).

While augmented realism have been introduced in several types of experiments in an attempt to address concerns about external validity of lab results, still little has been done in coordination game. Among the few exceptions, the framed experiments by Brandts and Cooper (2006a; 2006b) and by Brandts et al. (2007) and the artefactual experiment by Cooper (2006) on the minimum-effort game deserve to be mentioned. Quite surprisingly, no field or real-effort experiments[11] have been carried out for the minimum-effort game, or other coordination problems. New technologies, such as Mechanical Turk and virtual worlds, can also constitute a new testbed to study coordination in a more realistic context.

In the present chapter, I have discussed how pooled interdependence and the minimum-effort game can be linked. Even though many similarities between the two definitions are present, some discrepancies ought to be mentioned. In the minimum-effort game the action space is given and the strategy space is common knowledge; hence, efficient coordination reduces to the question: *which mechanisms is more efficient to induce all subjects to exert the maximum effort?* The complexity of coordination problems in actual organizations has led organizational scholars to take a slightly different perspective on the problem at hand: *which rules and structures are better suited for reaching efficient coordination, provided that employees will exert themselves at their best?* Let consider, once more, the take-off example. Experimentalists commonly assume that efficient rules are in place and efficient coordination is feasible; according to this perspective, if the plane does not take-off on time it is because efficient coordination is risky and employees in charge of the various activities (e.g., boarding passengers, security check, fueling, etc.) believe that some of the employees will decide to shirk. Hence, the focus is not on the (theoretical) feasibility of efficient coordination, but on individual attitudes toward risk, beliefs formation, and group trust. In contrast, organizational literature on coordination mainly focuses on which rules are needed to make sure that the plan can take off on-time (Okhuysen and Bechky, 2009). Thus, late departure can be caused by a bad activity scheduling, a problem in the definition of the procedures, or the lack of communication between different employees. Although very different, the two perspectives can complement each other hence laying out an interesting research agenda.

[11] All studies presented in this chapter implemented a choice of costs on a convex or linear function that is intended to be a proxy for real effort. For a review of real-effort experiments, see Chapter 2.

1.5. Discussion and conclusions

In this chapter, I have reviewed the experimental evidence on the minimum-effort; a steady growth in the number of experiments that studied how organizational elements affect coordination in the lab was observed. The effect of a host of factors, such as communication, leadership, competition, M&As, financial incentives, group's growth, on both preventing and overcoming coordination failure have been presented; in addition, the role of precedent transfer has been discussed.

Even though several variables of relevance for managing actual organizations have been studied in the minimum-effort game, there are still several issues that have been largely understudied in the experimental literature on coordination. Some topics suitable for experimentations are reported in the following:

Abilities and learning: ability heterogeneity and learning are common in the workplace, but quite difficult to capture in the lab; a nice exception is represented by Brandts et al. (2007; 2009) that introduced heterogeneous effort cost in the minimum-effort game. I believe that a better understanding of how unexperienced newcomers with low ability and learning affect coordination would prove useful.

Self-selection: a related issue is self-selection; indeed, employees self-selected into a job and are sorted by the employer. A few experiments addressed the self-selection issue in other games (e.g., Lazer et al., 2006; Dohmen and Falk, 2006), but to our knowledge nobody studied the relationship between self-selection and coordination (the most closely related paper is probably Cachon and Camerer, 1996, but the purpose of the paper was different).

Multi-tasking: Camerer and Weber (2007) noticed that, quite surprisingly, all organizational experiment in economics focus on a single activity, while multi-tasking is widespread in actual organizations. This is an extremely important area for research, also because it is very difficult, if not impossible, to write first best contracts for this type of situations.

Social and hierarchical structure: almost every organization relay on a formal structure and a hierarchy, this dimension, however, has been mainly overlooked in the experimental literature on coordination (for interesting exceptions, see Rick et al., 2007; Cassar, 2007).

Communication flows and conduits: organizational scholars have devoted considerable attention to the structure of the communication flows and the best means of communication. For instance, different types of communication (e.g., written vs. face-to-face) could be studied in coordination games.

Personality traits and cognition: personality characteristics, such as the so called Big Five, have since long been considered of primary importance to organizational performance and dynamics, but little if no experimental evidence considered this aspect. Moreover, as pointed put by Devetag and Ortmann (2007b), little is known about the underpinning cognitive processes leading to successful, or unsuccessful, coordination.

Chapter 2

Real Effort in Field and Lab Experiments

2.1. Introduction

Experiments have gained a prominent role in testing economic theories and the underlying assumptions: laboratory experiments, in particular, are said to offer a controlled context in which tough empirical tests can be conducted. Control, randomization, and replicability are obvious advantages relative to the noise naturally occurring data tend to be afflicted with. However, scholars have criticized laboratory experiments because internal validity is maximized at the cost of external validity: oversimplistic settings, lack of context, presence of small stakes, and the use of demographically unrepresentative subject pools (students) have been among the main reservations as regards the experimental methodology (e.g., Loewenstein, 1999; Ariely and Norton, 2007; for an early discussion in the organizational literature, see, for instance, Gordon et al., 1986; see also the Introduction of the present work for a discussion).

Concerns about external validity and generalizability of lab results have gained new momentum among experimentalists (e.g., Schram, 2005; Hogarth, 2005) and probably have contributed to the increased use of field experiments. Harrison and List (2004: 1012) have defined six characteristics identifying the various types of field experiments: nature of the subject pool, nature of information, nature of commodity, nature of task or trading rules, nature of stakes, and nature of environment. In the present work, I consider one specific aspect of the nature of the task not taken into account in previous reviews: real, rather than chosen, effort. One of the critical maintained assumptions in lab experiments is that a (almost effortless) choice of costs on a convex or linear function is a reliable measure for real effort. Notwithstanding, as pointed out by van Dijk et al. (2001), "[work] involves effort, fatigue, boredom, excitement and other affections not present in the abstract experiments" (189).

Although real and chosen effort have been implicitly assumed to lead to similar results, several factors might be responsible for differences between the two methods. The introduction of a real-effort task can trigger heuristics developed in everyday experience which might not be present when context is stripped away. In addition, intrinsic utility might be derived from work; there is some evidence that people are more willing to provide voluntary work, than contributing money, for a given purpose (see, van Dijk et al., 2001). Moreover, real clerical tasks allow us to observe how task learning, individual skills, and motivation interact with each other.

Despite the relevance of the issue, relatively few experimenters have operationalized real-effort tasks. Thus, important issues, such as how real effort can be measured and implemented and to what extent its results might differ from chosen effort, remain unsettled. In this chapter, I critically review both lab and field experiments using a real-effort task.

The chapter is organized around three main questions: first, *how can real effort be operationalized and measured in a controlled experiment*?; second, *can heterogeneity in* *innate ability be effectively controlled*?; third, *do real effort and chosen effort lead to similar results*?. In Section 2, I present a variety of clerical tasks which have been used in controlled experiments and discuss proxies used to measure effort and mechanisms used to control for heterogeneity in the distribution of abilities. In Section 3, I compare results obtained by using real-effort and chosen-effort tasks, while in Section 4, I discuss pros and cons of implementing a clerical task and consider classes of experimental games in which real effort has not yet been used.

2.2. Real-effort tasks and subjective abilities

Before presenting the clerical tasks used in the literature, some methodological caveats are in order. I extensively searched the literature on both lab and field experiments in an attempt to provide an extensive survey of the existing evidence on real effort tasks. To include as many studies as possible, I used EconLit (key words used during the search: real effort and experiment/ experimental/laboratory, real task and experiment/experimental/laboratory) and IDEAS (key words used for the search: real effort, real task);[1] I also took into account all publications and working papers not contained in the two aforementioned search engines I was aware of. In the present work only economics experiments[2] that used monetary incentives and contingent payments were considered. For expositional purposes, I review studies which used a real-effort task in the core part of the experiment and do not consider studies that implemented a real-effort task to legitimize subjects' initial entitlement (e.g., Cherry et al., 2002) or to test overconfidence (e.g., Cesarini et al., 2006). For instance, overconfidence experiments commonly use

[1] EconLit and IDEAS were last accessed on September, 2009.

[2] A review of real effort in the psychological literature is out of the purpose of this work; we only refer to a few psychological experiments that posed the bases for real-effort economics experiments hereto discussed.

general knowledge questionnaires, hence requiring subjects to exert some degrees of effort; however, in this class of games the main focus is on subjects' beliefs about their performance and not on the performance itself.

In standard laboratory experiments, since no real tasks or leisure are implemented, disutility is commonly associated with a monetary cost as well. For instance, disutility from work in a gift-exchange game is associated with a monetary cost, and the disutility from producing a good in an induced value market is a monetary cost. However, this is only one possible translation of disutility, commonly used in the lab for its simplicity and convenience. In economic theory, time spent working on the assigned task is typically associated with disutility of foregone leisure or with the opportunity cost of not working on a different task, while psychologists consider effort, and the associated disutility, as an unpleasant sensation produced when repeatedly performing an activity. In real-effort experiments, disutility cannot be modeled as a monetary cost, and other measures more in line with the one proposed by economic and psychological theories have to be used.

Thus, when moving from a chosen-effort to a real-effort task, how to measure disutility from work (or effort) is a crucial issue. Moreover, qualifications to the *non-satiability* concept proposed by Smith (1979) are of great relevance. Specifically, two qualifications deserve particular attention in real-effort experiments: (a) subjective costs associated with the decision at hand; (b) subjective game value attached to experimental outcomes. The problem of measuring disutility from effort and controlling for qualification (a) and (b) has been addressed in the experimental literature by using different techniques. In the following, I consider three different aspects of the problem: first, choice of the clerical task; second, measures of effort; third, task-relevant abilities and attitudes toward the task.

2.2.1. Real-effort tasks

The choice of the clerical task to be used in the experiment is a critical decision, and real-effort tasks implemented in the lab or in the field usually share the following characteristics: first, the task is easy enough to be performed by virtually everyone with no need for an extensive training; second, the task is uninteresting and quite boring so to minimize enjoyment and intrinsic motivation. These specific realizations of task dimensions have been chosen to minimize confounds potentially arising from learning effects and utility resulting from enjoyment and intrinsic motivation (e.g., Smith 1976; 1982). In the following, a list of the various tasks used in the literature is provided; for expositional purposes the tasks are categorized as cognitive and physical.[3]

Cognitive tasks: Tasks involving mathematical skills have often being used: Sutter et al. (2003), Dohmen and Falk (2006), Brüggen and Strobel (2007), Tomer, et al. (2007), and Kuhnen and Tymula (2009) all asked subjects to solve mathematical equations and multiplications; Heyman and Ariely (2004) and Niederle and Vesterlund (2007) used additions of two-digits numbers; and LevyGarboua et al. (2006) asked to decode a number from a grid of letters. Linguistic skills were needed for the task operationalized by Charness and Villeval (2009): in this experiment, subjects had to solve anagrams; while memory and logic were required for the following tasks: mazes (Gneezy et al., 2002; Gneezy et al., 2003), memory games (Ivanova-Stenzel and Kübler, 2005), sudoku (Calsamiglia et al., 2009), and the tower-of-Hanoi puzzle (McDaniel and Rutstrom, 2002). A two-variable optimization task was introduced by van Dijk et al. (2001): subjects had to search by trial and error the

3 The two categories have not to be intended as mutually exclusive; indeed, some tasks may lay in between of the two categories and share some elements of both categories. When performance does mainly depend on cognitive abilities (e.g., math or linguistics) or general knowledge, the task is classified as 'cognitive'. The category of physical tasks includes a more broader range of activities: some of them require physical strength to be adequately performed (e.g., picking oranges), while other are more mechanical and require more concentration than physical strength (e.g., entering data).

maximum of a function (the same task was used by Bosman and van Winden, 2005; Bosman and van Winden, forth.; Bosman, Sutter, and Winden, 2002; while a modified version of the game was operationalized by Dickinson and Villeval, 2004; Montmarquette et al., 2004; Sloof and van Praag, 2005).

Physical tasks: Real effort has quite a long tradition in psychology, Deci (1971) was one of the first in introducing mind-numbing tasks in a controlled experimental setting: specifically, subjects had to reproduce a series of figures with a puzzle called 'Soma'. In a field experiment, Deci (1971) asked subjects to write headlines for a college newspaper. Secretary tasks in general, and data entry in particular, have often been used in the lab: Hennig-Schmidt et al. (2005) let subjects type abstracts, Gneezy and List (2006) asked subjects to enter information about library books into a database (the same task was also used by Kube, et al., 2007; 2008), and Ottone and Ponzano (2008) asked participants to copy information about fictitious students in a database. Stuffing letters into envelopes was first operationalized by Konow (2000) and subsequently used by Falk and Ichino (2006) and Carpenter et al. (forth.). Azar (2009) asked subjects to find a letter on a specific page, line, and position, while Heyman and Ariely (2004) designed an experiment in which a computerized ball had to be dragged from one place to another. Manual tasks, such as picking oranges (Erev et al., 1993), beading little plastic pearls on a string (Eberlein and Przemeck, 2006), and crack walnuts (Fahr and Irlenbush, 2000), were also used. Finally, Gneezy and List (2004) used a door-to-door fundraising task. In Pokorny (2008), Abeler, et al. (forth.), and Falk et al. (2008), subjects had to count the number of zeros contained in a series of tables.

2.2.2. Measures of effort: quantity, quality, difficulty, and time

A major shortcoming of real-effort, relative to chosen-effort, tasks is that disutility that is associated with the task cannot be directly observed —let alone measured— by the experimenter; thus, proxies for the extent to which subjects exerted themselves (and experienced disutility) in a clerical task need to be identified. Specifically, what we usually observe, namely performance, is function of both effort and ability. Four proxies for effort, or some combination of them, have been used in the literature: quantity of the output, quality of the output, level of difficulty of the task, and time spent completing the task.

Output quantity is probably the most obvious measure of effort that can be used (e.g., number of stuffed letters, number of solved mazes, number of entered records, and number of solved equations). However, more quantity does not necessarily imply more effort if output can be of different qualities. While some studies did not allow for any variance in quality —that is, subjects could not move further in the experiment if they did not revise the wrong answer (e.g., Ottone and Ponzano, 2008) or the output was not counted in case of mistake (e.g., solving additions or counting zeros in a table) — other scholars considered both quantity and quality in their analysis. For instance, Henning-Schmidt et al. (2005) and Kube et al. (2007; 2008) considered the number of both entries and mistakes in their library tasks, contrary to Gneezy and List (2006) that used the number of entries only. Unlike Konow (2000), Hemming-Schmidt et al. (2005), and Falk and Ichino (2006) that used the number of stuffed envelopes to determine effort, Carpenter et al. (forth.) adjusted the output for quality.[4] It should be noticed that mistakes do not always reflect a lack of effort; indeed, there might be situations in which subjects that are very

[4] A letter carrier from the US Postal Service was asked to assess the deliverability of all the envelopes on a scale from 0 to 1. Envelopes' quality was also evaluated by other participants in the tournament to test how sabotage, in particular, and work politics, in general, interacts with monetary incentives.

close to their limit performance, are induced to make more mistakes than subjects with superior skills, no matter how much they exert themselves.

Another element that should be taken into account is the level of difficulty of a task. Some authors controlled for difficulty ex-ante by providing a sequence of tasks of the same difficulty over time; for instance, Kuhnen and Tymula (2009) made sure that the multiplications subjects were faced with were all of the same difficulty. In contrast, Azar (2009) preferred to increase the difficulty of task over time. In Gneezy et al. (2003) subjects could choose the desired mazes' difficulty level themselves.

Time can also be used as a proxy for effort, the assumption being that the longer time was spent on a specific task, the higher was the effort exerted (and hence the disutility experienced). However, note that time is also likely to convey information about task-relevant abilities. For instance, Falk et al. (2008) used time spent counting numbers in a series of tables as the main outcome variable of interest. A different approach was used by van Dijk et al. (2001) where effort was proxied by the time allotted to each of the two productive activities subjects were asked to do.[5] Azar (2009) took into account time spent working on the task in addition to quantity of the output and difficulty of the task. These examples illustrate how a commonly agreed upon proxy for real effort (and experienced disutility) has not yet been found and the most well suited variable to be taken into account often depends on the task at hand and the structure of the game. Moreover, most of the proxies I have discussed are more well suited to measure performance, rather than effort and the associated disutility. Even though taking into account different measures, such as

[5] Two tasks were used to account for the time actually spent working for the employer and time spent on other activities (productive for the worker only). A two-variable optimization task has been used for both tasks, but task A produced benefits for both the employer and the workers, while task B was rewarded individually on a piece-based rate and did not provide any utility for the employer. The authors decide not to operationalize different tasks because the benefit of 'doing nothing', surfing the net or enjoying some other activity cannot be controlled.

quantity, quality, difficulty, and time, helps in better understanding to what extent subjects exerted themselves, it is difficult to combine different measures and develop a compound index of effort.

2.2.3. Subjective abilities

As indicated in the preceding paragraphs, factors complicating the analysis of outcomes in real-effort experiments are innate abilities and previous experiences in the task at hand; this problem is related to subjective costs of decision discussed by Smith (1979), but is sharpened by the presence of heterogeneity of task-relevant ability. In particular, when skills are heterogeneously distributed in the sample, disentangling actual effort from abilities is not simple and invite systematic problems with the analysis of the data. In addition, virtually none of the variables discussed in the previous section can account for individual differences in innate abilities and skills in a satisfactory manner; thus, the problem has often been tackled by developing ingenious experimental designs. Before discussing these methods, two caveats are in order: first, in within-subjects designs it is possible to determine the net effect of the desired treatments on subjects' performance, but order effects might arise and demand effects may be more salient; second, in between-subjects designs we might expect subjective abilities to be randomly distributed across treatments, thus data at the aggregate, but not at the individual, level should not be affected by the heterogeneity of skills. In the following, I briefly present what, in my view, are the most interesting techniques —which usually combine the advantages of both within- and between-subjects design— that have been used to address the issue of the heterogeneity of skills.

Brüggen and Strobel (2007) developed an ingenious method to control for subjective abilities: in the first part of the experiment, subjects were asked to solve as many two-digit multiplications as they could within 5 minutes time and were paid on a piece-rate base. In the second part, a gift-exchange was played with the principal choosing a wage level for two agents, one of them working on a chosen-effort task and the other working on a real-effort task. Heterogeneity in individual ability was eliminated by translating real effort into the ratio of solved multiplications in the two parts of the experiment.[6] In the two-variable optimization task a within-subjects design was used by Montmarquette et al. (2004) and Sloof and van Praag (2007). Similarly, Henning-Schmidt et al. (2005) divided their experiment in two parts and varied the wage level in the second part, so as to compute the net effect of their manipulation. One possible shortcoming of this method is that, if subjects worked at their best in the first part of the experiment, an unexpected wage increase cannot indeed translate into a corresponding increase in performance since subjects cannot increase their effort.

Tomer et al. (2007) used the first part of their experiment to make sure that abilities were homogeneously distributed across treatments; subjects were in fact ranked according to their initial performance and evenly assigned to different experimental conditions. This method allows, in principle, to control for the distribution of abilities across treatments and, at the same time, for the net effect of the treatment by accounting for initial abilities.[7]

In economic theory, time spent working on the assigned task is typically associated with disutility of foregone leisure or with the opportunity cost of not working on a different

[6] The authors also controlled for learning that might have intervened between the first and the second part of the experiment (the two parts were not run during the same day to minimize carryover effects and hedging). However, it should be noticed that the number of solved equations in part 1 does not necessarily constitute the upper bound for performance, since the piece-rate compensation might not have been high enough to induce subjects to perform at their best.

[7] This approach, however, does not guarantee that the initial elicitation of ability is not itself noisy.

task. Along this line, Abeler et al. (forth.) considered the time spent on the relevant task, rather than measuring the extent to which subjects exerted themselves: indeed, participants could leave the lab whenever they wanted, that is, they had to decide how to allocate their time between labor and leisure. Van Dijk et al. (2001) proposed a different method in which the time spent on unproductive, rather than productive, activities was considered. Individual effort for the productive task was defined as the ratio between performance in the productive task and overall performance.

In Ottone and Ponzano (2008) and García-Gallego et al. (2008) subjects were given opportunity to familiarize themselves with the real task at the beginning of the experiment and then had to decide how many records to enter or envelopes to fill, respectively, under a specific payment scheme. This technique is quite close to the tradition of chosen-effort tasks, since subjects are required to decide in advance the quantity of output they are willing to produce during the experiment; notwithstanding, the choice is most probably affected by skills and subjects might have difficulties in correctly anticipate how much effort is needed to meet the chosen performance level.[8]

2.3. Real vs. chosen effort

Testing whether chosen effort is a reliable proxy for real effort is central to better understanding to what extent results obtained in a controlled lab setting can be generalized to the real world. Notwithstanding, none but two of the studies reviewed here carried out a direct comparison between real and chosen effort by implementing both conditions. Brüggen and Strobel (2007) directly juxtaposed real and chosen effort in a gift-exchange

[8] For instance, subjects could underestimate the economies of learning, hence leading subjects to chose a performance level lower than the one they would have chosen if they had foreseen that cost of effort was decreasing over time.

game, and their main findings can be summarized as follows: first, overall effort was higher in the real, as compared to the chosen, condition, hence possibly supporting previous evidence that individuals derive some form of utility from work or, possibly, that when real tasks are involved, an interiorized social norm of "working vs. shirking" takes place; second, reciprocity showed very similar patterns across the two conditions (i.e., real vs. chosen). Thus, the authors claimed that "chosen effort is an appropriate way of operationalizing effort in experiments" (233). García-Gallego et al. (2008) implemented an ultimatum game: the proposer had to make an offer on how to divide 10 Euros between him/herself and the responder. While in the chosen-effort treatment, in case of acceptance, the proposed division was automatically implemented, in the real-effort treatment acceptance of the offer required the responder to perform a task (filling envelopes) in order to implement the proposed split. The authors found both offers and rejection rates to be higher in the real-effort treatment as compared to the chosen-effort treatment. However, the observed results should not be surprising if we take into account the fact that in the real-effort treatment subjects incur a cost for performing the task, while virtually no cost, either explicit or implicit, is associated with the corresponding chosen-effort treatment: higher offers can thus be explained as a compensation for the effort to be exerted.

Even though very few experiments directly compared real and chosen effort, it is possible to gain further insights on the robustness of chosen effort by comparing results in real-effort experiments with patterns of behavior commonly observed in chosen-effort experiments for similar games. Three main classes of experiments are discussed in the following: gift-exchange games, intrinsic motivation, and tournaments.

Gift-exchange games: In line with previous evidence, Gneezy (2002) found positive reciprocity by using mazes instead of chosen effort in the lab. In addition, workers' effort was negatively correlated with employee' returns. In contrast, other studies on the gift-exchange game did not find positive reciprocity when real effort was operationalized. Gneezy and List (2006) found no reciprocal behavior in their library and fundraising tasks: indeed, the effect of a wage raise was only transient and lasted for a short time period. Even more surprisingly, for a subsample of the subjects that participated for two consecutive days in the fundraising task, the amount of money raised in the gift condition in the second day was lower as compared to the money raised in the no-gift condition during the second day. Similarly, Kube et al. (2007) did not find a wage rise to increase the number of logged books, while a wage cut was found to have a strong and large detrimental effect on performance. Kube et al. (2008) found little if no reciprocity for monetary gifts, while a gift-in-kind was found to produce results similar to the results commonly obtained in chosen-effort lab experiments. Henning-Schmidt et al. (2005) observed, in a field study, that work effort is by and large insensitive to an increase in the wage level and to positive and negative peer comparison. It should be noticed that all the experiments, presented so far, that did not find evidence of reciprocal behavior were conducted in the field and subjects were not aware of being part of an experiment.[9] Thus, the observed differences might be due to either real effort or to the more natural setting present in a natural experiment or both. To test this conjecture, Henning-Schmidt et al. (2005) conducted a lab study closely resembling their field experiment, and information about employer's surplus was varied. The authors reported evidence that a positive wage-effort relationship is observed only when complete information about employer's costs and surplus are common knowledge: a condition commonly used in lab experiments.

[9] Notwithstanding, not all field experiments did not observe positive reciprocity; for instance, Falk (2007) found support for the gift-exchange hypothesis in a natural experiment in which no real effort was involved.

Intrinsic motivation: We have so far taken for granted that repeatedly performing a task generates disutility; however, subjects can derive some enjoyment from performing the task at hand —this problem was also discussed by Smith (1979: 277)— and can be intrinsically motivated to perform well. The detrimental effect of positive reward (i.e., motivation crowd-out) has become very popular in the psychological literature and poses important questions about the use of financial incentives. This is especially true for real-effort experiments since intrinsic motivation is more likely to be present in real rather than chosen tasks. Indeed, a myriad of psychological experiments on motivation crowding theory used real-effort tasks, since it is extremely difficult to reproduce intrinsic motivation in an hypothetical task. Deci (1971) was one of the first scholars to test,, in a controlled setting, how monetary rewards crowd-out intrinsic motivation; in lab and field real-effort experiments, the author found that monetary incentives tend to decrease intrinsic motivation to perform the task at hand, while the converse was true for the relationship between verbal reinforcement and intrinsic motivation. More recently, Heyman and Ariely (2004) showed that monetary, but not social, markets are highly sensitive to the magnitude of compensation. The crowding-out effect, however, is not as robust as commonly thought; indeed, Eisenberger and Cameron (1996) showed in a meta-analysis that the effect is only marginal and occurs only under very restrictive circumstances. Contrary to Eisenberger and Cameron (1996), Deci et al. (1999) found that the detrimental effect of reward is significant if only “interesting” tasks are considered. Thus, the choice of the task is an important one and can interfere in non-obvious ways with monetary incentives. Also economists have recently tested in the lab the intrinsic motivation crowding-out theory with mixed results. Gneezy and Rustichini (2000) found that the absence of payment leads to better performance on an IQ test as compared to a small piece-rate payment. Unlike previous experiment, but in line with some other results in psychology (e.g., Eisenberger and Cameron 1996), Pokorny (2008) provided evidence that very small incentives do not have a detrimental effect on performance on both cognitive and physical tasks; in fact, an inverse U-shaped relationship between incentives and performance was observed. Even though comparison between real and chosen effort for this class of experiments is almost impossible, the cumulated evidence on intrinsic motivation suggests one ought to be careful when generalizing chosen-effort results to real world situations and especially when choosing the real-effort task to implement in the experiment.

Tournaments: Schotter, Weigelt, and co-authors have provided a series of early experimental tests of (chosen) effort provision under both symmetric and asymmetric tournaments.[10] While equilibrium predictions have been generally confirmed in symmetric tournaments and tournaments have been shown to induce higher effort levels than piece-rate payments (Bull et al., 1987; Weigelt et al., 1989; Nalbantian and Schotter (1997), oversupply has been commonly observed, especially for disadvantaged subjects, in case of asymmetric costs, and affirmative actions have been proven to reduce drop-outs (Weigelt, et al., 1989; Schotter and Weigelt, 1992). Erev et al. (1993) tested intergroup competition in a picking-orange task; similarly to previous lab studies, competition was found to mitigate production losses due to free-riding. Real effort under different incentive systems was also studied by van Dijk et al. (2001); data showed that individual (piece-rate) and team payments resulted in similar performance level: free-riding was indeed compensated by a large number of subjects working harder in the team as compared to the piece-rate compensation. In line with chosen-effort experiments by Nalbantian and Schotter (1997),

[10] In the typical experimental setup, subjects are matched in groups of two and choose a number — the cost associated with the decision are increasing in the number chosen— to which a randomly drawn number is added; the subject with the higher total number win the competition and get the highest prize. In symmetric tournaments, the cost of the chosen number is the same for all members of the group, while in the asymmetric tournament different cost functions are assigned to different subject, that difference being common knowledge.

competition in a real-effort setting induced higher levels of effort, but also more variance; interestingly enough, performance in tournaments was mostly raised by low ability workers (as classified according to behavior under piece-rate incentives). Carpenter et al. (forth.) found that effort is higher in tournament than in individual payment condition. However, real effort in tournament falls when competitors are able to sabotage each other (i.e., quality is evaluated by competitors and not by an external supervisor), thus suggesting that real effort is highly sensitive to ambiguity in evaluation. Asymmetric tournaments were studied in the field by Calsamiglia et al. (2009); the authors systematically varied the training received by subjects in the relevant task and showed that affirmative actions only marginally reduce performance in advantaged groups, while significantly increase the chances of disadvantaged groups to win the tournament. In Gneezy et al. (2003) women performed worse than men in a competitive environment, although showed similar skills in a non-competitive setting. Overall, tournaments, as compared to group payment, have a beneficial effect on performance for both real and chosen tasks in the lab and in the field; even though, little can be said about the magnitude of the effect, since precise point predictions cannot be derived for real effort experiments.

2.4. Discussion and conclusions

In the present chapter, I have tried to assess the causes and nature of the recent surge in real-effort experiments and discussed the new opportunities and problems the introduction of real effort brings about. While real-effort tasks have the clear advantage of increasing external validity of experimental results, this advantage does require to control for new and tricky variables, such as innate ability in solving the task at hand. I have discussed three main questions raised by the introduction of real effort: first, *how can real effort be operationalized and measured in a controlled experiment*?; second, *can* *heterogeneity in innate ability be effectively controlled*?; third, *do real and chosen effort lead to similar results*?

How can real effort be operationalized and measured in a controlled experiment? Quite a number of tasks, requiring either cognitive or physical skills, have been implemented, and several attempts to accurately measure effort have been made; notwithstanding, several issues are still unsettled. First, a single task has typically been implemented in each study and little can be said about the robustness of results across tasks requiring different abilities and skills (for a remarkable exception, see Pokorny, 2008). Second, it is not clear which proxy can be used to reliably measure effort —the subjective cost of effort is clearly unknown to the experimenter; despite several proxies, such as quantity and quality of the outcome, time allotted to the task, and difficulty of the task, have been used, it is still unclear how diverse measures can be integrated together. Finally, and possibly more importantly, it is a priori difficult to establish whether subjects will derive any utility from performing the task: that is, the fact that a task is judged as utterly boring by the experimenter, does not necessarily exclude that some of the participants may enjoy performing it, hence deriving positive utility. While some scholars used post-experimental questionnaires to understand whether subjects enjoyed performing the task, no incentivized and rigorous method to control for subjects' utility for different activities has been developed so far.

Can heterogeneity in innate ability be effectively controlled? I have provided evidence that it is possible, at least to some extent, to control for subjective abilities. A combination of a between- and within-subjects design has often been used to cancel out heterogeneity in the distribution of abilities, while other authors have evenly distributed abilities across treatments. An unsolved problem common to all the reviewed experiments is that performance is bounded from above, and, unlike in chosen-effort tasks, the experimenter does not, and possibly cannot, know the best performance a subject can achieve. This can constitute a major problem for experiments that aim to test the effect of a treatment (e.g., a wage increase) in a within-subject design. Another undesirable effect of real effort is an unusually high variance in the results. Even after controlling for subjective abilities, real effort produces noisier results as compared to chosen effort. For instance, Brüggen and Strobel (2007) in their direct comparison between real and chosen tasks observed higher variance in real effort treatments. I have discussed several problems arising from the heterogeneity in the distribution of cognitive or physical abilities in real-effort experiments; however, it is worth to remark that chosen-effort experiments can be affected by similar problems. Camerer and Hogarth (1999) were among the first in arguing that performance, even in a chosen task, is function of two elements: (a) cognitive effort exerted to complete the task and (b) cognitive abilities relevant for the task at hand. Even though individual cognitive abilities are likely to interact with relative payments in virtually every experiment, very few studies have addressed this issue (for some notable exceptions see Awasthi and Pratt, 1990; Palacios-Huerta, 2003; Rydval and Ortmann, 2004). For instance, Rydval and Ortmann (2004) assessed the relative importance of incentives and cognitive abilities in Gneezy and Rustichini' s (2000) experiment on intrinsic motivation and found that results are by and large driven by cognitive abilities, rather than monetary incentives. I argue that heterogeneity in task-relevant abilities is shared, to some extent, by all experiments but is more salient in real-effort experiments; quite ironically, more control may be obtained in real, as compared to chosen, tasks since data on individual abilities are usually collected and taken into account in the analysis.

Do real and chosen effort lead to similar results? In an attempt to answer the third question, I presented results from real-effort experiments for the following classes of games: gift-exchange, intrinsic motivation, and tournaments. Evidence from the gift-exchange game lead to conclude that the claim by Brüggen and Strobel (2007) that "chosen effort is an appropriate way of operationalizing effort in experiments" (233) is not robust at least, and more evidence is needed to understand whether chosen effort is a reliable proxy for real effort.

Introducing more realism in lab experiments by implementing a real-effort task often comes at the considerable cost of losing some control over important variables. I argue, however, that loss of control is not as severe as it might appear at a first glance; for instance, task-relevant abilities are likely to be relevant also in chosen-effort experiments, but have far too long been neglected. While stripping away context is commonly considered of paramount importance for control, there is a considerable amount of evidence from psychology that de-contextualization can hamper the use of efficient heuristics, hence affecting performance (see, for instance, Ortmann and Gigerenzer, 1997; for a more detailed discussion on field experiments and control, see Harrison and List, 2004; Ortmann 2005).

Despite a growing number of studies implementing real effort in the lab and in the field, several issues remain open laying out a promising research agenda. For instance, a satisfactory method to elicit utility/disutility from working has not been developed so far. Little is known about the heuristics and social norms that a real-effort task can bring about, and even less is known of the processes leading to different results for real and chosen effort tasks. Moreover, there is very little evidence on how real effort affects behavior in repeated games and when strategic uncertainty is present; to my knowledge, there is surprisingly no evidence of the role of real effort in coordination games or repeated public goods games.

Chapter 3

Exploring the effects of real effort and incentives in a weak-link experiment[1]

3.1. Introduction

Employees often fail to match the actions of "co-workers" even when incentive problems are not present, 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 (Van Huyck et al., 1990); ceteris paribus, a speedy downward drift to the minimum effort for this game (a process commonly called "coordination failure") has consistently been reported in the literature.

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

[1] Financial support by the Italian Ministry of Education, Universities and Research is gratefully acknowledged. This Chapter is based on Bortolotti, S., Devetag, G., and Ortmann, A. (2009). Exploring the effects of real effort in a weak-link experiment. CEEL *Working Paper* 1-09. We wish to thank Marco Tecilla and the CEEL staff for their help in implementing and conducting the experiments. The usual disclaimer applies.

seem to increase external validity, to engineer coordination success in the lab. Even though a number of organizationally-relevant variables have been tested in a variety of coordination games, little has been done to inject more realism in this class of games, and in the weak-link game in particular (see Chapter 1). For instance, 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.

Concerns about realism of lab experiments have been advocated by both organizational scholars and experimental economists. Indeed, the experimental methodology has often been discarded among organizational scholars based on the lack of generalizability (e.g., Gordon et al., 1986), and similar questions have recently gained new momentum among experimentalists (e.g., List, 2006; Levitt and List, 2007; Schram, 2005). The gap in abstraction between controlled experiments and external reality (e.g., organizations) has been reduced in several ways, and the use of real-effort, rather than chose-effort, tasks is one of these ways. Indeed, experiments in other classes of games have demonstrated that the introduction of real effort can make a significant difference (Gneezy and List, 2006; Hennig-Schmidt et. al., 2005; García-Gallego et. al., 2008). Thus, we develop a new testbed to explore whether coordination failure in weak-link games survives real-effort implementations.

Our new testbed also provides an interesting and more natural environment to test different incentive provision mechanisms in case of task interdependence. Incentive provision is a central issue to all organizations and, despite the surge of theoretical models, “there has been little empirical assessment of incentive provision for workers” (Prendergast, 1999: 7). Following this advice, a number of experimental studies tested the effectiveness of different incentives on performance. For instance, van Dijk et al. (2001) and Erev et al. (1993) compared individual, group, and relative payments in real-effort experiments. However, in previous experiments group based incentives were set so that free-riding was possible. To our knowledge, there is so far no evidence on the relative effectiveness of individual and group incentives in coordination games in general, and in the weak-link game in particular. Whereas individual-based incentives are by and large considered more effective by practitioners, Pfeffer (1998) classifies this belief as one of the most dangerous myths about pay. Even though a variety of incentive mechanisms have been explored in games of the weak-link type none of them directly compared group and individual incentives. Indeed, previous studies focused on group-based incentives (e.g., Brandts and Cooper, 2006b; Hamman et al., 2007), a mix between individual and group incentives (Fatas et al., 2006) or incentives based on relative performance (e.g., Bornstein et al., 2002; Myung, 2008).

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).[2] 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

[2] 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.

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. Effort in both *Group* and *Individual* treatments were largely indistinguishable in our experimental study, hence providing support to some organizational evidence that individual incentives are not necessarily superior to group incentives (Pfeffer, 1998).

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 effort tasks. Second, we show that in our testbed a history of coordination failure in performing a real task can be overcome in the large majority of cases. We thus show that chosen effort might not be a reliable measure of real effort hence laying out a promising research agenda. Third, we contribute to the organizational and economic debate on the effectiveness of different incentives schemes. Specifically, we provide evidence that incentives based on the worst performance in the firm may not necessarily lead to lower performance as compared to individual incentives.

The chapter 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. Section 4 presents the results and Section 5 concludes.

3.2. The real-effort weak-link game

3.2.1. 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 period, all employees decide their individual levels of effort.[3] The employees are guaranteed a fixed wage, *F*, and a bonus rate, *B*, which is announced by the experimenter at the beginning of each period. Employees independently select an effort level $e_i \in \{0, 10, 20, 30, 40\}$ and incur a cost C_{e_i}, which increases linearly in the level of effort chosen. The per-period payoff to employee *i* is given by the following function:

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

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

[3] 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.

TABLE 3.1: Brands and Cooper (2006) Payoff Table, *B*=6 and *C*=5

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

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 3.1). 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-period reasoning might be applied by employees. Consider, for example, an employee that is paid according to Table 3.1 and has experienced a history of coordination failure in previous periods (minimum effort = 0).[4] 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.[5] 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*.

[4] 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).

[5] 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).

3.2.2. 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.[6] The task was, quite intentionally, highly 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

[6] 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 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.

complete their work.[7] 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.

3.2.3. 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 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 period, up to four slots of extra time could be bought. Thus, the per-period payoff received by any of the employees in our experiment is represented by the following payoff function:

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

where $\left[max_{j\in\{1,2,\ldots,N\}}(e_j)\right] \in \{0,1,2,3,4\}$ 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*.

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

TABLE 3.2: 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 3.2 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 3.1 and 3.2 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 period 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 period 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 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).[8] 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.[9]

3.3. Experimental design

3.3.1. Design

The aim of the present experimental design is twofold. First, our design 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. Second, we confront effort levels under two different incentives schemes; in one case, employees are paid according to worst performance in the firm (*Group* condition), while in the other case are paid according to individual performance (*Individual* condition).

[8] 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.

[9] This example makes reference to B=60 and C=50 as reported in the first panel of Table 2.

Sixteen periods 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 periods. Part 1, therefore, lasted eight periods rather than the four periods each that constituted Parts 2 and 3. We expected a decrease in the number of errors in the initial periods and a marginal role of task learning in the subsequent periods.

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 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 3.3). Our sequencing of the bonus was designed to reproduce in the lab plausible real-world scenarios,[10] following Brandts and Cooper (2006b) and Hamman et al. (2007).

TABLE 3.3: 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

[10] 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).

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.[11] 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 3.4).

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-period 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 (see Table 3.4).[12]

[11] 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.

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

The introduction of a real-effort task in the experiment necessarily calls for the need to control for subjective abilities (see Chapter 2 for a discussion). While a one-to-one relationship between performance and disutility from effort is present in standard chosen-effort experiments, performance in the present study (measured as number of errors) is the result of both effort and ability. The comparison between Group and Individual conditions allows to control, at least to some extent, for the relative contribution of effort and ability to final performance. Ability, on the one hand, tends to increase with task repetition and hence, ceteris paribus, causes the number of errors to decline over time (what we call *task learning*) in both Group and Individual conditions. On the other hand, the level of effort exerted in the Group condition 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), while effort in the Individual condition is function of individual behavior only and should be kept constant over time (since no changes in the incentives is implemented). It seems plausible to assume that the presence of uncertainty about others' effort and performance ("strategic uncertainty") tends to decrease effort with respect to a situation in which an employee is only rewarded according to his/her own performance. Hence, the comparison between Group and Individual conditions serves two main purposes. First, we ran baseline treatments without strategic uncertainty (Individual condition), for which the assumption that effort of employees is constant over time was therefore easier to rationalize, in order to disentangle the relative contribution of strategic learning and task learning. Second, contrasting performance in Group and Individual conditions should provide a measure of the relative effectiveness of the two incentive schemes, while controlling for other confounding factors, such as task learning.

TABLE 3.4: 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.

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 period(s) in our real-effort setting?

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

Moreover, we juxtapose effort over time in Individual and Group conditions to grasp a better understanding about the relative effectiveness of these two incentive schemes. We also 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? Are individual or group bonuses more effective?

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.

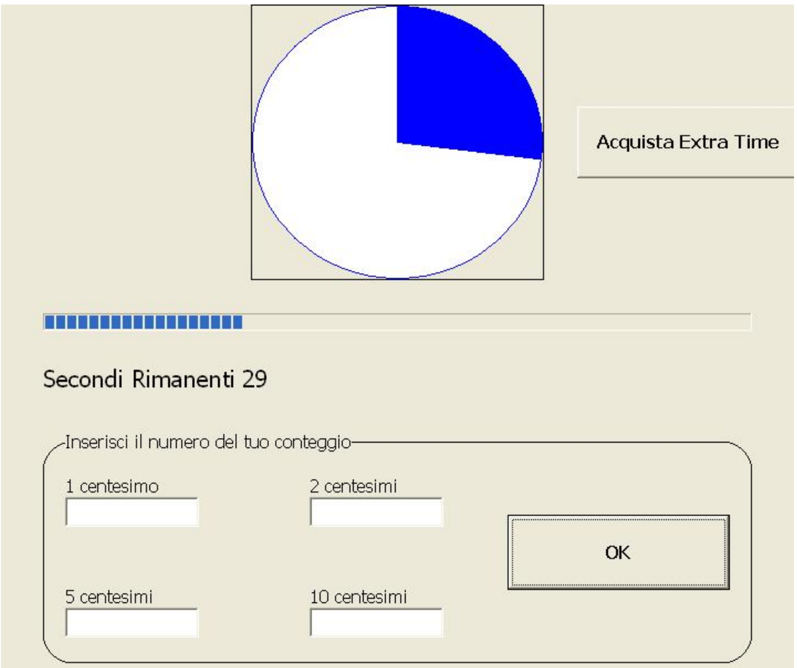
3.3.2. Implementation

A total of 144 subjects participated in the experiment (Study I), 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 period started, and before they received a small bag containing coins worth 1, 2, 5, and 10 cents, subjects were asked to wear headphones. The total number of coins contained in the bag varied from period to period and from firm to firm, while the types of coins were the same throughout the experiment.

FIGURE 3.1: Sample Screen for Counting Task



In each period, subjects had 45 seconds of regular time to perform the task; a sound informed them of the beginning of the period and a clock visualized the passing of time on each participant's monitor. In each period, subjects had the option of buying at most four slots of extra time, lasting 12 seconds each, simply by clicking a button (see Figure 3.1). 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).

3.4. 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.

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

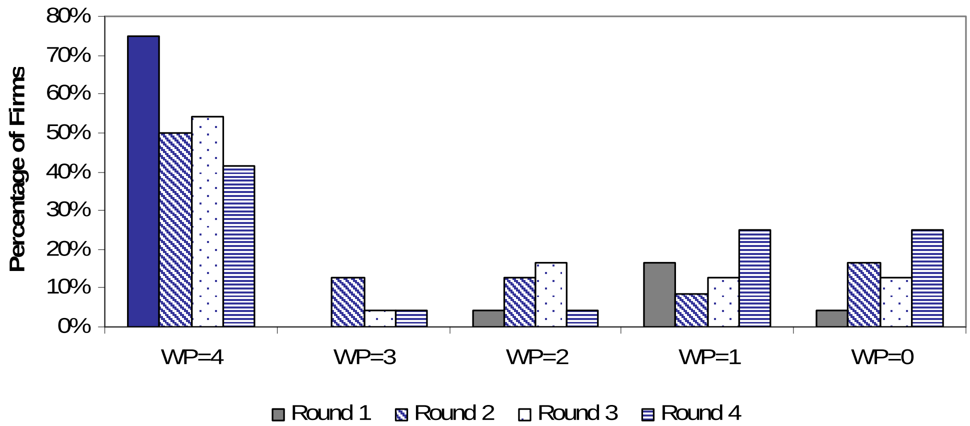
Repeated observations at the employee level are not independent, and, furthermore, employee-level data also show correlation for employees of the same firm. Therefore, when testing our hypotheses (and if not otherwise specified), our independent unit of observation is the mean value of the variable for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

3.4.1. Initial coordination

Is failure to coordinate on the efficient equilibrium observed in the initial period(s) in our real-effort setting? At the firm level, coordination failure was pervasive in the first period: 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 period; 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 period 1 for both worst performance and extra time. A clear tendency toward coordination failure at the firm level is also shown in subsequent periods.

Figure 3.2 presents the distribution of worst performances over the first four periods, by pooling all Group treatments. In each of the first four periods, modal worst performance was equal to 4 errors, while average worst performance over periods 1 to 4 was equal to 2.7 and it was greater or equal to 3 in 14 out of 24 firms.

FIGURE 3.2: Distribution of the Worst Performance in Rounds 1 through 4, all sessions pooled



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 3.5a and 3.5b (in Appendix) report the distributions of errors and extra time, in period 1 and in periods 1 through 4 respectively, divided by "types" of performers. In the first period, 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 period, 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 period (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 ($p = 0.043$);[13] 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 period is affected by strategic uncertainty. We note that more extra time does not translate into an improvement in average performance.

In subsequent periods, the frequency of employees that perform very poorly tends to decrease. However, in periods 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 3.5b 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 periods; 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.

More generally, our data suggest a positive answer to our first research question: coordination failure is observed in the initial period(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.

[13] 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.

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.

3.4.2. Dynamic Analysis: Group Conditions

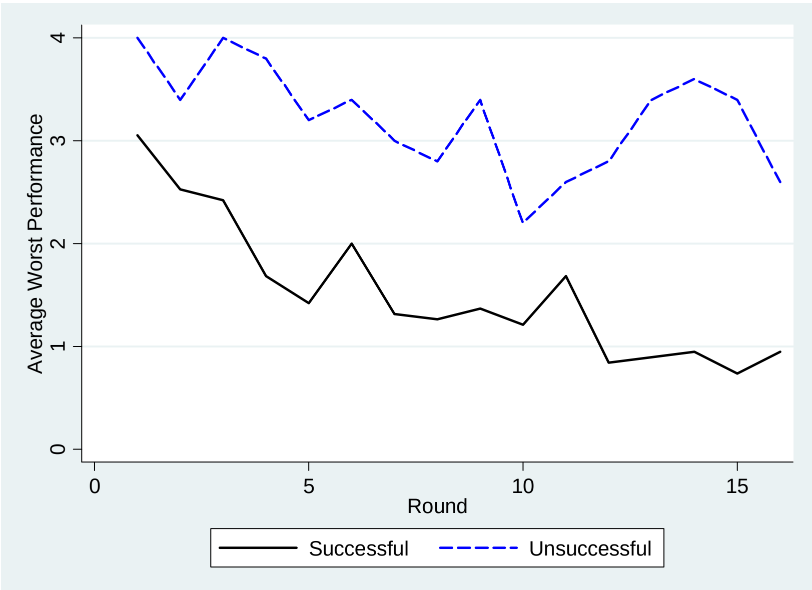
Is speedy unraveling toward the inefficient equilibrium observed in later periods, or is it possible to overcome coordination failure by introducing real effort in the lab? How and why exactly does coordination evolve over time? We now turn to a dynamic analysis of firm-level data to answer our second set of research questions, and test whether coordination failure is overcome in our real-effort experiment. Table 3.6 (in Appendix), 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 periods 5 through 8 was lower than in the first four periods for all conditions (1.85 as compared to 2.70, all treatments pooled). The average worst performance in Parts 2 (periods 9-12) and 3 (periods 13-16) was less than 2.1 for all treatments (see bottom panel in Table 3.6). 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.

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.[14] A Mann-Whitney ranked-sum test reveals that the distribution of firms' average worst

[14] 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.

performance differs for successful and unsuccessful firms (p = 0.001, two-sided).[15] Figure 3.3 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 3.3: Worst Performance over Time, by Successful and Unsuccessful Firms



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.

---

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

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 periods; second, individual adjustment dynamics.

Table 3.7 (in Appendix) reports average worst performance, errors, and extra time in periods 1 through 4, conditional on firms being successful or unsuccessful. Relative to successful firms, unsuccessful firms performed poorly in initial periods. However, if only best performers (i.e., the employee with the smallest number of errors within a firm in a period) are considered, there is no difference between successful and unsuccessful firms in periods 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).

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 period. 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. [16]

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 periods 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.

Indeed, as shown in Table 3.8 (in Appendix), 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.

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 3.9 we report the average number of errors and

[16] 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.

the average extra time bought in period *r* separately for minimum and non-minimum employees in period *r-1*.

TABLE 3.9: 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

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

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.

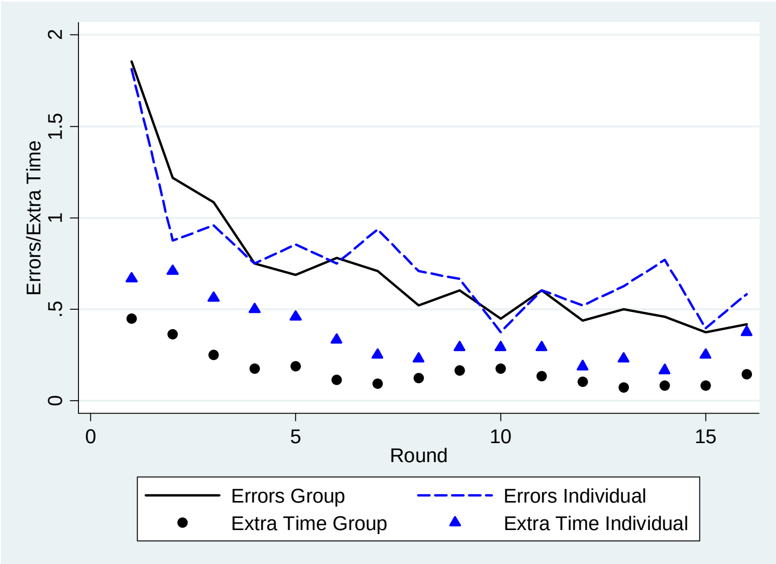
Learning Direction Theory fares poorly at explaining how and why coordination failure is overcome, while the presence of "leaders" that tend to hold steady better explains the observed behavior in successful firms.

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.

3.4.3. Dynamic Analysis: Group vs. Individual Conditions

Does strategic uncertainty undermine the learning process in the counting task? Are individual or group bonuses more effective? We now turn to the comparison between Group and Individual conditions to answer our third set of research questions on the effectiveness of different incentives schemes and 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 3.4 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 3.4: Extra Time and Errors over Time, by Group and Individual conditions



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. That is, there is evidence that incentives based on the worst performance in a group are as effective as incentives tailored on the individual performance.*

In the previous section we presented results from the Learning Direction Theory for Group data; the right column of Table 3.8 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).[17] 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 period. In the Individual conditions, only 52 percent of the employees who made 4 errors in a period made fewer errors in the following period, 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 period 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.[18]

[17] Observations from the latter range of errors have to be compared with both minimum and non-minimum players.
[18] 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-

3.4.4. 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[19]. As expected, only a few 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 periods of Part 1 to Part 2 (from 0.20 to 0.22), after the high bonus was introduced,[20] 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 3.7 shows no effect of the increase in the bonus rate on errors.

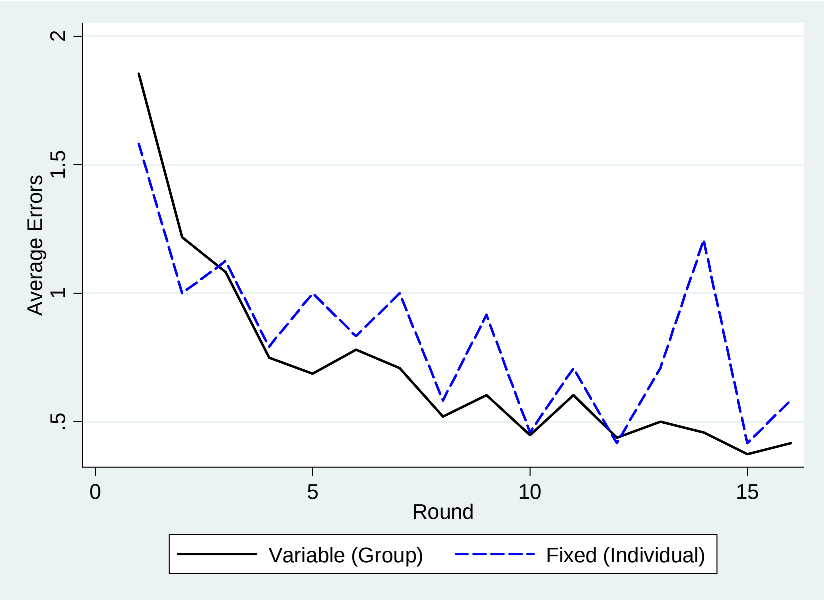
Whitney ranked-sum test. Only the difference for non-minimum employees that did 0 errors at time *t* is statistically significant at the 10% level.

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

[20] 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.

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

FIGURE 3.5: Comparison between Fixed (Individual) and Variable (Group) Bonus, all treatment pooled



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.

3.4.5. 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 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.

3.5. 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 period(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.[21]

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 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. It is quite surprisingly to notice that, despite the extreme task interdependence and highly unforgiving payoff structure of the weak-link game, "virtuous" dynamics emerging in the Group condition 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 (i.e., Individual condition). Self-esteem or self-efficacy (Bandura, 1977) may have played a role in the present set-up; indeed, the fact that a firm member managed to successfully complete a task may have led other firm members to think they could have been successful as well. 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 and to what extent these virtuous dynamics are triggered by strategic teaching, competition, social norms, or other motivations.

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

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

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 period 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.

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 (2006a) have found, partial feedback was efficiency-enhancing, but only in treatments with low deviation costs.

Appendix 3.A: Experimental Instructions

Instructions (Part 1)[22]

Instructions for Full Information and High Deviation Costs

Group condition: *(italic in brackets)*

Individual condition: [highlighted in yellow and in squared brackets]

General information: The purpose of this experiment is to study how people make decisions. [The experiment is strictly individual] 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.

Parts, periods (*and groups)*: This experiment will have 3 parts. In Part 1 there will be 8 periods. After these 8 periods have finished, we will give you instructions for Part 2 of the experiment. *(Before the experiment proper starts, the software will randomly match you*

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

with three other participants in this room. The composition of the groups, made of 4 participants each, will be the same in all periods.)

The following instructions are about Part 1 only.

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

In each period:

- *(each of the employees is)* [you are] 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 [your performance] *(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.

Performance measurement: (*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***.) [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.].

Payoffs: The payoff you will receive in a period 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 period, 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 period (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 period; 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.

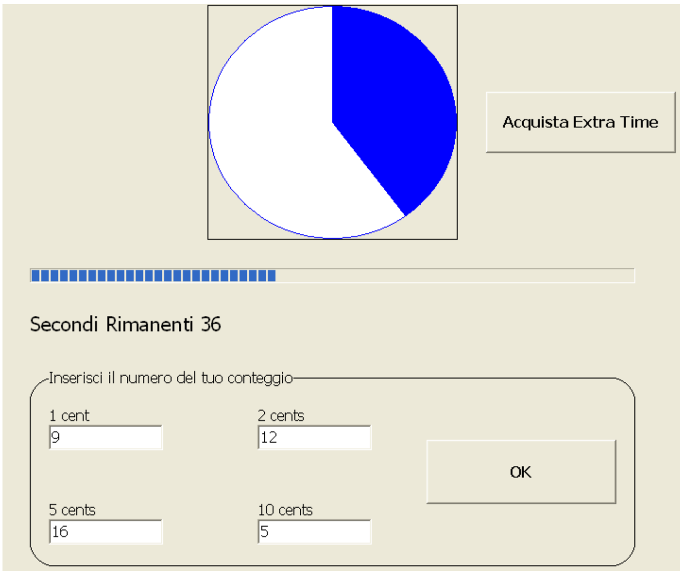
On your desk you can find a paper copy of the payoff table in order to have the possibility to easily check it whenever you want. Please, make sure you understand how to read the payoff matrix since your earnings are based on it. If you have doubts that you understand how to read the payoff table please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

<u>Practical implementation</u>:

Before each period starts you have to wear the headphones you have on your desk.

Before each period 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 period to period 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 period of the experiment, a screen like the one shown below will be displayed on your monitor. A sound will inform you that the period has started and you can start performing the task.



You are allowed to open the bag with the coins only after the period 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 period.

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 period 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

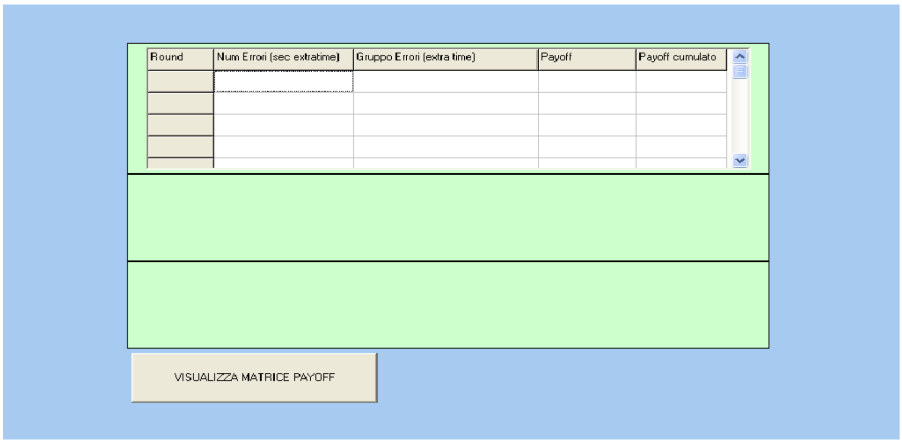
ENTER THE NUMBERS. In the case that you have not filled the cells with numbers, or you have NOT CLICKED ON 'OK', all cells will be automatically filled with zero. This will translate automatically into *(the firm*) [you] having 4 or more mistakes. *(Please recall that your earnings do not depend on your performance only, but they depend on* ***the worst performance in the group****.)*

During the period 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 period 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 period is finished and you can NEITHER BUY EXTRA TIME NOR INSERT THE NUMBER OF OBJECTS YOU HAVE COUNTED.

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

<u>Information that you will receive</u>: After each period 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 period, and your cumulated payoff is shown in the last two columns.



At the end of the 8 periods, information about that period 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.

Appendix 3.B: Results

TABLE 3.5a: Distribution of Errors and Extra Time in Period 1, by Group and Individual Conditions

		Period 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 3.5b: Distribution of Errors and Extra Time Period 1 through 4, by Group and Individual Conditions

		Period 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 3.6: Errors, Extra Time and Worst Performance by Treatment, all Session Pooled

Errors		Full50	Partial50	Full20	Partial20	Ind50Var	Ind20Var	Ind50Fixed	Ind20Fixed
Period	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
Period	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
Period	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 3.7: Worst Performance, Errors, Extra Time, and Gap (WP-BP) in Periods 1 through 4, by Successful and Unsuccessful Firms

Periods 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 periods 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 period.

TABLE 3.8: 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%		17.42%
% unchanged errors	12.50%	16.13%	11.76%	26.13%	11.36%
% decreased errors	70.11%	48.38%	74.51%	(p=0.020)	71.21%
NON-MINIMUM EMPLOYEES					
Errors <4 & >0 (Obs)	(139)	(61)	(78)		(136)
% increased errors	16.54%	27.87%	7.44%		17.42%
% unchanged errors	17.99%	18.03%	17.95%	22.82%	11.36%
% decreased errors	65.47%	54.10%	76.92%	(p=0.095)	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 period *t* to period *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.

Chapter 4

Group Pride in the Minimum-(Real)Effort Game[1]

4.1. Introduction

Matching best performers together in "elite groups" is common across different types of organizations. Delta force, Army Rangers, and SAS are only a few of the many examples of elite military groups. Similarly, organizations often group together their best employees to work on the core business or on an innovative project, and recruiting and motivating talent is becoming increasingly important among practitioners (see, for instance, Lewis and Heckman, 2006; Ready et al., 2008). Top universities and research teams can also be thought of as interesting examples of elite groups.

Best performers (in management literature, often referred to as As players) are commonly offered competitive pay, continual training, and challenging jobs in order to fully exploit their potential (Ready et al., 2008). However, as Berglas (2006) has noticed, "if you

[1] I wish to thanks Giovanna Devetag and Andreas Ortmann for their help in designing the experiment and their comments on earlier versions of this chapter, Matteo Ploner and participants to the Cifrem seminar for their comments, Marco Tecilla for programming the software, Giovanna Devetag, Sibilla di Guida, and the CEEL staff for their help in implementing and conducting the experiment. The usual disclaimer applies.

do not carefully manage the often unconscious need of these A players for kudos and appreciation, they will burn out in a way that is damaging to themselves and unproductive for you" (106). The need for motivation suggests why elite groups are often featured prominently —for instance, elite military groups are located in separate locations and have different training programs. Along the same line, firms organize meetings and social events to foster identification with a group.

Despite the relevance of the problem, assessing the relative contribution of motivation on performance in elite groups is difficult because of the paucity of field data and the presence of many confounds (performance can be affected by factors such as individual abilities, learning spillovers, higher pay, special training) hence rendering the study of motivation problematic. I thus follow Falk and Fehr's (2003) advice to intend laboratory experiments as an important complement to field evidence and a useful setting to test different organizational institutions in vitro.

In this chapter, I experimentally test how the motivation derived from being part of an elite group, or "group pride", affects coordination. In coordination games, employees may fail to match the actions of co-workers because of strategic uncertainty even when incentive problems are not present, thus causing organizations to drift into, or stay locked in, inefficient equilibria (see, for reviews, Camerer, 2003; Devetag and Ortmann, 2007b). I hypothesize that strategic uncertainty hampering coordination can be reduced in elite groups throughout two psychological elements laying the foundation of group pride: self-esteem and group trust. Being selected as a member of an elite group has a strong impact on self-esteem which has long been thought as an important human impulse and a strong motivation to act (Salancik, 1977; James, 1890). It has been shown that the desire to enhance self-esteem and reduce cognitive dissonance[2] can lead people to be more prone in

[2] Cognitive dissonance is defined as the unpleasant feeling of holding conflicting thoughts.

engaging in challenging tasks, taking risky actions, and persevere in the pursuit of an objective (Köszegi, 2006; Bénabou and Tirole, 2003; 2002).[3] Group trust, unlike self-esteem, is observed at the group level; evidence from social psychology (especially from social identity theory, SIT hereafter) and experimental economics suggests indeed that higher levels of trust are developed in cohesive groups with strong identities (for a review, see Brown, 2000). Despite the ample evidence on both self-esteem and group trust, to our knowledge, there is no evidence on the role of group pride in a coordination game.

I implemented a minimum-effort game (Study II), for which speedy unraveling toward inefficient equilibria is commonly observed, as first shown by Van Huyck et al. (1990). A real-effort experiment recreating the prototypical features of the minimum-effort game in which subjects had to count the number of letters contained in a sequence of tables was designed. There were two real-effort treatments —*Info* and *NoInfo*— and in both treatments subjects (employees, hereafter) were asked to solve as many tables as they could during the learning phase; they were then ranked accordingly. Groups of four (firms, hereafter) were formed based on the ranking —each firm was formed either by high or low performers, defined according to a median split— and ten periods of the minimum-(real)effort game were played. Following Kuhnen and Tymula (2009), the feedback about relative performance in the clerical task was varied to induce different levels of group pride. While each employee was informed about his/her type (i.e., high or low performer), and the matching protocol was common knowledge in the *Info* treatment, no information about

[3] Although self-esteem is commonly viewed as a strong motivator and a desirable attribute, important drawbacks are also present. Berglas and Jones (1978), for example, have discussed one of the possible negative consequences of self-esteem: “self-handicapping”. That is, the tendency to avoid potentially beneficial actions when there is a risk of lowering your own self-esteem. For instance, managers can dismiss a project in a new and promising area simply because they fears to feel inexpert or inadequate, thus lowering their self-esteem. Self-handicapping can also explain why performance-based incentives are underused in actual organizations (see, for example, Cowen and Glazer, 2007); high self-esteem can indeed shy away from incentives contingent upon performance because they fear to receive feedback contrasting with their self-image.

both type and matching was provided in the *NoInfo* treatment, this being the only difference between the two treatments. *NoInfo* was meant to test if any correlation between ability in solving our clerical task and strategic behavior in the stage game is in place and allows to determine the net effect of group pride on coordination. Informed by previous evidence from Study I, I also explored whether differences in behavior are present when a real-effort, rather than chosen-effort task is implemented . A *Chosen* treatment with ten periods of a standard minimum-(chosen) effort game was run.

The results show that when information about types and matching is provided, high performers manage to coordinate on significantly higher levels of effort as compared to low performers; specifically, half of the firms composed by high performers in the *Info* treatment coordinated on the efficient equilibrium, while none of the firms composed by low performers achieved the efficient equilibrium. Interestingly enough, no differences in initial choices and, as a consequence, beliefs are detected between high and low performers' firms, hence suggesting that high performers are more prone to perceive feedbacks in a more positive way and adjust their subsequent behavior accordingly. Successful coordination was not driven by differential ability in solving our clerical task, since no difference in behavior and coordination between high and low performers was found in *NoInfo* treatment, and both effort and minimum effort for high performers was statistically higher in the *Info* as compared to the *NoInfo* treatment. As for the comparison between real and chosen effort, contrary to evidence from Study I, no difference was observed; if anything, higher coordination when the chosen-effort task was implemented was found.

The present experiment contributes in several ways to the existing literature. First, I show that successful coordination is achieved when employees with high self-esteem are matched together and receive feedback about their initial performance (i.e., group pride is artificially induced). This finding is primarily interesting as a possible explanation of why many real world groups (e.g., firms, sport teams, military corpses, etc.) seem considerably more successful at achieving coordination compared to what the available experimental evidence would seem to suggest (e.g., Weber, 2006). Hence, group pride may be an element that positively affects coordination, and which may contribute to explain the gap between the level of coordination success observed in the world and in the lab. Group pride may apply to all those groups that have some (real or imaginary) reason to be proud of themselves, and these are far more numerous than so-called elite groups *strictu sensu,* including successful firms, successful sport teams, successful work teams of any kind. Group pride manipulation, moreover, can also be used as an effective tool for groups regardless of their actual performance status (high or low), since we demonstrate its effect to be present in a case in which ability on which group pride is based is irrelevant per-se for achieving coordination (as even a low ability level suffices).[4] Second, I provide a laboratory set-up which allows the study of self-esteem and group trust by using a real-effort task. Contrary to previous experiments on segregation and SIT, subjects' classification is based on the same task used in the stage game, rather than an unrelated task. This methodological innovation renders the classification closer to the one experienced in actual organizations, hence increasing the external validity of our lab experiments. Finally, the present testbed allows for the study of coordination games with real effort rather than chosen effort. As compared to previous studies on coordination and real effort, namely Study I, the testbed presents a tighter control over the payoff structure and equilibria of the game. In this study, I do not find any difference between real and chosen effort, hence suggesting that chosen effort can be a reliable proxy for real effort, at least in this specific context.

[4] While the task was the same across the two parts of the experiment, we constructed the experiment in such a way that both high and low performers could achieve the Pareto-efficient equilibrium.

The chapter is organized as follows: in Section 2 I review the relevant literature on self-esteem, SIT, and coordination games. In Section 3 I describe the minimum-effort game, while in Section 4 I present the experimental design and the implementation. In Section 5 I show the results divided by treatments and in Section 6 I discuss managerial implications of our findings and further researches.

4.2. Literature review

In this section, I summarize three streams of literature relevant to the experiment: first, I briefly summarize both theoretical and experimental evidence on the role of self-esteem; second, I present evidence on the effect of salient group identity in a strategic setting; finally, I link finding on self-esteem and SIT to the experimental evidence on the minimum-effort game.

4.2.1. Self-esteem

The role of self-esteem in a variety of work-based situations has attracted major attention from organizational scholars. While a unique and agreed definition of self-esteem does not exist (see, for instance, the concept of global, role-specific, situation-specific, and organization-based self-esteem), it is commonly believed to be an important mediator of behavior. For instance, high self-esteem employees have been shown to be more productive, to be more prone to take risk, to persevere in the face of a problem (see Bandura, 1977), and to overcome failure more easily (Lane et al., 2002). Self-esteem is also believed to be an important moderator of role ambiguity, role conflict, job satisfaction, and performance (Pierce et al., 1993). However, in some contexts, high self-esteem can be counterproductive and not desirable. A high level of self-esteem can indeed lead to an excess of risk-taking, hence leading to failure (Josephs et al., 1996), or, in other cases, to strategies of self-handicapping. It is important to notice that self-esteem's perception is constantly modified by external cues, and organizations can play a crucial role in enhancing self-esteem. For example, McAllister and Bigley (2002) have found a positive relationship between self-esteem, job fairness, and job authority.

Economists have only recently acknowledged the role of self-esteem, and a few but growing number of theoretical models have been proposed. Concerns about the maintenance of self-esteem and the desire to preserve a positive self-image have been shown to have important consequences on how agents search and process information. Köszegi (2006) incorporated the concept of "ego utility" (i.e., the agent derives utility from a positive self-image) to determine choices over different tasks characterized by different degrees of ambitiousness. Bénabou and Tirole (2002) have studied under which circumstances rational agents may want to apply self-deceptive strategies, such as avoiding new information about their ability and self-handicapping, to maintain their self-esteem; the resulting self-serving biases may have both positive and negative welfare effects. Bénabou and Tirole (2003) have shown that extrinsic incentives can be detrimental when they undermine self-esteem by revealing information about ability. Similarly, Cowen and Glazer (2007) have demonstrated that agents with ego utility may opt for jobs that do not disclose information about ability. To the best of my knowledge, only one economics experiment has tested the role of self-esteem in the lab: Kuhnen and Tymula (2008) tested the role of self-esteem, induced by feedback on relative performance in a clerical task, on performance. The authors have found that high self-esteem subjects decrease their effort and, at the same time, expect to improve their position in the ranking in subsequent periods. This finding suggests that high self-esteem and information about ranking do not necessarily have a positive impact on performance.

4.2.2. Group identity and strategic behavior

Since the path-breaking work by Tajfel and Turner was published in 1979, myriad studies in social psychology have investigated the role of social and individual identity on behavior. According to SIT, when acting as a member of a group, subjects develop —and behave according to— a social, rather than individual, identity. In psychological experiments, social identity has often been artificially induced in the lab by using the minimal group paradigm: trivial and unrelated tasks are generally used to classify subjects —deception has been widely used in these studies to ensure that the classification was irrelevant. Trivial categorization has been shown to induce in-group favoritism and out-group discrimination (for a review of the psychology literature see, for example, Brown, 2000)

The potential of social identity has been recently acknowledged also by economists. For instance, Akerlof and Kranton (2005; 2000) have proposed a neoclassical model in which norms of behavior (and relative disutility from not abiding the norm) are associated with social categories and identity. A growing number of economists, using various games, has artificially induced salient social identity in the lab and tested how it affects beliefs and preferences. Different procedures have been implemented to induce social identity without deception, such as randomly assigning column and row players to different rooms (Charness et al., 2007), dividing subjects according to their preferences for painters (Chen and Li, 2008), or randomly distributing stars (Eckel and Wilson, 2007). Contrary to previous findings in the psychology literature, some of these studies have found that the minimal-group paradigm does not have significant effects; however, when the categorization is made more salient (e.g., introduction of observers or participation in a common task) an appreciable and significant group effect is observed. In the following, we review the evidence on the effectiveness of both equal and unequal social identity.

Equal social identity: Charness et al. (2007) have found a significant effect of social identity both in the battle of the sexes and the prisoner dilemma; Chen and Li (2008) have shown that subjects are more altruistic toward an in-group member; and Chen and Chen (in prep) have observed coordination on higher strategies when matched with an in-group, rather than an out-group member in the minimum-effort game. In Hargreaves-Heap and Zizzo (2009) subjects played a trust game with in- and out-group member and had the chance to trade group membership; overall, the welfare effect of group membership was at best neutral. Güth et al. (2008) teased out group affiliation from guilt aversion in a modified dictator game and found that group affiliation explains most, but not all, the in-group bias.

Status inequality: High status subjects earned more in a market setting (Ball et al., 2001), were offered more in an ultimatum game (Ball and Eckel, 1996), contributed more in a voluntary contribution game (Kumru and Vesterlund, 2005), and gave more in the dictator game (Duffy and Kornienko, 2005). Moreover, information from high status members is valued more and results in enhanced coordination (Eckel and Wilson, 2007); high status members trust, and punish lack of trust, more than low status members do (Butler, 2008).

4.2.3. Minimum-effort game

Van Huyck et al. (1990) were the first to document initial failure to coordinate on the Pareto-dominant equilibrium and subsequent speedy unraveling to inefficient equilibria (i.e., coordination failure) in the minimum-effort game (see Chapter 2 for an extensive review of the literature). [5] This game is a very interesting framework to study the effect of group pride since both self-esteem and group trust can interact in several non-trivial ways with beliefs and coordination. For instance, it is not a priori clear whether high self-esteem will have a positive impact on coordination: in particular, will self-handicapping or propensity toward risk prevail? Although high status has been shown to have an overall positive effect on several games, high status subjects have also been shown to punish lack of trust more severely (Butler, 2008), a characteristic that could hamper efficient coordination in the game under analysis.

While many mechanisms have been implemented in the lab to enhance coordination, such as incentives, communication, and transfer of precedent, very few studies focused on the relationship between group membership and coordination. Chen and Chen (2009) applied a group-contingent social preference model to coordination games with multiple Pareto-ranked equilibria, which includes the minimum-effort game, where utility depends on both own and other's payoffs. In their model, the weight assigned to others' payoff depends on the group category of the other members of the group: the stronger the identification with the group, the greater the weight assigned to others' payoffs hence leading to coordination on higher levels of effort. The authors experimentally tested the model and found that group identity significantly increases coordination when subjects are matched with other members of their group; however, no out-group bias is observed. Bornstein, Gneezy, and Nagel (2002) introduced intergroup competition in the minimum-effort game; while the authors acknowledge competition as the main cause for enhanced coordination, competition might have induced a strong sense of belonging and

[5] Following previous literature on the minimum-effort game, we define 'coordination failure' as the inability to coordinate to the Pareto-dominated equilibria, and 'mis-matching' as the inability to coordinate on any equilibria of the game.

identification with the group (see Feri 2009 for a related discussion on group decision making and social identity).

The present experiment departs from previous evidence in at least three important ways: first, I induce asymmetric, rather than symmetric, identities in the minimum-effort game, hence reproducing in the lab the concept of high and low status groups; second, and in an attempt to enhance the external validity of the results, subjects were grouped based on performance in a real-effort task that is also used in the stage game Third, to further enhance the external validity of the study, I employed a minimum-(real)effort game.

4.3. The minimum-effort game

Firms of four employees are randomly formed before the proper experiment starts. In each period, the employees of each firm have to simultaneously choose how to allocate four units of time between two different projects: A and B. While project B pays a fixed amount for each unit devoted to the project, project A yields —possibly— a higher amount based on the minimum contribution of time to project A within the firm. Formally, employees are paid according to the following earnings function:

$$\pi_i = G\left[min_{j \in \{1,\ldots,4\}}\,(a_j)\right] + I(b_i) \qquad (1)$$

where, π_i are the earnings for employee $i \in \{1, 2, 3, 4\}$, $min_{j \in \{1,\ldots,4\}}\,(a_j)$ is the minimum contribution to project A by a member of the firm, b_i are time units devoted to project B by employee i, and G and I are the bonus for the minimum contribution to A and the individual contribution to B, respectively. In each period, exactly four units of time have to be divided between the two projects.

Table 4.1 reports the earnings table for $G = 80$ and $I = 50$; the parameters are set in such a way that coordinating on any contribution level to project A is an equilibrium. All employees choosing to contribute four units of time to project A is the Pareto-dominant strategy, while all choosing to contribute zero to project A (which means contributing four to project B) is the safe strategy. (See § 3.2.1 for a more detailed discussion of the game).

TABLE 4.1: Earnings table

		Minimum Effort by Employees				
		4	3	2	1	0
Effort by Employee *i*	4	320	240	160	80	0
	3		290	210	130	50
	2			260	180	100
	1				230	150
	0					200

Following Brandts and Cooper (2006 a; 2006b; 2007), we used a firm, rather than neutral, context and had employees to decide over different allocations of time between two projects, instead of a single level of effort; we opted for the contribution between projects for the ease of comparability between chosen- and real-effort conditions.

4.3.1. Minimum-(real) effort game

Following up on Study I, I further explore whether the choice of nominal effort level, or chosen effort, is a reliable proxy for real effort. This critical assumption has recently been questioned for a number of other game situations (see Chapter 2), and the results of Study I seems to suggest indeed that chosen effort is, at least for coordination games of the minimum-effort variety, not a good proxy for real effort. Because legitimate questions were asked about the testbed used in the previous work, in the experiment reported here a new clerical task and a new experimental set-up which allows for a better comparison between real-effort and chosen effort was introduced.

Our new real-effort task consists of counting the number of either As or Bs contained in a table; where correctly counting the As contributes to project A, while counting the Bs contributes to project B. Pokorny (2008), Abeler et al. (forth.), and Falk et al. (2008) implemented a closely related task, the difference being the dimension of the tables and the fact that they used numbers rather than letters. Four tables, for a total of 60 letters each, could be solved in each period and three trials were available for each table;[6] As and Bs were randomly drawn from a uniform distribution and the sequence of tables was the same for all employees. This task was chosen for a series of reasons: first, it is pointless and arguably not engaging, hence suggestions that no positive utility should be derived from solving the task; second, it is easy to switch from one task to the other (i.e., it is easy to switch from counting As to counting Bs, and vice versa); finally, it seems reasonable to assume that the cost of performing the two tasks is exactly the same.

This new setup was introduced to overcome two potential confounds present in Study I: namely, task learning and lack of control over perceived cost of effort. Strong economies of learning can indeed lead to a change in the structure of the game over time; specifically, a decline in the perceived cost of effort can determine a change in the riskiness of the Pareto-dominant strategy. To control for learning, and the implied cost variations, in solving the clerical task, employees had the chance to familiarize with the counting activity before playing the minimum-(real)effort game. In addition, the time allotted to the completion of the task was long enough to minimize time pressure and almost everyone could successfully complete the task within the time limits.

[6] Following previous studies using the same task, we decided to limit the number of trials for each table to prevent employees from entering numbers at random in an attempt to guess the correct number of letters contained in the table.

The second, and possibly more relevant, problem when introducing real effort in the lab is that the experimenter cannot directly observe the cost of effort and the extent to which each subject exert him/herself (for a discussion, see Chapter 2). To overcome this problem, an alternative way to capture effort can be used in the lab: measure the time spent working on the assigned, rather than other, tasks (for an example, see van Dijk et al., 2001). Hence, the experimental employees were asked to allocate their time between two tasks or projects.[7] As in the minimum-(chosen)effort game, employees can choose between project A and B: the first being paid according to the minimum contribution to the project within the group, and the second being paid on an individual base according to equation (1). This approach does not only allow for tight control over the payoff function but also well captures a variety of real world situations. For instance, employees are often assigned to both group and individual tasks and must decide how to allocate effort and time among these different projects. In decentralized organizations, where no close monitoring is possible, employees may work for private purposes, rather than working for the firm.

The earnings table is the same for both chosen and real effort, the only difference being that for the real-effort task it holds only in case all four tables are completed within a period —if one or more tables are not completed, the earnings function (1) must be used. Note that tables that were not completed, or not completed correctly, were not paid. Even though nominal earnings are the same for both real and chosen effort, the subjective and unobservable cost of effort ought to be taken into account in the analysis when real effort is implemented. Under the assumption that the subjective cost of effort for completing one table, e_i, for project A or B is the same, the new set of equilibria can be derived. In particular, three cases can be distinguished. First, the cost of solving one table for at least

[7] We decided not to have employees to allocate their time between work and leisure since leisure in repeated game is extremely complicated to implement in the lab and it is a potential source of major loss of control.

one employee in the firm is higher than the bonus for the minimum contribution to project A: that is, $e_i > G$. In this case, the employee has no interest in completing any table for neither project A or B, and hence the only equilibrium in the game is to contribute zero to project A. Second, the cost of solving one table for an employee is lower than the bonus for project B, but not A: $I < e_i < G$. Therefore, for such an employee is optimal to contribute zero to project B and match the minimum number of tables contributed to project A within the firm. Third, the cost of completing a table is lower than the lower bonus: that is, $e_i < I$. It is straightforward to see that in this case the set of equilibria is exactly the same as for the chosen-effort game.

4.4. Experimental design and behavioral predictions

In the present study (Study II), I address two main questions: first, I investigate to what extent different levels of group pride influence behavior in a minimum-(real)effort game; second, I provide further evidence on the role of real, rather than chosen, effort in the context of a coordination game. To address the first question I run two real-effort treatments, *Info* and *NoInfo*, while a treatment with a chosen-effort task was run to answer the second research question (see Table 4.2).

The two real-effort treatments, summarized in the second and third column of Table 2, were composed by two parts each. In Part 1 (also referred to as "learning phase"), employees were given three minutes to solve how many tables as they could; the objective of the learning phase was twofold: first, familiarize employees with the task so to reduce task learning in the stage game; second, rank employees according to their ability to quickly solve the clerical task.

TABLE 4.2: Experimental design (Study II)

	Info	*NoInfo*	*Chosen*
Effort	Real	Real	Chosen
Part 1 (learning phase)	Yes	Yes	No
Matching in Part 2	Based on performance in Part 1	Based on performance in Part 1	Random
Information about the matching protocol	Yes	No	Yes
Total # of subjects	60	48	12
# of sessions	4	3	1

NOTE: Sixteen subjects participated in all but two sessions (one for the *Info* and one for the *Chosen* treatment) in which only twelve subjects participated.

In Part 2 (also referred to as "coordination phase"), employees were matched in firms of four according to their position in the ranking and played ten periods of the minimum-effort game. Specifically, employees were divided in high and low performers according to a median split, and each firm was composed only by either high or low performers (HP and LP, respectively). To test for the role of group pride on coordination, we information received by employees before playing the coordination phase was varied. In the *Info* treatment, employees were informed about the matching procedure and their type — either HP or LP. In the *NoInfo* treatment, employees did not receive information about both their type and the matching protocol; this treatment was meant to control for any possible correlation between performance in Part 1 and strategic behavior in Part 2.

Since employees were not informed about the matching procedure, the *NoInfo* treatment also provides us with a measure of coordination under real effort. In order to test for differences between real and chosen effort, I also run a standard minimum-effort game. In the *Chosen* treatment, employees were randomly matched in firms of four and played ten periods of the minimum-(chosen)effort game: the learning phase was not implemented since employees were not classified according to their ability, and there was no need to familiarize with the counting task, since no real-task was involved in the coordination phase.

4.4.1. Implementation

A total of 120 subjects, divided in 8 sessions as detailed in Table 4.2, participated in the experiment, which was conducted at the Computable and Experimental Economics Lab (CEEL) at the University of Trento. All subjects were students from the University of Trento and were recruited through emails, 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, the instructions were read aloud to assure common knowledge. After reading the instructions, a questionnaire was distributed to ensure correct understanding; in case of incorrect answers, the instructions were read again.

Info and *NoInfo* treatments were divided in two parts. In Part 1, subjects were asked to solve as many tables as they could within three minutes time. Fifteen tables containing As and Bs (see Figure 4.1) were displayed on the screen one after the other; for each table, subjects could choose to work for either project A —which implied counting the number of As— or project B —which implied counting the number of Bs. Subjects had to enter the number of either As or Bs in the corresponding cell and confirm their choice by clicking "OK"; a new table was shown in case the result of the counting activity was correct. Subjects had three trials for each table, after three incorrect trials the next table was automatically displayed on the screen. A bar on the bottom part of the screen displayed the time left and no number could be entered after the time was expired. Subjects were paid a fixed amount of 3 Euros for Part 1 no matter how many tables (and to which project) they completed and received information about the number of tables correctly completed for each project.[8]

When analyzing the learning phase an important caveat is in order: a technical problem in five of the real-effort sessions was encountered.[9] Specifically, there were errors in the sequence of the number of As and Bs contained in each table in the software; namely, the number of letters contained in table 2 and some of the subsequent tables was wrong. As a consequence, all the employees in the "non-regular" sessions received a wrong feedback when completing table 2 —the similar problems in subsequent tables were not noticed by participants since none of them reached that point; this was quite possibly a consequence of the repeated attempts at getting the counting for table 2 right. More troublesome is that (some) subjects might have thought that we were playing games with them, as it is likely that the second or third time around they paid particular attention to get the numbers right. We will return to this issue in the Results section.

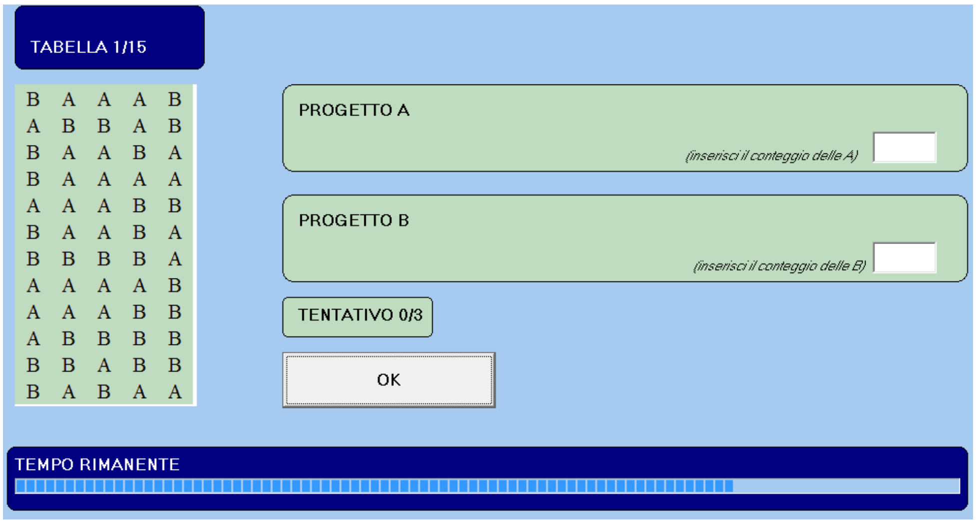
In Part 2, subjects were matched in groups of four and played 10 periods of the stage game; the implementation was exactly the same as in Part 1 the only difference being that the maximum number of table displayed was four per period, rather than fifteen. In both treatments —*Info* and *NoInfo*— the matching depended on the performance in Part 1. Specifically, subjects were ranked according to the number of tables successfully completed in the learning phase (i.e., performance); in each session, subjects ranked 1 through 8 and 9

[8] Subjects were not paid on a piece-rate in Part 1 since I did not want to have HP to start off from a higher wealth level as compared to LP. Therefore, the classification might be marginally influenced by the 'docility' and the desire to please the experimenter of the subjects and not only on their ability in performing the task; however, I can control how this potential confounding affect behavior in the coordination game, since the effect should cancel out when confronting the two treatments.

[9] Two sixteen-subjects sessions were not affected by the technical problem: one *Info* and one *NoInfo* sessions.

through 16 were referred to as “high performers” and “low performers”, respectively.[10] Each group was composed only by subjects of the same type (that is, only HP or LP);[11] while in *NoInfo* treatment subjects were not informed about the matching procedure, in the *Info* treatment subjects received an envelope in which their type was reported. At the end of each period, information about individual contribution to the two projects, distribution of contribution to project A within the group, and earnings (both cumulated and for the current period) were displayed on the screen. Subjects were paid according to the earnings table reported in Table 4.1.

FIGURE 4.1: Screenshot



In the *Chosen* treatment, subjects played 10 periods of the standard minimum-effort game. In each period, subjects had to allocate four units of time to project A and B; only integer numbers could be entered in the two cells and the sum had to be equal to four. As for the real-effort treatment, information about individual contribution to the two projects, distribution of contribution to project A within the group, and earnings (both cumulated

[10] In the *Info* treatment, only twelve subjects took part in one of the experimental sessions; in this case, subjects ranked 1 through 8 were classified as HP, while the remaining four subjects (ranked 9 throughout 12) were classified as LP, and this was common knowledge among participants.
[11] Employees ranked 1st, 3rd, 6th, and 8th were grouped together in firm A; 2nd, 4th, 5th, and 7th in firm B; 9th, 11th, 14th, and 16th in firm C; 10th, 12th, 13th, and 15th in firm D. This protocol was adopted to ensure that ability was evenly distributed across firms of the same type.

and for the current period) were displayed on the screen at the end of each period. Subjects received a show-up fee of 3 Euros plus the amount of money they made during the experiment (each period was paid according to Table 4.1) At the end of the experiment, relevant socio-demographic information was collected for all treatments.

Sessions involving real effort averaged about 90 minutes, while for the *Chosen* treatment the average session lasted about 60 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 200 ECUs. Subjects knew the conversion rate in advance and were paid privately their earnings plus a fixed amount for Part 1 (in the *Chosen* treatment, subjects received a show-up fee) of 3 Euros.

4.4.2. Behavioral predictions

Based on pilot experiments, we calibrated the time allotted to the completion of the four tables so that there was no time pressure and virtually everyone could complete the task within the time limits. Therefore, the contribution to project A should be influenced by strategic motives only and not by the impossibility to complete all four tables; in this sense, the grouping of HP and LP is based on a characteristic that is only seemingly-relevant. Moreover, as long as the subjective cost of completing one table is lower than the bonus for project B, the decision to solve a table for either project A or B exclusively depends on the expectations about the minimum and not on the perceived cost of effort itself. Thus, heterogeneity in the distribution of abilities should not affect strategic uncertainty per se.

Since experimental employees are assigned to each category at random, the *NoInfo* treatment controls if and to what extent our segregation mechanism per-se affects coordination. Specifically, I expect the ability to quickly solve the task, and any other dimension which may affect coordination levels —such as ability to see more deeply into game structure, attitudes toward risk, and trust— not to be correlated.

> HYPOTHESIS 1: *When no information about types and matching procedure is provided (NoInfo treatment) there is no difference in both average effort (henceforth simply "effort") and minimum between HP and LP in either first or subsequent periods.*

In the *Info* treatment, employees' group pride was manipulated by providing information about type and matching in the learning phase. Although ability as such is expected to have no influence on coordination (Hypothesis 1), I hypothesize that beliefs and trust, and therefore coordination, are influenced by the initial performance through group pride. More specifically, we conjecture that, in our coordination game, group pride helps achieving higher levels of coordination over time; in contrast, I expect a low level of group pride to further hamper coordination.

> HYPOTHESIS 2: *When information about the matching procedure is provided (Info treatment), I expect: (a) both effort and minimum to be higher in HPfirms, as compared to LPfirms; (b) both effort and minimum to be higher in HPfirms, as compared to HPfirms in NoInfo; (c) both effort and minimum to be lower in LPfirms, as compared to LPfirms in NoInfo.*

Finally, real and chosen effort are juxtaposed. Informed by Study I, I expect higher coordination when a real, rather than chosen, task is implemented. *NoInfo* provides a measure of coordination in presence of real effort; indeed, it can be assumed that the matching protocol per-se does not bias the results if no difference is observed between different types of firms (i.e., Hypothesis 1 holds), despite the fact that employees are not randomly assigned to firms.

HYPOTHESIS 3*: Both effort and minimum are higher in NoInfo as compared to Chosen treatment.*

4.5. Results

This section is organized around the three hypotheses and the corresponding results are provided. Evidence on real-effort and chosen-effort experiments is discussed separately, with behavior in real-effort treatments presented first. Since we are interested in the effect of ability and group pride on both initial beliefs and subsequent coordination dynamics, initial behavior and behavior in subsequent periods are discussed separately.

Observations at the employee level show correlation across both periods and employees of the same firm, with the only exception being the first period; therefore, when testing our hypotheses in periods other than the first (and if not otherwise specified), our independent unit of observation is the mean value of the variable for a single firm. Thus, there a total of 30 independent observations, divided across treatments as follows: 15 observations for *Info* (8 HP + 7LP), 12 observations for *NoInfo* (6 HP + 6LP), and 3 observations for *Chosen*. All tests are two-sided. In the following, I use "effort" and "minimum" to refer to individual and minimum contribution to project A, respectively. The term "mistake" refers to the number of tables not successfully completed for one of the two projects within a period. It might be useful to recall that in the real-effort treatments, the sum of the individual contribution to project A and B can be smaller than four in the case one or more mistakes in the counting task have occurred —in the *Chosen* treatment the sum of the contribution to project A and B is always equal to four. With the term "regular" we indicate the two sessions in which no technical problems occurred, the remaining sessions are labeled "non-regular".

4.5.1. Real-effort treatments: Learning phase

During the learning phase, employees were asked to complete as many tables as they could to familiarize with the task; a maximum of fifteen tables could be completed within the three minutes time allotted to this phase. The number of solved tables was used as a proxy for ability; indeed, the number of solved tables in the learning phase depends on two main factors: first, innate ability in solving the clerical task; second, employees willingness to work hard even if no contingent payment is present —employees were, in fact, paid a fixed amount of money, regardless of the performance.

The number of tables solved during the learning phase ranged from 0 to 11, the average number of solved tables was 3.23 and the median was 2 with more than 30 percent of the employees solving two tables. LP solved 2.15 tables, while HP solved 4.23 tables on average.

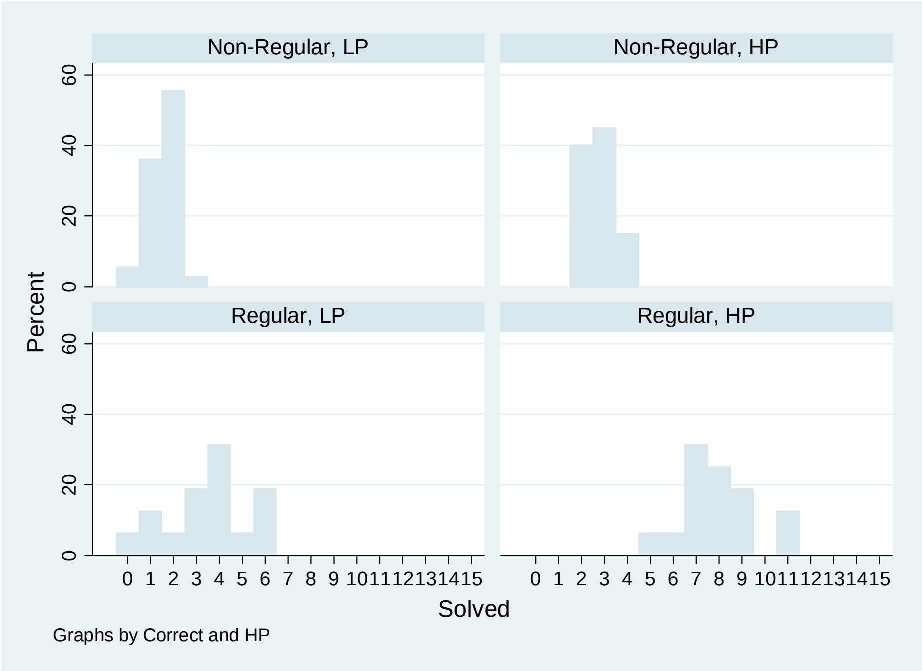
As mentioned, a technical problem was encountered in the learning phase and feedback about the counting task was not correct for table 2, affecting 70 percent of the employees participating in the real-effort sessions. It is not possible to say how exactly this problem affected the number of completed tables and the ranking. Since employees had to count three times table 2 before moving to the subsequent table, the number of solved tables in non-regular sessions was very low (it ranged from 0 to 4); therefore, the difference in performance between HP and LP was small, hence casting doubts on whether such a classification reflects a real difference in ability or is mainly due to the technical problem.

The technical problem may have interacted with behavior and beliefs in the coordination phase in a number of ways. First, it may have lowered employees' trust in the feedback they received, although this probably lost importance as the experiment progressed and as employees saw that the results of their subsequent counting were correctly recognized by the software. Second, the technical problem may have affected the correct measure of ability, as employees wasted time repeatedly counting table 2, hence not allowing to properly control for correlation between ability and strategic behavior in the *NoInfo* treatment. Finally, the technical problem may have demotivated LP and possibly surprised HP of their being classified among the best despite the very low number of tables completed in the learning phase. Since the technical problem may have influenced the data in a number of non-obvious ways, I discuss results both at the aggregate level and divided by regular and non-regular session.

In regular sessions, the number of correctly solved tables ranged from 0 to 11. Figure 4.2 reports the distribution of solved tables by type of session (i.e., regular and non-regular) and type (i.e., HP and LP); an eyeball examination reveals important differences between regular and non-regular sessions for both HP and LP (two-sided Wilcoxon rank-sum test, $p < 0.001$).[12] Notwithstanding, it is important to notice that the difference in the number of solved tables between HP and LP, according to a two-sided Wilcoxon rank-sum test, is statistically significant at one percent level also for non-regular sessions.

[12] Since the task in Part 1 was individual, we used employee level data to test our hypothesis.

FIGURE 4.2: Distribution of solved tables by session and performance



4.5.2. Real-effort treatments: Coordination phase

Table 4.3 reports effort and minimum in the first period separately by treatment (i.e., *Info* and *NoInfo*) and type (i.e., HP and LP) for real-effort treatments. Initial failure to coordinate on the efficient equilibrium is consistently observed across treatments. While no firms coordinated on the Pareto-dominant equilibrium and only one on an effort level of 3, the majority of the firms coordinated on an intermediate or low effort level (i.e., 2 or 1); in particular, 10 firms out of 27 started out with a minimum of 2 and 8 firms coordinated on the secure strategy (i.e., minimum equal to 0). The average effort was 2.60, while the modal effort was 4 with more than 30 percent of the employees choosing the highest contribution to project A. At the employee level, the first period behavior presents a high degree of heterogeneity with the great majority of effort choices (88 out of 108) ranging from 2 to 4: 26 percent of the employees exerted an effort level of 2, 23 percent an effort of 3, and 31 percent an effort of 4. A two-sided Wilcoxon rank-sum test[13] reveals no statistically significant differences in the effort (p = 0.471) between *Info* and *NoInfo* treatments, while the difference in the minimum between the two treatments is significant at 5 percent level (p = 0.048). By pooling all real effort sessions, effort provision in the first round is not different for HP and LP, as found by a two-sided Wilcoxon rank-sum test (p = 0.591).

TABLE 4.3: Effort and minimum by treatment and type: Period 1

	NoInof	*Info*
High Performers	2.29	2.97
	(1.42)	(1.63)
Low Performers	2.66	2.61
	(1.50)	(0.86)
Overall	2.48	2.80
	(1.46)	(1.33)

NOTE: Minimum effort by treatment and type reported in brackets. Real-effort treatments only.

In line with previous experiments, a decline in the effort provision over time is observed; indeed, average effort goes from 2.60 in the first period to 1.41 in the last period of the stage game. Table 4.4 shows a speedy unraveling toward the Pareto-dominated equilibrium for the majority of the firms (17 out of 27 firms by the last period). A two-sided Wilcoxon rank-sum test reveals no statistically significant differences in both effort (p = 0.880) and minimum (p = 0.471) between *Info* (15 observations) and *NoInfo* (12 observations) treatments over the ten periods of the stage game.

More generally, aggregate level data for our real-effort treatments follows by and large the pattern commonly observed in previous experiments, and speedy unraveling toward the Pareto-dominated equilibrium is observed. In the following paragraphs, I will

[13] All tests for the first period are based on employee level data.

discuss *Info* and *NoInfo* treatments separately to gain some insights about the effectiveness of our group pride manipulation.

TABLE 4.4: Number of firms by period and minimum, all real-effort sessions pooled

	Period									
	1	2	3	4	5	6	7	8	9	10
0	8	11	12	14	15	16	15	15	17	17
1	8	4	7	5	3	3	4	3	1	3
2	10	9	4	2	3	2	2	3	4	2
3	1	3	1	4	4	2	2	2	1	1
4	0	0	3	2	2	4	4	4	4	4
Total	27	27	27	27	27	27	27	27	27	27

NoInfo treatment: initial coordination and dynamic analysis

I now focus on the *NoInfo* treatment to test Hypothesis 1; that is, to see whether any correlation between ability in the learning phase and coordination exists when no information about performance in the learning phase and firm composition is provided. As hypothesized, no difference in initial behavior between HP and LP was detected. If anything, LP performed better as compared to HP in period one; in fact, effort in the first period was higher, albeit not statistically significant at any conventional level ($p = 0.686$, two-sided Wilcoxon rank-sum test), for LP (24 observations) as compared to HP (24 observations). The same behavior is observed if only regular session are considered.

Turning to a dynamic analysis, the data provide evidence that coordination failure is common in the experimental firms. In the *NoInfo* treatment, average error in period 1 throughout 10 was 1.71 and average minimum was 0.87. Nine out of 12 firms converged to the worst equilibrium and the highest minimum achieved throughout the experiment was 2. Effort supplied by HP lays below the effort provided by LP for most of the periods; though, a two-sided Wilcoxon rank-sum test reveals no statistically significant differences, (p = 0.522), in average effort between HP (6 observations) and LP (6 observations) when no information about types and firm composition is provided. Partially in contrast with the evidence at the employee level, the minimum effort for HP is higher, albeit not significant (p = 0.326, two-sided Wilcoxon rank-sum test), as compared to LP. Notwithstanding, the aforementioned difference is mainly determined by two experimental firms and, most importantly, no difference between HP and LP firms is observed in the regular session (i.e., session in which no ranking problems occurred in Part 1). In the non-regular sessions, almost no heterogeneity in the number of solved tables is observed, thus weakening the relationship between innate ability in solving the task and classification as HP or LP —since ties were very common in the learning phase, some of the positions in the ranking were determined at random. More generally, there is no evidence that ability in the counting task have any influence on coordination if no information about type and firm composition is provided; hence, HP and LP do not differ in some characteristics that systematically influence coordination.

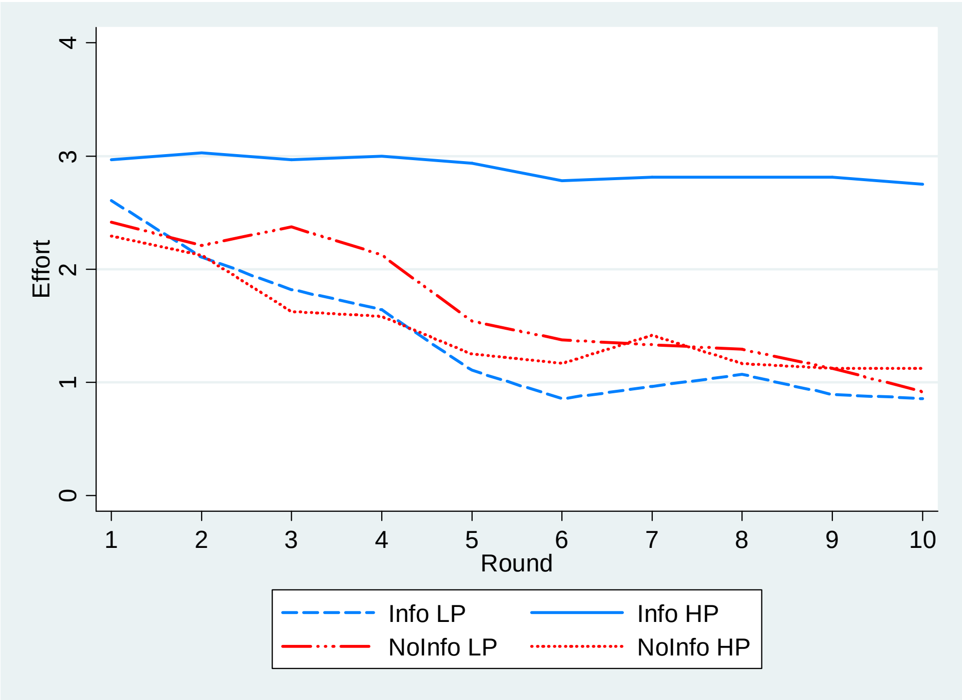
RESULT 1: When no information about types and matching procedure is provided (NoInfo treatment), there is no difference in both effort and minimum between HP and LP.

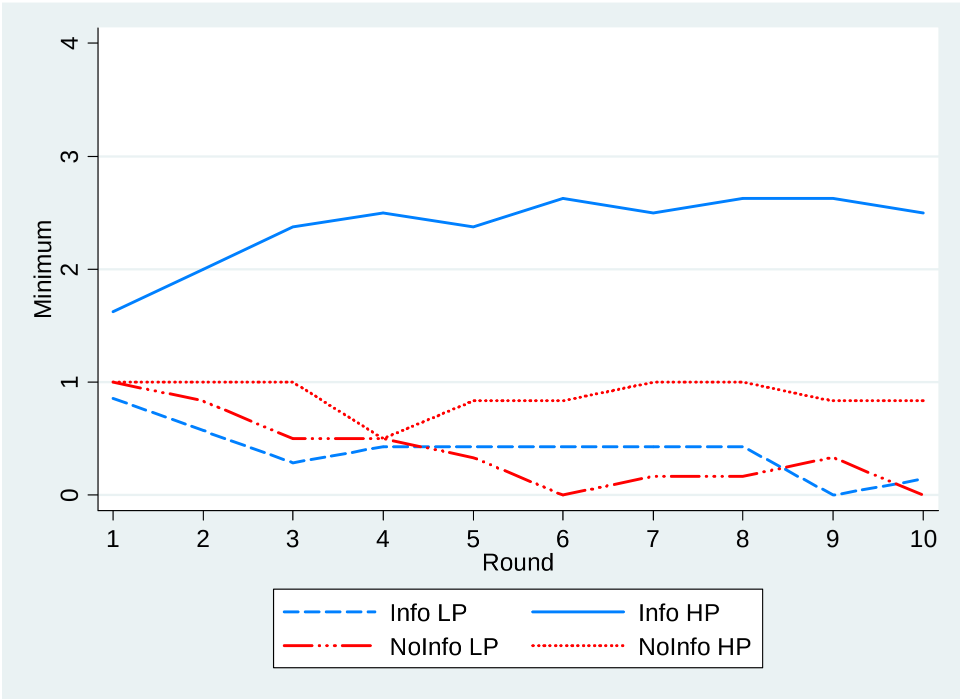
Info treatment: initial coordination and dynamic analysis

To test Hypothesis 2a throughout 2c, I consider the *Info* treatment and the role of group pride in our minimum-(real)effort game. It might be useful to remind that, if any difference in behavior between high and low self-esteem employees exists, it is not due to differences in ability. In fact, no correlation between ability in solving the task and behavior in the coordination phase has been found in the *NoInfo* treatment (see Result 1). Therefore, any difference in effort provision or minimum results from the inducement of different group pride levels.

I first consider the difference between HP and LP firms in the *Info* treatment, to test if any difference exists between firms with high and low group pride (Hypothesis 2a). In the first period, group pride played a marginal role on coordination. Average effort in period one was 2.96 for HP firms and 2.61 for LP firms; this difference was not statistically significant according to a two-sided Wilcoxon rank-sum test. Similarly, higher, albeit not significant ($p = 0.128$), levels of coordination were achieved in HP as compare to LP firms (1.63 vs. 0.87).

FIGURE 4.3: Effort (upper panel) and minimum (bottom panel) by treatment and performance





A dynamic analysis of the data reveals that coordination failure is commonly observed in the experimental firms; notwithstanding, important behavioral differences between HP and LP are present when information about types and matching is provided. Indeed, similarly to firms in *NoInfo* treatment, LP*Info* firms commonly converged to inefficient equilibria, while HP*Info* firms presented an upward trend for both effort and minimum (see Figure 4.3). According to a two-sided Wilcoxon rank-sum test, both effort ($p = 0.037$) and minimum ($p = 0.019$) are statistically different at five percent level between HPInfo and LPInfo. The data suggest that differences in group pride play an important role on coordination.

RESULT 2a: *When information about the matching procedure is provided (Info treatment), there is difference in both effort and minimum over time between HP and LP firms. Specifically, high levels of group pride positively affect coordination.*

To test Hypothesis 2b, I now compare HP firms in the two real effort treatments. Half of the HP firms in the *Info* treatment coordinated on the efficient equilibria (i.e., minimum equal to 4); interestingly enough, none of the HP firms in the *NoInfo* treatment coordinated on the Pareto-efficient equilibria. Moreover, a speedy unraveling toward the inefficient equilibrium (i.e., minimum equal to 0) is observed only in 2 out of 8 HP*Info* firms, while this behavior was observed in the vast majority of the remaining experimental firms (15 out of 19 firms). The provision of feedback about performance to HP and, hence, the inducement of high group pride had a beneficial effect on effort provision and coordination; indeed, both effort ($p = 0.053$) and minimum ($p = 0.092$) are statistically different for HP*Info* and HP*NoInfo* according to a two-sided Wilcoxon rank-sum test. The observed behavior is even more surprising if we consider that effort and minimum in the first period were almost identical across types and treatments.

RESULT 2b: *Higher levels of effort and minimum are observed over time when HP received information about performance and group composition (Info treatment) as compare to the situation in which no information was provided (NoInfo treatment).*

Hypothesis 2c is about the impact of feedback on LP. As shown in Figure 4.3, firms composed by LP follow a similar pattern, and coordination failure is pervasive, in both *Info* and *NoInfo* treatment. It is worth to notice that the knowledge of being part of a LP firm had a negative, even though small, effect of effort provision; average effort across periods was 1.39 for LP*Info* as compared to 1.69 for LP*NoInfo*. Despite the difference is not significant at any conventional level ($p = 0.253$, two-sided Wilcoxon rank-sum test), it has to be noticed that effort is bounded form below and the effort supplied in LP*NoInfo* was already very small, hence rendering the difference of some interest.

RESULT 2c: *Lower levels of effort are observed over time when LP received information about performance and group composition (Info treatment) as compare to the situation in which no information was provided (NoInfo treatment).*

Regular and non-regular sessions are overall very similar in terms of both effort and minimum. Notwithstanding, an important difference deserves to be mentioned. In non-regular sessions there is no difference in initial behavior between HP and LP. In line with Pierce et al. (1993), in non-regular sessions, LP showed grater negative reaction to initial effort and minimum as compared to HP that, starting off from the same coordination level, were able to climb the strategy space and achieved successful coordination. In contrast, in the regular session, the provision of feedback and the inducement of high and low group pride had an impact already in the first period. Specifically, in the regular session, initial effort and minimum in HP firms is higher as compared to LP firms; moreover, LP firms showed a lower level of coordination as compared to LP firms in the *NoInfo* treatment. This evidence suggests an immediate effect of high and low group pride in the regular session, while the same effect was delayed in the remaining sessions.

Mistakes and equilibria

The previous analysis focused on the number of tables correctly solved for project A. However, a low level of contribution to project A do not necessarily implies a high level of contribution to project B, since employees can fail to correctly solve one or more tables. In order to minimize this potential confound, and through pilot experiments, the time interval allotted to each period was calibrated so that virtually everyone could easily solve the four tables. In the following, I analyze to what extent inability to solve the task affected behavior and equilibria in the game.

As expected, only few mistakes were made across the ten periods of the stage game; indeed, the overall number of mistakes was 60, 1.39 percent of the total number of table to be solved. Mistakes were evenly distributed over time, with the first two periods being the only exception; in particular, mistakes were slightly higher in the first two periods (11 mistakes out of 432 tables in each period) as compared to subsequent periods (on average, 4.75 mistakes per period). Even though a difference between the first two and subsequent periods is present, the limited number of mistakes suggests that the effect of learning in performing the task is negligible. Data at the employee level show that almost all the employees were able to solve —and actually solved— the clerical task within the time limits. In fact, only 12 out of 108 employees made more than one mistake over the ten periods of the stage game.[14] Appreciable differences in the distribution of mistakes across treatments or firms are not present. The number of mistakes is very small for all types of performers, but it is higher for LP as compared to HP (average mistakes 0.92 vs. 0.21 out of 40 tables). Interestingly enough, mistakes were fewer in non-regular as compared to regular sessions; this result seems to suggest that the technical problem occurred during the learning phase induced employees to count even more carefully during the coordination phase.

Evidence on mistakes distribution allows us to gain some insights into the cost of exerting effort. Although it is difficult to precisely quantify the perceived cost of solving one table, it is possible to define an upper bound for the aforementioned cost, hence having tight control over the payoff function and the structure of the equilibria. Since the number

[14] Specifically, four employees made four or more mistakes each, three made three mistakes, and five made two mistakes.

of the tables not completed was very small and everyone decided to complete at least one table, it is possible to safely assume that the perceived cost of completing one table was lower than the bonus for completing a table for project A (i.e., $e_i < G$). Moreover, 27 out of 29 of the employees who made a mistake contributed to project B at least once. Thus, the perceived cost of effort for almost all the employees was lower than the bonus obtained for completing a table for project B (i.e., $e_i < I$). The latter inequality being satisfied assures that the structure of the real-effort game is comparable to previous minimum-(chosen) effort experiments. While the ex-post analysis of mistakes' distribution suggests that the structure of the game was successfully recreated, it is not possible to have control over employees' beliefs about the number of mistakes in their firm. This issue can be of some relevance since some employees may have thought that the Pareto-dominant equilibrium could not be achieved since not all the members of the groups were able to correctly solve all four tables.

4.5.3. Chosen-effort treatment

A baseline treatment with chosen effort to control for possible behavioral differences between real and chosen effort was run. Average minimum in the first period was equal to 2.00 and none of the three experimental firms coordinated on the Pareto-efficient equilibrium. At the employee level, average effort was 3.08, and the modal choice was 3 (50 percent of the employees chose the second highest effort level) in the first period.

In line with previous evidence, minimum effort declines over time and converges toward inefficient equilibria. As shown in Figure 4.4, data at the firm and employee level follow similar patterns, thus determining a roughly constant wasted input.[15] Figure 4.5,

[15] Wasted input is defined as the difference between effort and minimum.

however, reveals important differences across firms. One of the three firms achieved and maintained almost successful coordination (i.e., minimum equal to 3), while speedy unraveling toward dominated equilibria was observed in the remaining firms; that was the case also in firm “2” (see Figure 4.5) where almost no wasted input was present in the first period. Interestingly enough, behavior of non-minimum employees in firm 2 was quite surprising and inconsistent with learning direction theory, in both period two and three.

FIGURE 4.4: Effort and minimum over time, all firms pooled: *Chosen* treatment

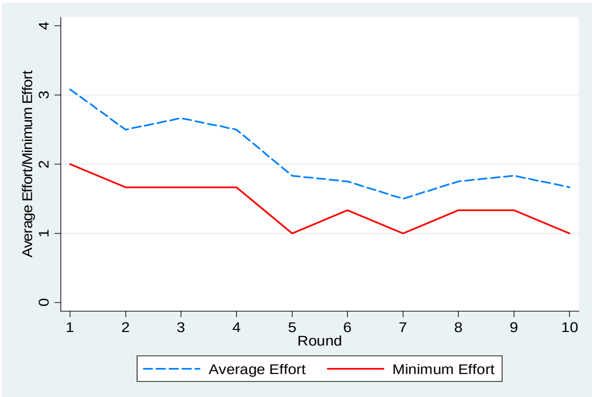
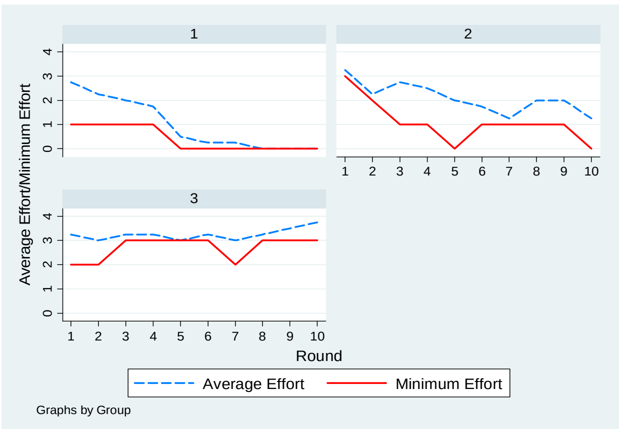


FIGURE 4.5: Effort and minimum over time by firm: Chosen treatment



I now turn to the comparison between real and chosen effort. Given that there were no differences between HP and LP when no information about performance and firm composition was given to the employees, *NoInfo* treatment can be compared to the *Chosen* treatment to test whether chosen effort is a reliable proxy for chosen effort. Contrary to Hypothesis 3 and previous evidence (Study I), we found no difference in either effort or minimum between real and chosen effort in our testbed; if anything, we observed higher levels of coordination in *Chosen* as compared to *NoInfo* treatment.

> RESULT 3*: Contrary to what hypothesized, neither effort nor minimum are higher in NoInfo as compared to Chosen treatment. If anything, coordination is more successful when chosen, rather than real effort is implemented.*

4.6. Discussion and conclusions

In this chapter, I tested how matching and initial ability in performing a clerical task influence coordination in a minimum-effort game with a real task. Specifically, employees were divided into HP and LP based on the relative performance during the learning phase, and homogeneous firms were formed. Information about both individual performance and matching procedure received by the employees was varied. Before summarizing our main results, it should be noticed that, given the technical problem affecting the majority of the data, the present study has the character of a pilot study and further experiments are needed to understand if the results are robust.

In line with our first hypothesis, the data seem to suggest no differences in the coordination level between firms composed by HP and LP when the employees were not informed about performance and matching (*NoInfo* treatment). While in real firms initial ability is likely to affect subsequent performance, I constructed a testbed in which this potential confound is stripped away, since we wanted to isolate the effect of group pride.

The present results provide evidence, although not definitive, that group pride alone can enhance coordination in experimental firms. Coordination failure is common in the present experiment; notwithstanding, important behavioral differences between HP*Info* firms and all other firms (i.e., LP*Info* and all *NoInfo* frims) are present. Indeed, a speedy unraveling toward the inefficient equilibrium (i.e., minimum equal to zero) is observed only in two out of eight HP*Info* firms, while the same behavior was observed in the vast majority of the remaining experimental firms (15 out of 19). In line with psychological and organizational evidence on self-esteem and group identity, the priming of individual and group traits seems to affect significantly beliefs and trust within the firm. It is worth to notice that, in-group bias is observed even though, and unlike most previous experiments on SIT, no direct comparison and competition among different groups is present. In addition, LP*Info* showed grater negative reaction to initial effort and minimum as compared to HP*Info* that, starting off from the same coordination level, were able to climb the strategy space and achieved successful coordination: this effect is in line with the one documented by Pierce et al. (1993).

The experiment also provides a comparison between real and chosen effort, an issue that stems from recent concerns about external validity of lab data. In the present setting and for the specific game and task, chosen effort is a reliable proxy for real effort.

The aforementioned results might provide some interesting insights and prescriptions for organizations; the priming of positive characteristics can indeed result in gains for organizations. Group pride may apply to a wide range of groups, not only elite groups properly defined; indeed, there is evidence that ability on which grouping, and hence group pride, is based do not need to be relevant per-se for achieving successful coordination (that is, even low ability employees could have coordinated on the efficient strategy).

The present framework can be further extended in several ways; in the following, I lay out three possible follow-up studies building on the present testbed. First, it would be interesting to investigate whether the beneficial effect of group pride on coordination survive when segregation is based on a task different from the task used in the stage game. This proposed extension serves two main purposes: (a) better understand which tasks are well suited for inducing group pride and which are not; (b) control for the possibility that differences in perceived strategic uncertainty between HP and LP present. Even though all employees, both HP and LP, were able to correctly solve the task at hand (that is, could correctly count four tables in a round), employees may have developed different beliefs about the ability of each and every member of the firm to correctly solve the four tables. Therefore, LP may have thought that the Pareto-efficient equilibrium could not be attained because of ability, rather than strategic, reasons. Grouping based on a task not correlated with the counting task can account for the aforementioned possible confound (see Chapter 5).

The second proposed extension is aimed to disentangle the effect of the two components of group pride: self-esteem and group trust. In fact, in the *Info* treatment of the present study, I provide information about both others' performance (which is responsible for group trust) and self-performance (which is responsible for individual self-esteem). Although in elite groups both levels —individual and group— are commonly present, it would be interesting to sort out the relative contribution of the two elements. That is, to what extent people value other's self-esteem? Is self-esteem more important than trust? The answer to the aforementioned questions is of interest from both an academic and practitioners perspective. Notably, little research has been done on the interaction between these two levels; the literature on SIT and status focused almost exclusively on the group level, while studies on self-esteem rarely considered the interaction between individual and others' self-esteem from a strategic perspective. Moreover, a better understanding of the interaction between self-esteem and group pride may provide a useful guidance to practitioners that have to decide whether to invest more on the individual or the group level. The effect of self-esteem could be easily accounted for by adding a new treatment —*Self-esteem*— equal to the *NoInfo* treatment the only difference being that, at the end of the learning phase, employees are informed about their performance, and their performance only. Specifically, employees will know whether they ranked among the top or the bottom eight during the learning phase but no information about the matching protocol will be provided. I plan to keep the same matching procedure (i.e., homogeneous firms) not to change more than one variable at a time and to avoid the potential confounding coming from the interaction among employees with different levels of self-esteem.

Finally, the present set-up can also be used to address an important organizational design issue: group composition. Whereas several characteristics of group composition, such as age, gender, and race, have been analyzed, talent disparity have being mainly overlooked (see Franck and Nüesch, 2008; Hamilton et al., 2004 for some notable exceptions). However, heterogeneity in the distribution of talent and performance is universally observed in actual organizations, and it is a priori unclear how talent heterogeneity affects group productivity. The presence of more talented employees in one group may affect performance in several ways; first of all, best employees can teach to less skilled employees, hence producing learning spillovers; second, the presence of talented employees can establish a social norms of high productivity. Nevertheless, the presence of less talented employees in a group may demotivate best employees that can adapt to a lower level of productivity. Optimal talent distribution is also affected by the degree of task interdependence, and, in theory, homogenous groups have to be preferred over heterogeneous groups in case of a weak-link technology. In the proposed follow-up experiment, I intend to test whether homogenous groups are more efficient, as compared to heterogeneous groups, by running an additional treatment in which each firm is composed by HP and LP in equal proportion. Similarly to the *Info* treatment, the position in the ranking (i.e., HP or LP) and the matching protocol will be common knowledge among our experimental employees. As in the present experiment, initial ability does not affect coordination per se, and we can concentrate on a single aspect of talent disparity —social norms— by eliminating the potential confounding of learning spillovers. Specifically, we want to test whether a high rather than a low productivity norm prevail when HP and LP interact with each other: will HP be able to act as leaders hence sustaining coordination? Three previous experiments addressed the issue of group composition in the weak-link game. First, Brandts, et al. (2007) implemented an experiment with asymmetric costs of effort to test whether more talented employees (i.e., employees with lower costs of exerting effort) are more likely to become leaders. It is important to notice that in this set-up the Pareto-dominant action is less risky for talent employees, while in our testbed no such difference would be present. Second, Brandts et al. (2009) tested the effect of learning spillovers in heterogeneous group by allowing high ability employees to help less able ones. As already mentioned, our proposed experiment tackle the problem of talent disparity in situation where learning spillovers are not possible, hence concentrating on the social norm aspect of the problem. Finally, Engelmann and Normann (2007) showed that the subject-pool composition affects the level of coordination in the minimum-effort game, and, in particular, the minimum effort increased as the share of Danish subjects in the group increased. In this case, the group composition was not known by subjects in advance meaning that there is a correlation between being Danish and coordination. From and organizational point of view, this findings should suggest to change the workforce composition trying to hire as many talented employees as possible, while we are mainly interested on how to enhance coordination given the existing workforce.

Appendix 4: Experimental Instructions

Instructions (Part 1) [Real effort][16]

General information: The purpose of this experiment is to study how people make organizational decisions. 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 privately.

You will be able to earn additional 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 200 ECUs. Upon completion of the experiment the amount that you earned will be paid to you in cash. Payments are private; no other participant will be told the amount you earned.

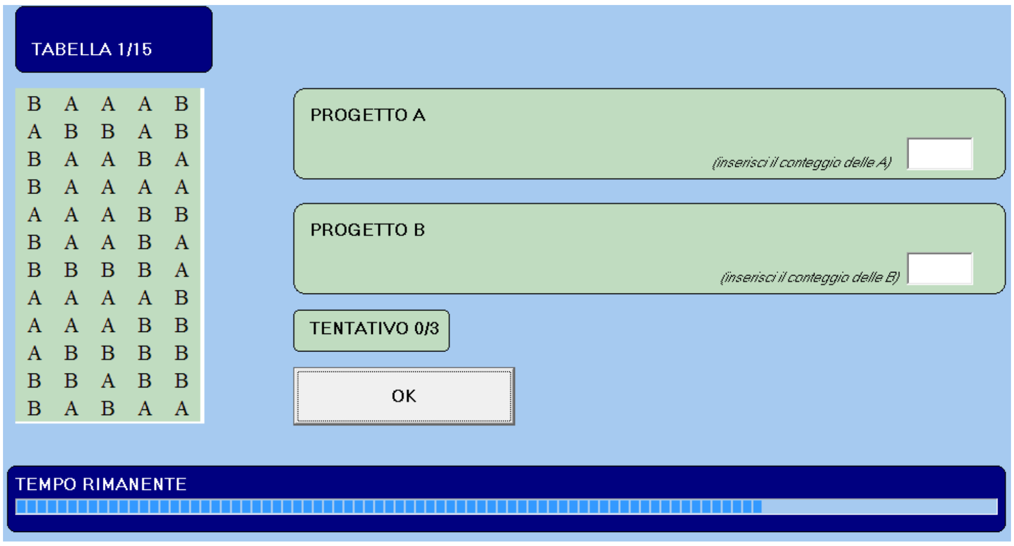
Parts and rounds: This experiment will have 2 parts. Part 2 has 10 rounds. In part 2, you will be matched with 3 other participants in this room. Part 1 will last 1 round and you will not be matched with anyone. In the first part you have the chance to familiarize with the task you will be assigned in the second part. The following instructions are about part 1 only.

[16] Translation from Italian; the original instructions are available upon request to the author.

Description of the task : You can think of yourself as an employee of a firm that is assigned the following task:

- 15 tables containing A and B will be displayed on the screen one after the other; your task consists of counting the number of either As or Bs contained in each table;
- for each table, you may choose to work either for project A (which implies that you will be counting As) or for project B (which implies that you will be counting Bs);
- you will have 3 minutes to complete as many tables as you can; this time will not be extended.

Practical implementation: a screen like the one shown below will be displayed on your computer.



A table containing A and B will appear on the left side of your screen; you can start completing your task as soon as the table is displayed on the screen. If you decide to contribute a table to project A, you have to count the number of A and report this number in the blank space labeled "Number of A"; if you decide to contribute to the project B, you have to count the number of B and report this number in the blank space labeled "Number of B". You have to confirm your number by clicking on "OK". If the number (of either A or B) you entered is correct and you clicked on "OK", the next table will be displayed. In case the number you entered is not correct, a pop-up window will open right above the two green fields and will inform you that the number is wrong; you will have 3 trials for each table, after the third trial, no matter if the number is correct or not, the next table will be automatically displayed.

For each table, you can independently decide whether to contribute to either the project A or B. That is, if you start working for project B you can decide to switch to project A for next table. You can switch from one project to the other as many times as you want. However, note that, for each table, you can contribute to ONLY ONE of the two projects.

In part 1, you will have 3 minutes of time to complete at most 15 tables. The time left is reported on the bottom part of the screen. The number of the tables and trials are also reported on the screen.

<u>Earnings:</u> you will be paid a fixed amount of 3 Euros for this part, no matter how many tables (and for which project) you completed (them). We ask you to try to complete as many tables as you can

<u>Information that you will receive</u>: you will be informed about the number of tables you successfully completed for both project A and B.

Please remember that if any type of communication is detected during the experiment, the session will be immediately concluded and nobody will be paid. If you have questions, please raise your hand and one of the experimenter will come to your desk to answer it privately.

Instructions (Part 2) [Real effort]

Info treatment: *italics and marked in yellow*

Parts, rounds, and groups: Part 2 has 10 rounds. In part 2, you will be matched with 3 other participants in this room; *the matching will depend on the performance in part 1. Specifically, participants in this room will be divided in two categories:*

1. *Best Performers: the 8 employees who completed the highest number of tables in part 1;*
2. *Worst Performers: the 8 employees who completed the smallest number of tables in part 1.*

Ties will be resolved at random. You will be matched with 3 other employees from the same category. For instance, if you are among the Best Performer (Worst Performer) the other 3 members of your group will be Best Performers (Worst Performers) as well.

Before part 2 will start, you will receive an envelope containing information about your performance in part 1 (Best Performer or Worst Performer) and the type of group your will be working for (Best Performers Firms or Worst Performers Firms)

Once determined, the composition of a firm will be the same for all rounds. The following instructions pertain to part 2.

Description and implementation of the task : You can think of yourself and the other members of your group as employees of a firm.

Task and implementation are the same as in part 1; the only difference is that in each round four tables at most can be completed. The time allotted for each round is 3 minutes, as in part 1.

Payoffs: Your earnings in each round depend on how many tables you and the other employees of the firm will successfully complete: a table will be counted only in case the number of either A or B which have been counted is correct. Moreover, your payoff will depend on how many tables you and the other members of your firm will decide to solve for the project A and B. Please, recall that completing a table for project A means that you correctly counted A, and completing a table for project B means that you correctly counted B.

For each successfully completed table you contribute to project B, you will get 50 ECUs no matter what other employees of the firm do. The payment for project A will be 80 ECUs multiplied by the MINIMUM number of tables a member of your firm (you included) has contributed to project A. That is, in order to determine the minimum contribution to project A (*min(A)*), you have to consider the number of tables solved for project A by each of the employees of the firm (you included) and pick the MINIMUM number. For example, if the number of completed tables for project A is 4, 3, 2, and 0, the smallest number is 0 and hence *min(A)=0*.

Formally, your earnings are determined by the following payoff function:

$$\pi = 80*[min(A)] + 50*[B] \qquad \text{(a)}$$

Please note that incorrectly or not complete tables will not be paid.

To help you understand better how your earnings are determined, below please see we report the payoff table for a particular case: the following payoff table is, in fact, **ONLY FOR THE CASE IN WHICH YOU COMPLETE ALL OF THE FOUR TABLES** IN THE ROUND. PLEASE, REMEMBER THAT THE FOLLOWING TABLE DOES NOT EXHAUST ALL POSSIBLE CASES, which can be derived by using the payoff function (a). Cases where not all 4 tables are completed will be discussed later on. All numbers in the non-shaded cells represent ECUs.

		Minimum contribution to project A: *min(A)* (number of tables)				
Your contribution to		**4**	**3**	**2**	**1**	**0**
project A	**4** (0)	320	240	160	80	0
(in brackets the number	**3** (1)		290	210	130	50
of tables you	**2** (2)			260	180	100
contributed to project	**1** (3)				230	150
B)	**0** (4)					200

For easy reference we also have provided you with a paper copy of the earnings table and function on your desk.

Focus initially on the first row (shadowed in grey) in the earning table; in that row, the possible MINIMUM contributions to project A are listed. *min(A)* ranging from 4 [the case where all employees of the firm contributed four tables to project A] to 0 [the case one or more members contributed zero tables to project A]. For example, if the number of tables completed for project A are: 4, 3, 2, and 2, then the minimum is 2; hence, you have to check the column identified by "**2**" in the first row for your potential earnings.

Focus next on the first column (shadowed in grey) in the above payoff table; in that column, the number of tables YOU can contribute to project A is reported in bold. In brackets, the number of tables you can contribute to project B is reported. For instance, you had successfully completed 4 tables for project A (and, as a consequence, 0 for project B): you would have had to check the row identified by "**4** (0)" in the column for your potential earnings.

Notice that the sum of tables contributed to project A and B is always four; that is because **ONLY** THE SPECIAL CASE OF FOUR TABLES SUCCESSFULLY COMPLETED IS REPORTED IN THE PAYOFF TABLE. You can determine your earnings in a round by crossing the relevant rows and columns. Consider the previous examples in which *min(A)*=2 and your contribution to project A was 4; your earnings are then determined at the intersection of row "**4** (0)" and column "**2**" yielding a payoff of 160. Please, notice that this payoff can be computed also by using the payoff function (a):

$$\pi = 80*[min(A)] + 50*[B]$$

$$= 80*[2] + 50*[0] = 160$$

Consider the same example, but suppose you contributed 2 to project A, 1 to project B, you failed to complete one table, and the minimum is still 2. You cannot determine your earnings by using the payoff table, since you did not completed all of the four tables; hence, you have to use the payoff function (a):

$$\pi = 80*[min(A)] + 50*[B]$$

$$= 80*[2] + 50*[1] = 160 + 50 = 210$$

<u>Information that you will receive</u>: Each round, after all employees have finished their task or the time is over, the computer will display a screen like the one shown below. In the first column the number of the round is reported. The actual number of tables you completed for project B is reported in the second column, and the number of tables you completed for the project A is displayed on the third column. In the fourth column the number of tables each member of the group (you included) contributed to project A is displayed; in curly brackets *min(A)* is reported. Your payoff for each round, and your cumulated payoff is shown in the last two columns. When you are ready to start a new round, please click on "Go to next round" and a new table will be displayed.

ROUND	INDIVIDUAL	FIRM	GROUP	PAY ROUND	PAY TOT
1	0	0	0,0,0,0 {0}	0	0

At the end of the 10 rounds , information about that round will be displayed and a message on the screen will alert you that the experiment is over.

<u>Payment</u>: At the end of the experiment, you will be paid, in cash, the sum of the payoffs that you will have earned in all the rounds of the experiment. The conversion rate is one Euro for 200 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.

Chapter 5

Group pride and reduction of strategic uncertainty

5.1. Introduction

In Study II, I tested how the motivation derived from being part of an elite group, or "group pride", affects coordination in a real-effort experiment. For that purpose I designed and implemented a real-effort experiment in which subjects had to count the number of letters contained in a sequence of tables and through with which I attempted to capture the prototypical features of the minimum-effort game . Groups of four were formed based on performance in the first part of the experiment, and the feedback about relative performance in the clerical task was varied to induce different levels of group pride. When groups composed by high performers received information about their type and matching with other high performers, higher levels of effort and efficient coordination were achieved.

In this chapter, I follow up on Study II and test whether the beneficial effects of group pride on coordination survive when grouping and coordination are based on different tasks. Unlike Study II, high performers (HP, hereafter) are divided from low performers (LP, hereafter) according to performance on either a math or general knowledge quiz, while the task used in the coordination phase was the same as the one used in the previous chapter (counting numbers in a sequence of tables).

This extension of the testbed is important for at least two reasons. First, exploration of the robustness of group pride further contributes to economic studies on Social Identity Theory (SIT, hereafter) that found the beneficial effects of group identity in a variety of games to critically depend on the manipulation used. For instance, Chen and Chen (2009) found that dividing subjects according to their preferences for painters does not affect behavior, while enhancing the salience of group identity (subjects with similar preferences for painters had to help each others to solve a quiz) enhances coordination among in-group members. Similarly, Charness et al. 2007 found no effect of group identity when subjects were simply classified in row and column players, but the introduction of observers, either in- or out-group, was highly effective. Overall, economic experiments have shown so far that the magnitude of the in-group bias depends on the salience of the categorization procedure. Along the same line, I test to what extent pride generated in one context can be transferred to another context; indeed, group pride may apply to all those groups that have some (real or imaginary) reasons to be proud of themselves even if these reasons are only loosely related to the task at hand. For instance, think of a group of employees hired on the base of their college achievements for a job that does not require any particular ability and knowledge (i.e., performance does not depend on the education level): under these circumstances, would priming the scholastic achievements affect performance on the workplace?

Second, in Study II I hypothesized that the strategic uncertainty undermining coordination can be reduced in elite groups through two psychological components: self-esteem and group trust. However, previous results might have been confounded, at least partially, by strategic uncertainty reduction. Even though I provided evidence that all subjects, both HP and LP, were able to correctly solve the task at hand (that is, could correctly count four tables in a round), they may have developed different beliefs about the ability of each and every member of the firm to correctly solve the four tables. Therefore, LP may have thought that the Pareto-efficient equilibrium could not be attained because of ability, rather than strategic, reasons. As in standard minimum-(chosen) effort games, actions and strategy space are common knowledge; however, the introduction of a real-effort task introduces in the game additional uncertainty: feasibility of action and strategies. To control for the strategic uncertainty reduction hypothesis, the task used to divide experimental subjects into HP and LP was varied; specifically, in one treatment a task similar to counting numbers was used (i.e., solving additions), while in the other treatment the task was completely uncorrelated with counting (i.e., answering general knowledge questions).

Thus, the strategic uncertainty reduction argument can be measured by comparing the outcomes of the two treatments: *Math-I* in which subjects were grouped according to their performance in the addition task and *GK-I* in which subjects were grouped according to their performance in the general knowledge quiz. In both treatments, subjects received information about their type and the matching protocol. Since uncertainty reduction is present in the treatment where a correlated task is used but not in the uncorrelated task treatment (i.e., no information about abilities in solving the counting task is conveyed by general knowledge questions), if strategic uncertainty reduction is responsible for results in Study II, differences between HP and LP should be observed in *Math-I* but not in *GK-I*. To control for any spurious correlation, a treatment without information was run for the mathematical task (*Math-NI*).

The main findings of the present experiment (Study III) can be summarized as follows. First, correlation between performance in the grouping and the coordination task was as conjectured; namely, there is a positive (actual and perceived) correlation between adding up numbers and solving tables, and a negative correlation between answering general knowledge questions and solving tables. Second, performance in the math task can affect subsequent behavior when no information is provided; that is, HP coordinated on higher effort levels as compared to LP in *Math-NI*. Surprisingly enough, information about type and matching had, if anything, a negative impact on HP in the math treatment. Third, in line with results in Study II, and contrary to evidence in *Math-I*, the data show a positive and significant effect over time of group pride on high performers in our GK sessions. Thus, reduction of strategic uncertainty alone does not play a role in enhancing coordination.

This chapter is organized as follows: Section 2 describes the stage game and Section 3 details the experimental design and the behavioral hypotheses. Section 4 reports the results for each of the three parts of the experiment and Section 5 concludes.

5.2. The minimum-(real)effort game

The stage game and the clerical task used in this chapter are exactly the same as in Study II; thus, I report only the main characteristics of the game at hand (for a detailed discussion, see §4.3). Subjects (employees, hereafter) are matched in groups (firms, hereafter) of fours and the composition of each firm remains the same throughout the experiment. In each period, the employees of each firm have to simultaneously choose how to allocate four units of time between two different projects: A and B. While project B pays a fixed amount for each unit devoted to the project, project A yields – possibly - a higher amount based on the minimum contribution of time to project A within the firm. Formally, employees are paid according to the following earnings function:

$$\pi_i = G\left[min_{j\in\{1,\dots,4\}}\,(a_j)\right] + I(b_i) \qquad (1)$$

where, π_i are the earnings for employee $i \in \{1,2,3,4\}$, $min_{j\in\{1,\dots,4\}}\,(a_j)$ is the minimum contribution to project A by a member of the firm, b_i are units of work devoted to project B by employee i, and G and I are the bonus for the minimum contribution to A and the individual contribution to B, respectively. In each period, exactly four units of time have to be divided between the two projects. Table 5.1 reports the earnings table for $G = 80$ and $I = 50$.

As for Study I and II, a real effort task was used; specifically, I let experimental employees count the number of either As or Bs contained in a table. Correctly counting the As contributes to project A, while counting the Bs contributes to project B. Four tables are presented in sequence in each period, and for each table employees can choose between project A and B where project A brings about a payoff pay according to the minimum contribution to the project within the group while project B brings about an individual payment according to equation (1). If all four tables are correctly solved within a period, Table 5.1. shows the earnings for the current period.

TABLE 5.1: Earnings table

		Minimum contribution to project A: *min(A)* (number of tables)				
		4	3	2	1	0
Your contribution to project A (number of tables you contributed to project B)	4 (0)	320	240	160	80	0
	3 (1)		290	210	130	50
	2 (2)			260	180	100
	1 (3)				230	150
	0 (4)					200

5.2.1. Experimental design

This experiment is meant to address two main questions: first, does reduction of strategic uncertainty, as induced by the grouping mechanisms, enhance coordination?; second, is the beneficial effect of group-pride observed in Study II a function of the particular grouping mechanism?

The experimental design was varied along two dimensions: feedback and task used for Part 1. The main focus is on how different tasks used to group employees affect group pride and coordination; therefore, in Part 1 subjects were asked to either solve mathematical additions (*Math* treatment) or answer general knowledge questions (*GK* treatment). Following the protocol used in Study II, the feedback provided to employees about type and matching were also varied (*I* vs. *NI*). Table 5.2 summarizes the experimental design; due to budget and time constraints I decided not to run a full 2 [*I* vs. *NI*] x 2 [*Math* vs. *GK*] between-subjects design and opted for not having a *NI* control for the *GK* treatments.

TABLE 5.2: Experimental design

	Math-NI	*Math-I*	*GK-I*
Effort	Real	Real	Real
Part 1 (grouping phase)	Additions	Additions	General Knowledge
Matching in Part 2	Based on performance in Part 1	Based on performance in Part 1	Based on performance in Part 1
Information about the matching protocol	No	Yes	Yes
Total # of subjects	48	48	48
# of sessions	3	3	3

NOTE: Sixteen subjects participated in all sessions.

Informed by our previous study, we decided to implement two new tasks in Part. This new design help accounting for a potential confound present in the previous design: strategic uncertainty reduction Indeed, by using the same task for both grouping and coordination, differences in observed coordination between HP and LP could be driven by both group pride and reduction of strategic uncertainty: while being grouped with HP might have generated positive expectations about everyone in the firm being able to complete all four tables, LP might have thought that the other employees within their firm were not able to successfully complete all four tables, hence ruling out some equilibria from the strategy space. Moreover, I aimed to test the robustness of group pride across different domains in order to provide some evidence on priming strategies that foster coordination.

To disentangle group pride from reduction of strategic uncertainty I varied the correlation between ability needed for performing the task in Part 1 and in the stage game. Both adding up numbers and counting letters are related to mathematical skills, performance in the two tasks is thus expected to be, and to be perceived as, correlated. In contrast, I conjecture that there is no correlation —actual or perceived— between performance in the general-knowledge quiz and the counting task.

> *HYPOTHESIS 1: There is a positive correlation, actual or perceived, between adding up numbers and counting letters while there is no correlation, actual or perceived, between answering general knowledge questions and counting letters.*

In line with the evidence from Study II, I do not expect to find any correlation between coordination and type (HP or LP) when no information is provided to the experimental employees (i.e., *Math-NI* treatment). In addition, I conjecture that group pride or, possibly, reduction of strategic uncertainty lead firms composed by HP to coordinate on higher levels as compared to LP when information is given in the *Math* treatment.

> *HYPOTHESIS 2: (a) effort and minimum over time are indistinguishable between HP and LP in Math-NI treatment; (b) effort and minimum over time are higher for HP as compared to LP in Math-I* treatment.

The *GK*-I treatment is meant to disentangle the relative contribution of group pride and strategic uncertainty reduction. Thus, in case no correlation between performance in the GK quiz and counting is detected (Hypothesis 1 is satisfied), any difference in behavior between HP and LP in *GK-I* cannot be attributed to strategic uncertainty reduction. In contrast, if strategic uncertainty reduction affects coordination, we should observe little, if no, difference between HP and LP firms in *GK-I*.

HYPOTHESIS 3: I conjecture that the group pride looms larger than the strategic uncertainty reduction argument; thus, both effort and minimum over time are higher for HP as compared to LP in GK-I treatment.

5.2.2. Implementation

A total of 144 subjects (56 percent male, average age 20.36), equally distributed among the three treatments, participated in the experiment which was conducted at the Computable and Experimental Economics Lab (CEEL) at the University of Trento. All but one subject were students from the University of Trento and were recruited through emails, ads, and posters; each subject was allowed to participate in at most one session and nobody participated in a similar experiment before.[1] Participants were mostly enrolled in economics (122 out of 144) and the great majority was either freshmen (75 out of 144) or second year college students (48 out of 144).

Upon arrival, subjects were randomly assigned to a cubicle and no form of communication was allowed from that moment on. The experiment was divided in four parts and a paper copy of the relevant instructions (see Appendix) was distributed at the

[1] Subjects that previously participated in a minimum or median effort experiment, with either real or chosen tasks, were excluded from the subject pool.

beginning of each part; the experimenter read the instructions out loud to assure common knowledge. In Part 3, and after reading the instructions, subjects were asked to answer a short questionnaire to ensure correct understanding of the earnings table; in case of incorrect answers, subjects were asked to revise their answers.

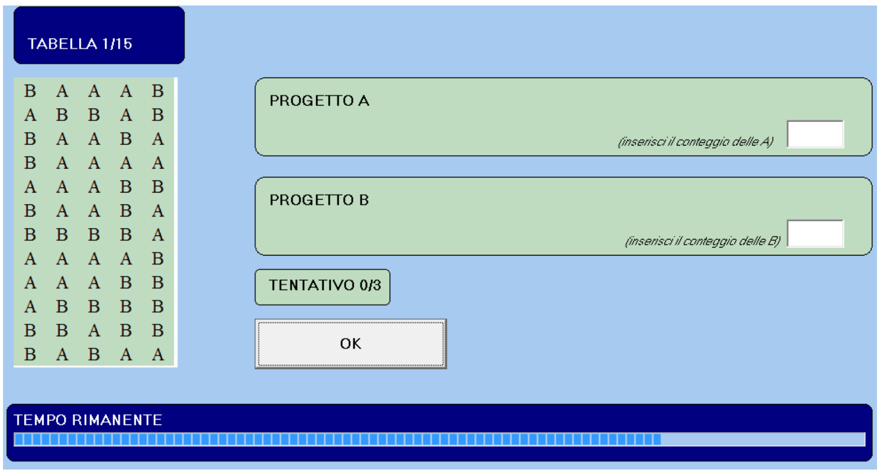
All treatments were divided into four parts. In Part 1 (or *grouping phase*, subjects were asked to either perform additions of pairs of two-digit numbers or answer general knowledge questions. In *Math* treatments, subjects were allotted 2 minutes time to solve as many additions as they could; a decision-sheet containing 44 additions was placed on each desk and covered with a blank paper. The experimenter informed subjects when they could turn over the decision-sheet and start working; when the time allotted to the task was over, subjects were asked to stop working without delay and the decision-sheet was collected by the experimenter. The procedure was the same for the *GK* treatment, the only difference being that subjects were given 3 minutes to answer as many general knowledge questions as they could.[2] The general knowledge questionnaire was composed by 36 questions equally divided in 6 broad categories: sport, music, art/cinema, geography, curiosities, and history/politics. The order in which the questions were presented was determined at random, with the additional constraint of two questions from the same category not following each other immediately. To minimize information leakage between sessions, the order in which the questions were presented was changed; subjects were allowed to solve the additions or answer the questions in the order they liked. Subjects were paid a fixed amount of 2 Euros for Part 1 no matter how well they solved the assigned task. [3]

[2] The time interval was longer in *GK* as compared to *Math* since answering questions is likely to take a longer time than adding up numbers and we wanted to maintain a reasonable variance in the performance data.

[3] We did not pay subjects on a performance base in Part 1 since we did not want to have HP to start off from a higher wealth level as compared to LP. Therefore, our classification may also depend on

Part 2 and 3 were the same as the learning and coordination phases in Study II. In Part 2 (or *learning phase*) fifteen tables containing As and Bs (see Figure 5.1) were displayed on the screen one after the other; for each table, subjects could choose to work for either project A —which implied counting the number of As— or project B —which implied counting the number of Bs. Subjects had to enter the number of either As or Bs in the corresponding cell and confirm their choice by clicking "OK"; a new table was shown in case the result of the counting activity was correct. Subjects had three trials for each table, after three incorrect trials the next table was automatically displayed on the screen. A bar on the bottom part of the screen displayed the time left and no number could be entered after the time was expired. At the end of Part 2, subjects received information about the number of tables correctly completed for each project.

FIGURE 5.1: Screenshot



In Part 3 (or *coordination phase*), subjects were matched in groups of four and played 9 periods of our stage game; the group composition depended on the performance in Part 1. Specifically, subjects were ranked according to their performance in Part 1 (i.e., the number of correct answers); in each session, subjects ranked 1 through 8 and 9 through 16

the subject's desire to please the experimenter ; however, this effect , if it exists, is likely to be the same across the two treatments.

were referred to as “high performers” and “low performers”, respectively.[4] Each group was composed only by subjects of the same type (that is, only HP or LP).[5] We varied the information provided to subjects about their position in the ranking and the group composition; in the no information treatments (NI) subjects did not receive any feedback about ranking and matching, while in the information treatment (*I*) subjects received an envelope in which their type and the composition of the group was reported. The implementation of Part 3 was exactly the same of Part 2 the only difference being that the maximum number of tables displayed in each period was four, rather than fifteen. At the end of each period, information about individual contribution to the two projects, distribution of contributions to project A within the group, minimum, and earnings (both cumulated and for the current period) were displayed on the screen. Subjects were paid according to the earnings table reported in Table 5.1.

In Part 4, a measure of risk aversion was collected by implementing the instrument used by Holt and Laury (2002). In order to avoid having subjects to switch from one alternative to the other multiple times, we asked them to simply cross the first line in which they preferred the alternative with the higher outcome variance over the other alternative. At the end of the experiment, relevant socio-demographic information was collected and subjects were asked to indicate, on a scale from 1 to 10, the perceived similarity between counting letters in a table and solving additions, and that between counting letters and answering general knowledge questions.

The average session lasted about 90 minutes: all values were expressed in Experimental Currency Units (ECUs) and were converted at the end of the experiment at

[4] Ties were broken at random.
[5] Employees ranked 1st, 3rd, 6th, and 8th were grouped together in firm A; 2nd, 4th, 5th,and 7th in firm B; 9th, 11th, 14th, and 16th in firm C; 10th, 12th, 13th, and 15th in firm D. This protocol was adopted to ensure that ability was evenly distributed across firms of the same type.

the rate of 1 Euro for 200 ECUs. Subjects knew the conversion rate in advance and were paid privately their earnings for each part. Both Part 1 and 2 paid 2 Euros, while earnings for Part 4 ranged from 0.10 to 3.85 Euros.

5.3. Results

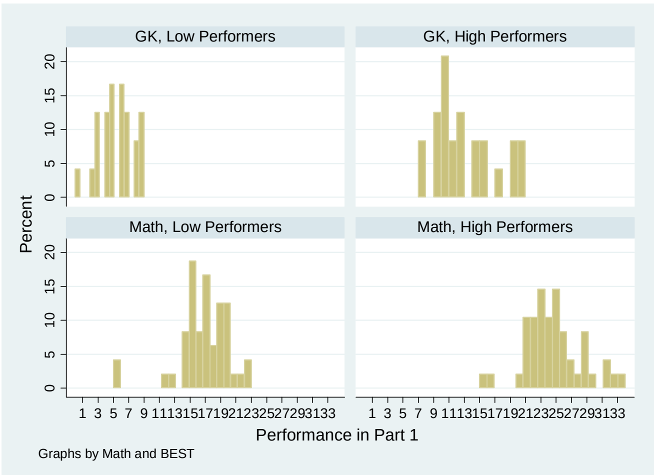
It might be useful to recall that observations at the employee level in the coordination phase show correlation across both periods and employees of the same firm, with the only exception being the first period; therefore, when testing the hypotheses in periods other than the first (and if not otherwise specified), the independent unit of observation is the mean value of the variable for a single firm. Thus, there are a total of 36 independent observation equally distributed among treatments and types. All tests are two-sided. In the following, I use "effort" and "minimum" to refer to individual and minimum contribution to project A, respectively. The term "mistake" refers to the number of tables not successfully completed for one of the two projects within a period (the sum of the individual contribution to project A and B can indeed be smaller than four in case one or more mistakes in the counting task were made).

This section is organized as follows. First, I discuss behavior in the grouping and learning phases. Second, I analyze the number of mistakes made by our experimental employees. Finally, I provide evidence on behavior in our minimum-(real) effort game. Since we are interested in the effect of group pride on both initial beliefs and subsequent coordination dynamics, initial and subsequent periods are discussed separately.

5.3.1. Part 1: Grouping phase

Figure 5.2 shows the distribution of correct answers in the grouping phase divided by task (Math and GK) and type (HP and LP); an eyeball examination suggests there is considerable heterogeneity in the distribution of performance for both tasks. The number of correctly solved additions in Math sessions ranged from 5 to 34, and the average was 20.71 (Std. Dev. 5.43). The average number of correctly solved additions was 24.69 for HP and 16.73 for LP. The number of correct answers in the GK sessions ranged from 0 to 21, and the average was 9.04 (Std. Dev. 4.85). The average number of correct answers was 12.36 for HP and 5.46 for LP. As expected, the number of answers was lower in GK as compared to Math sessions, but the variance was large enough to observe only few ties. No information leakage was observed in GK sessions; indeed, difference in the number of answered questions across sessions is not statistically significant according to a K-sample test on the equality of medians (p = 0.706) and to a Kruskal-Wallis rank test corrected for ties (p = 0.138).

FIGURE 5.2: Distribution of correct answers in Part 1 by task (Math vs. GK) and type



NOTE: HP and LP were divided according to a median split. Since some across-session variance was observed, it might be the case that employees with the same performance (e.g., same number of solved addition or GK questions) were classified as HP in one session and as LP in another session.

It is interesting to note that males performed better than females in both tasks used in the grouping phase. Table 5.3 summarizes the number of males and females classified as HP and LP. Several reasons can be at the basis of the observed difference: first, males are often thought to be, on average, better at solving mathematical quizzes (Shih et al., 1999); second, the situation might have been perceived as a sort of competition and, as suggested by the work of Niederle and Vesterlund (2007), females tend to shy away from competition more than males. However, the analysis of the average number of solved additions does not reveal any statistical difference between males and females for either HP (two-sided Wilcoxon rank-sum test, p = 0.138) or LP (two-sided Wilcoxon rank-sum test, p = 0.318). Gender difference was more marked in the GK; while the difference was not significant for HP (two-sided Wilcoxon rank-sum test, p = 0.178) it was statistically significant at the five percent level for LP (two-sided Wilcoxon rank-sum test, p = 0.049).

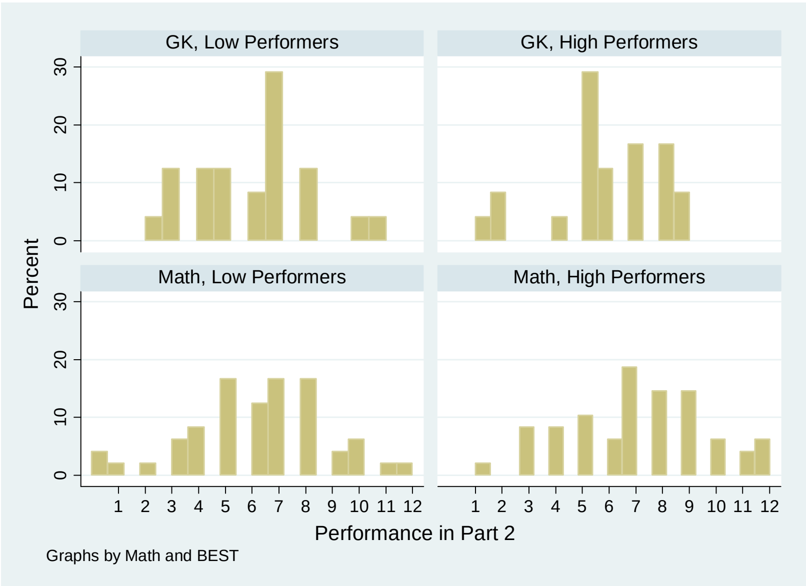
TABLE 5.3: Gender differences between HP and LP for both Math and GK tasks

	Math		GK	
	# of male	# of female	# of male	# of female
HP	*31*	*17*	*16*	*8*
LP	*23*	*25*	*10*	*14*

5.3.2. Part 2: Learning phase

Figure 5.3 shows the distribution of correctly solved tables during the learning phase divided by task in Part 1 (Math and GK) and type (HP and LP). The number of correctly solved tables ranged from 0 to 12 and the average was 6.40. The average number of tables was higher in Math (6.64) as compared to GK (5.92) sessions, albeit the difference was only weakly significant according to a Wilcoxon rank-sum test (p = 0.095, two sided). The aforementioned difference was mostly due to HP that solved on average 7.13 tables in Math sessions and 5.75 tables in GK sessions (p = 0.067). These data seem to suggest that adding up numbers might have served as a sort of training for the subsequent task. While performance in the learning phase was higher for HP as compared to LP in Math sessions, the same was not true for GK sessions in which the reverse pattern was observed.

FIGURE 5.3: Distribution of solved tables in Part 2 divided by task in Part 1 (Math and GK) and type (HP and LP)



As predicted, correlation between adding up numbers and counting letters was positive, although small (corr. 0.264). On the contrary, negative correlation between general knowledge and counting letters was found (corr. - 0.052). Thus, scoring high in the GK questionnaire has no, if not negative, effect on performance in the clerical task.

TABLE 5.4: average (std.dev.) perceived similarity between Math, GK, and counting tables

	Overall	Math-NI	Math-I	GK-I
Math-	6.01	5.10	5.85	7.06
Counting	(2.59)	(2.61)	(2.72)	(2.04)
GK-	2.78	2.71	2.81	2.81
Counting	(2.30)	(2.14)	(2.33)	(2.47)

NOTE: similarity is based on a scale from 1 to 10, where 1 is the lowest and 10 the highest level of similarity

In order to control for the potential interaction between performance in the grouping phase and strategic uncertainty in the coordination phase, it is needed to consider both actual and perceived correlation between the tasks used in Part 1 and 3 of the experiment. At the end of the experiment, employees were asked to state the perceived similarity, on a scale from 1 to 10, where 1 is the smallest and 10 the greatest degree of similarity, between summing up numbers and counting letters and between GK questions and counting letters. As shown in Table 5.4, not only the actual but also the perceived correlation was higher for Math as compared to GK. While perceived similarity between GK and counting was similar across treatments, an interesting difference in perceived similarity between Math and counting deserves to be mentioned: indeed, the similarity was perceived to be stronger by those employees which did not undertake the adding up task.

RESULT 1: There is a positive correlation, both actual and perceived, between adding up numbers and counting letters while there is a negative, albeit small, correlation, both actual and perceived, between answering general knowledge questions and counting letters.

5.3.3. Part 3: Mistakes

I now turn to the analysis of the coordination phase. Before discussing initial coordination and subsequent dynamics, I focus on the number of mistakes (the number of mistakes is defined as the number of tables not completed or not completed correctly) made by employees in the 9 periods of the stage game. The cause of mistakes can be twofold: first, inability to perform the task within the assigned time limits; second, unwillingness to perform the assigned task since the subjective cost of solving a table is higher than the bonus (either *G* or *I*) for completing the table (for a more detailed discussion of the equilibria structure, see § 3.2.1).

In line with evidence from Study II, only a negligible number of mistakes was observed; 90 out of 5184 tables (1.74 percent) were not successfully completed, and mistakes were evenly distributed across periods. Two or more mistakes were made within a period by an employee in only 16 cases (out of 1296). At the employee level, 112 out of 144 employees successfully completed all the tables in each of the 9 periods of the game; of the remaining 32 employees, only 10 made more than two mistakes in total.

HP tend to perform better as compared to LP: the average number of mistakes per period was 0.040 vs. 0.099, respectively (two-sided Wilcoxon rank-sum test, $p = 0.296$). The same pattern was observed in both Math and GK sessions; differences in the number of mistakes between HP and LP in GK sessions are quite surprising since a negative correlation between the number of correct answers in Part 1 and the number of solved tables in Part 2 was found. By digging into the data we find that the difference between HP and LP is mainly driven by the provision of information about type and matching; indeed, average mistakes were 0.032 and 0.088 in sessions with information and no information, respectively (two-sided Wilcoxon rank-sum test, $p = 0.101$). This evidence suggests that mistakes are negatively affected by the knowledge of being classified among the LP and being matched with other LP.

5.3.4. Part 3: Initial coordination

I first consider behavior at both employee and firm level in the initial period to grasp a better understanding on how treatments affected initial beliefs. Indeed, it ought to be noticed that initial effort choices should mainly be affected by expectations about other employees behaviors and the treatments might have interfered with these expectations. Table 5.5 reports effort and minimum divided by treatment and type; overall, employees started off by exerting a medium or high level of effort: the average effort across all treatments was 2.66, the modal choice was 4 (52 out of 144 employees chose 4), and the second most chosen action was 2 (36 out of 144). Notwithstanding, little coordination was observed; in fact, none of our experimental firms managed to coordinate on the Pareto-efficient equilibria and the modal minimum was 1 (in 15 out of 36 firms).

TABLE 5.5: Effort, minimum, and mistakes by treatment and type, initial period only

	Overall	*Math-NI*		*Math-I*		*GK-I*	
		HP	LP	HP	LP	HP	LP
Effort	2.66	2.71	2.33	2.50	2.67	3.04	2.71
Minimum	1.17	1.50	1.00	0.83	1.00	1.83	1.00

The breakdown of effort and minimum by treatment and type (Table 5.5) can shed some light on the strategic uncertainty reduction hypothesis. Indeed, if grouping induces negative expectations among LP about the possibility of every employee solving all four tables, coordination failure could be determined by ability concerns, rather than coordination problems. The strategic uncertainty reduction argument can clearly apply to information treatments only; hence, we should observe higher levels of effort and coordination among HP in *Math-I* as compared to the other treatments. However, we do not find any statistically significant difference in both effort and minimum between information and no information treatments. The strategic uncertainty hypothesis also predict the difference between HP and LP to be larger when grouping conveys meaningful signal about ability (i.e., the effect of strategic uncertainty reduction should be larger in *Math-I* as compared to *GK-I*). Remarkably, LP performed better —in terms of both effort and minimum— than HP in *Math-I*, while the converse was true for *GK-I*. First period behavior provides some preliminary evidence against the uncertainty reduction hypothesis.

5.3.5. Part 3: Coordination dynamics

Table 5.6 describes effort (at the employee level) and minimum (at the firm level) over time, all sessions pooled. While the modal effort choice in the initial period was 4, it went down to 0 by the last period of the game; however, the modal effort choice across all nine periods was 4. Although most of the firms, in line with previous evidence, coordinated on the inefficient equilibrium (i.e., minimum equal to 0), it is interesting to notice that 6 firms managed to coordinate on the efficient equilibrium (i.e., minimum equal to 4) from period 5 throughout the end of the experiment.

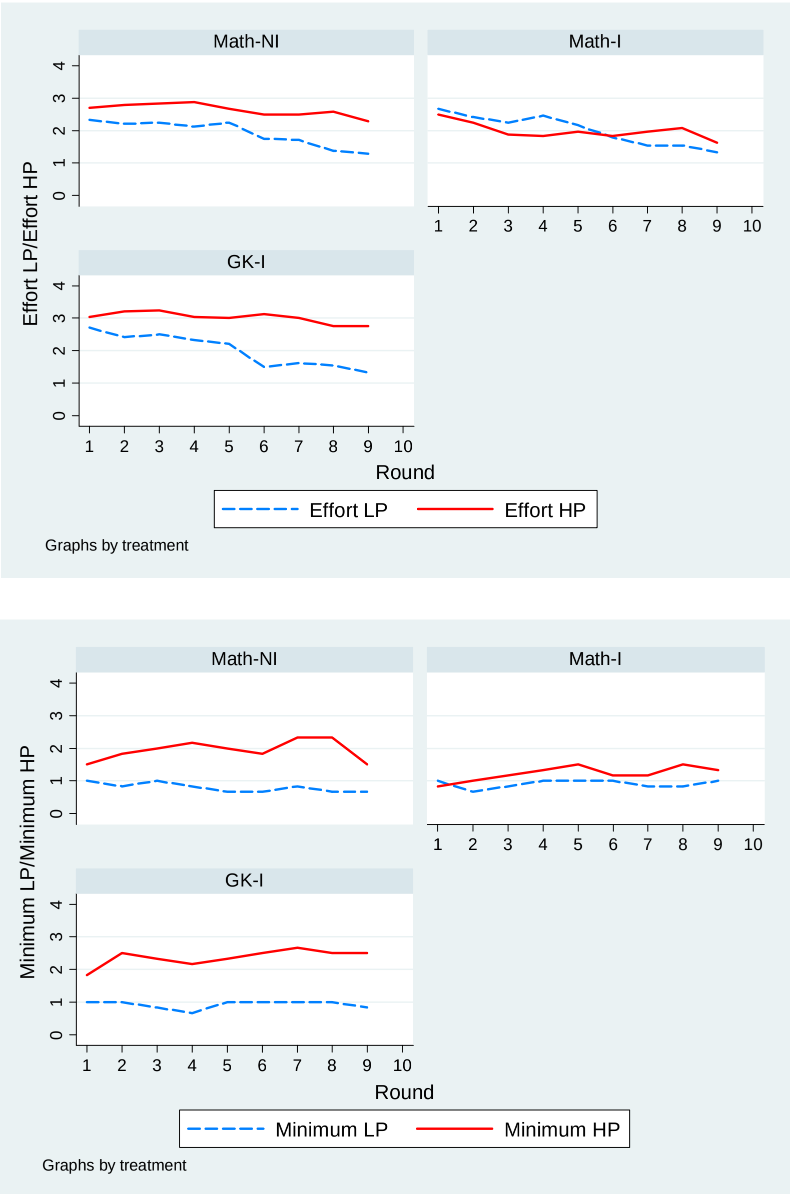
TABLE 5.6: Number of employees choosing each effort level (upper panel) and number of firms coordinating on each minimum level (bottom panel), all sessions pooled

		Period									Total
		1	2	3	4	5	6	7	8	9	
	0	9	15	20	26	24	38	35	39	48	254
	1	19	15	15	12	17	18	21	21	22	160
Effort	2	36	34	33	28	34	26	31	26	24	272
	3	28	36	26	28	19	18	15	20	15	205
	4	52	44	50	50	50	44	42	38	35	405

		Period									Total
		1	2	3	4	5	6	7	8	9	
	0	9	13	12	16	15	16	15	14	16	126
	1	14	6	7	3	6	6	4	7	7	60
Minimum	2	10	11	11	8	6	5	8	5	5	69
	3	3	5	4	6	3	3	3	4	2	33
	4	0	1	2	3	6	6	6	6	6	36

Turning to the differences between HP and LP, by pooling all treatments no statistically significant difference in effort was found according to a two-sided Wilcoxon rank-sum test effort ($p = 0.831$), while the minimum was different and significant, albeit weakly, according to the same test (p = 0.066). Interesting behavioral regularities can be found by considering the three treatments separately (see Figure 5.4).

FIGURE 5.4: Effort (upper panel) and minimum (bottom panel) over time, by treatment and type



Math-NI: this treatment is aimed to control for any possible correlation between math abilities and coordination in the stage game. We find some evidence that performance in the math task can affect subsequent behavior: in fact, HP exerted higher levels of effort (2.64 vs. 2.92) and achieved higher levels of coordination (1.94 vs. 0.80) as compared to LP, although the difference is not significant for both effort (p = 0.262) and minimum (p = 0.171) according to a two-sided Wilcoxon rank-sum test. To back empirical regularities observed in the *Math-NI* treatment, I also use a more powerful specification. Since the latent variables (effort and minimum) are ordinal, I use a ordered probit regression. The regressions both at the employee level (2) and at the firm level (3) include robust standard errors clustered at the firm level. The variable of most interest is the dummy for HP that test whether the difference in effort and minimum is higher among HP as compared to LP. Controls for time and individual characteristics (only for data at the employee level) are also included. Data from period 1 throughout 9 for *Math-NI* are included in the analysis.

$$Effort_{i,k,t} = \alpha + \beta_1 HP_i + \beta_2 Period_t + \beta_3 Male_i + \beta_4 RiskAversion_i + \\ + \beta_5 Age_i + e_{i,k,t} \quad (2)$$

$$Minimum_{k,t} = \alpha + \beta_1 HP_i + \beta_2 Period_t + e_{k,t} \quad (3)$$

Table 5.6, reports the results for both effort and minimum specifications; the main conclusion to be drawn from these results is that HP performs better, both at the employee level (i.e., exert higher levels of effort) and firm level (i.e., achieve higher minimum effort level) when no information about type and matching procedure are provided; however the difference is marginally statistically significant only at the firm level (i.e., minimum). As expected, *Period* and *Risk Aversion* have negative coefficients, although none of them is significant. In line with Dufwenberg and Gneezy (2005), no gender differences are detected.

Table 5.6: Ordered probit regression on data from *Math-NI*, all periods

	Effort (individual level)	Minimum (firm level)
HP	0.524	0.885*
	(0.384)	(0.512)
Period	-0.064	-0.013
	(0.041)	(0.054)
Male	-0.031	--
	(0.250)	
Risk aversion	-0.091	--
	(0.067)	
Age	0.037	--
	(0.042)	
Log-likelihood	-649.078	-151.367
#Observations	432	108

NOTE: * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

Two main explanations can account for better, although not significant, coordination in firms composed by HP: first, given the correlation between adding up numbers and counting letters, there is a difference in the number of mistakes made in the coordination phase; second, ability in solving additions is related with determinants of coordination, such as trust, beliefs about the other members of the firm or ability to understand the game. The first argument can be ruled out, given the negligible number of mistakes made during the coordination phase which was very unlikely to hamper coordination alone. As for risk aversion, we found no difference in risk attitudes between HP and LP in Part 4.

RESULT 2(a): Contrary to Hypothesis 2(a), effort and minimum over time are higher, albeit not statistically significant, for HP as compared to LP in Math-NI treatment.

Math-I: either group pride or the reduction of strategic uncertainty was expected to enhance coordination among HP. Figure 5.4 shows that both effort and minimum in *Math-I* treatment are basically indistinguishable across HP and LP. Surprisingly enough, information about type and matching had, if anything, a negative impact on HP in the math treatment; indeed, a comparison between *Math-NI* and *Math-I* reveals that effort and minimum were higher when HP did not receive any information. However, the difference in both effort and minimum between HP in *Math-NI* and *Math-I* was not significant at any conventional level.[6]

$$e_{i,k,t} = \alpha + \beta_1 LP_i + \beta_2 NoInfo_i + \beta_3 LP_i * NoInfo + \beta_4 Period_t + \beta_5 Male_i + \beta_6 RiskAversion_i + + \beta_7 Age_i + e_{i,k,t} \quad (4)$$

Table 5.7: Ordered probit regression on data from *Math* treatments, all periods

	Effort (individual level)	Minimum (firm level)
LP	0.004	-0.344
	(0.429)	(0.644)
NoInfo	0.515	0.568
	(0.472)	(0.553)
LP*NoInfo	-0.546	-0.562
	(0.572)	(0.794)
Period	-0.075***	0.006
	(0.026)	(0.030)
Male	-0.219	--
	(0.169)	
Risk aversion	-0.029	--
	(0.034)	
Age	0.006	--
	(0.038)	
Log-likelihood	-1311.42	-290.558
#Observations	864	216

NOTE: * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

[6] Two-sided Wilcoxon rank-sum test (p = 0.262 and p = 0.261 for effort and minimum, respectively)

The same results are confirmed by an ordered logit regression on *Math* data (see, Table 5.7). The dummy variable *NoInfo* capture the effect of the absence of information on HP; indeed, the positive coefficient, even though not significant, suggests that the absence of information about type and matching had a negative effect.

This evidence clearly suggests that reduction of strategic uncertainty alone cannot account for the higher level of coordination. In addition, not every task is likely to promote group pride and thus enhance coordination.

> *RESULT 2(b):* The provision of information had a negative effect on both effort and minimum on HP; that is, both effort and minimum among HP were higher in *Math-NI* as compared to *Math-I*.

GK-I: in line with results in Study II, and contrary to evidence in *Math-I* treatment, a positive and significant effect over time of group pride on HP was found in the GK sessions. The average effort was 3.02 and the minimum was 2.37 for HP, while effort was 2.02 and minimum 0.93 for LP; the difference between HP and LP was significant at 5 percent level according to a two-sided Wilcoxon rank-sum test for both effort ($p = 0.037$) and minimum ($p = 0.036$). Support to the observed regularities is also provided by results of an ordered probit regression reported in Table 5.8. In this case, the enhanced coordination among HP firms must be due to group pride rather than reduction of strategic uncertainty, since both actual and perceived similarity between the task used to group employees and the task used in the stage game was extremely small. The present results suggest that grouping is not enough to induce group pride; moreover, the task used to group employees does not necessarily need to be correlated with the main task, but most probably need to be perceived as valuable. We find support for this conjecture by comparing the two treatments with information (I.e., *Math-I* and *GK-I*); in HP firms, both effort (p = 0.078) and minimum (p = 0.092) are higher and significant at 10 percent level in *GK-I* as compared to *Math-I*.

Table 5.8: Ordered probit regression on data from *GK-I*, all periods

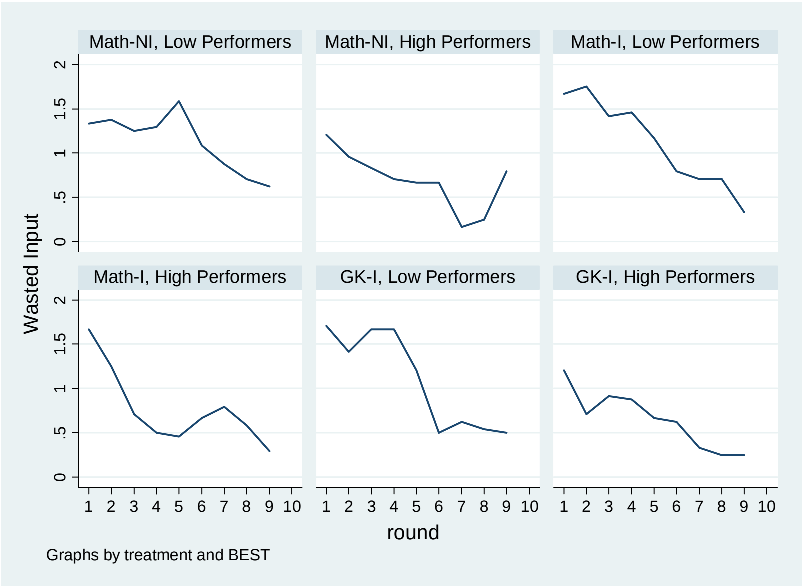
	Effort (individual level)	Minimum (firm level)
HP	0.633**	1.702***
	(0.277)	(0.553)
Period	-0.106***	-0.046
	(0.028)	(0.058)
Male	0.255	--
	(0.227)	
Risk aversion	-0.001	--
	(0.050)	
Age	-0.008	--
	(0.090)	
Log-likelihood	-640.825	-129.212
#Observations	432	108

NOTE: * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

RESULT 3: Both effort and minimum over time are higher for HP as compared to LP in GK-I treatment, hence, as conjectured, group pride looms larger than the strategic uncertainty reduction argument.

To better understand the difference in behavior let consider how type (I.e., HP and LP) affects efficiency and, more specifically, wasted input over time. Figure 5.5 compares wasted input over time broken down by treatment and type; it can be noticed that wasted input was smaller for HP as compared to LP (two-sided Wilcoxon rank-sum test, p = 0.002 by pooling all treatments). However, the only treatment where the difference was significant was the *GK-I* (p = 0.025).

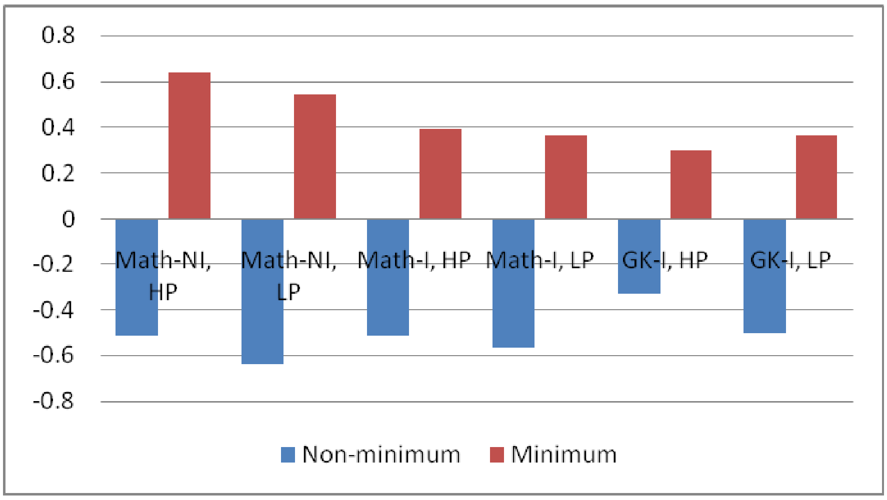
FIGURE 5.5: Wasted input over time, by treatment and type



Finally, I consider how subjects adjusted their behavior over time. Specifically, I divided employees in two categories: employees that in the previous period exerted an effort higher than the minimum are defined 'non-minimum', while those who exerted an effort equal to the minimum are labeled 'minimum'. Figure 5.6 reports the average change in effort from one period to the other divided by minimum/non-minimum employees, treatment and type.[7] Variation for HP non-minimum employees is smaller as compared to LP non-minimum employees across treatments. Indeed, the effort increase for minimum employees is larger or almost identical for LP as compared to HP in both treatments with information.

[7] Only non-equilibrium periods are included in the analysis.

FIGURE 5.6: Wasted input for minimum and non-minimum employees, by treatment and type



5.4. Discussion and Conclusions

Building on Study II, I explored whether varying the task used to divide employees in HP and LP affects coordination in the minimum-(real)effort game. Manipulation of the degree of similarity between the task in Part 1 and 2 (and 3) was implemented to disentangle the relative contribution of group pride and uncertainty reduction. In line with Hypothesis 1, correlation between performance in Part 1 and 2 (and 3) was different between *Math* and *GK* treatments. Correlation between adding up numbers and counting letters was positive, while correlation between general knowledge and counting letters was negative, although small. Thus, performance in the general knowledge quiz does not provide any valuable information about ability in counting letters in a table. Since the uncertainty reduction hypothesis is based on perceived similarities, and not on actual similarities that are not known to the subjects, measures of perceived similarity were collected. Similar results for actual and perceived correlation were found; thus, performance in the GK quiz is not perceived as a valuable information to make any inference about performance in the counting task.

Based on evidence from Study II, the beneficial effect of grouping was expected to apply to *Math-I* but not *Math-NI* treatment. Contrary to Hypothesis 2(a), in *Math-NI* HP exerted higher levels of effort and achieved higher levels of coordination as compared to LP. Although the difference was not significant, it suggests that some sort of correlation between solving additions and coordination are in place. Two main explanations can account for the observed behavior. Fist, HP may have been satisfied with their performance although they did not knew how they ranked. Satisfaction might have induced some changes in the level of self-esteem (or any other positive affective state) that interacted with coordination. Second, adding up numbers may be a proxy for cognitive abilities which have shown to interact with behavior in different classes of games (e.g., Devetag and Warglien, 2003). Surprisingly enough, information about type and matching had, if anything, a negative impact on HP in the math treatment. Indeed, a comparison between *Math-NI* and *Math-I* reveals that effort and minimum were higher when HP did not receive any information. The negative effect of information among HP, is very puzzling and neither group pride nor uncertainty reduction can account for the observed behavior. Behavior among HP in the *Math-I* treatment shares some commonalities with results reported in Butler (2008); the author have found that high status members punish lack of trust more than low status members do. The steady decline in wasted effort observed in the very first periods of the game suggests that non-minimum employees decided to dramatically reduce their effort since they did not "forgot the lack of trust" in the first round. The observed negative reaction and the steady decrease of effort is at odd with evidence form Study II (and from some psychological studies, see Pierce et al., 1993) where firms with a higher group-pride level were more patient and managed to hill-climb the strategy space.

In contrast, the expected effect of grouping was found in the *GK-I* treatment. Since the uncertainty reduction argument only applies to correlated tasks, but not to general knowledge questions, enhanced coordination among HP cannot be determined by uncertainty reduction. The present results suggest that grouping is not enough to induce group pride; moreover, the task used to group employees does not necessarily need to be correlated with the main task, but most probably need to be perceived as valuable. Further research is thus needed to better understanding how different tasks affect group pride.

Appendix 5: Instructions

Instructions

General information: The purpose of this experiment is to study how people make organizational decisions. 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 privately.

The experiment will have four parts. Parts 1 and 2 both last one round. . These two parts are strictly individual; you will not be matched with anyone else in the room. Part 3 will have nine rounds and you will be matched with 3 other participants in this room for all nine rounds. Part 4 has a brief questionnaire. The instructions for each part will be read just before the corresponding part.

In each part of this experiment you are able to earn money. In part 3 what you earn results from your decision and the decisions of the 3 other participants in this room that you will be matched with. In the other parts your earnings are not dependent on the decisions of others. 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 200 ECUs. Upon completion of the experiment, the amount that you earned will be paid to you in cash. Payments are private; no other participant will be told the amount you earned.

Instructions (Part 1)

Description of the task: You will be asked to [solve as many additions [answer as many general knowledge questions] as you can. You will have YY minutes of time to [solve additions/answer questions] and fill out the task sheet. The experimenter will let you know when the time allotted to this part has elapsed and she wants you to stop working on the task assigned. Please do so immediately and without delay. If you do not follow the experimenter's s instructions, she will have to dismiss you from the experiment. In that case, you will not be paid anything.

The experimenter will also tell you when you can turn over the task sheet on your desk and start working on it. You will have to report your answer in the blank space just below each question. You can answer the questions in any order you like. The only important thing is that you try to answer as many questions as possible.

Earnings: you will be paid a fixed amount of 2 Euros for this part, no matter how many (additions you solved correctly) [questions you answered correctly]. Although the payment is fixed, we ask you to do the task at your best.

Questions?

Turn over the task sheet now!

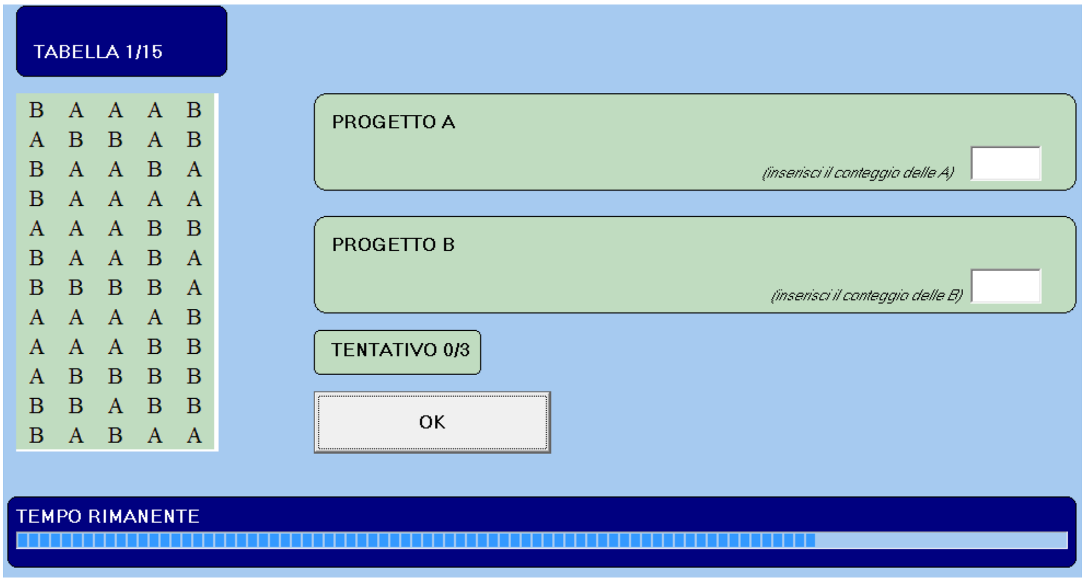
Instructions (Part 2)

Parts and rounds: Part 2 will last 1 round during which you will not be matched with anyone. Part 2 has the purpose of giving you the chance to familiarize yourself with the task you will be assigned in part 3. The following instructions are about part 2 only.

Description of the task : You can think of yourself as an employee of a firm that is assigned the following task:

- 15 tables containing the letters A and B will be displayed on the screen one after the other; your task consists of counting the number of either As or Bs contained in each table;
- for each table, you may choose to work either for project A (which implies that you will be counting As) or for project B (which implies that you will be counting Bs);
- you will have 3 minutes to complete as many tables as you can, that is to count either the number of As or the number of B2 contained in each table; this time will not be extended.

Practical implementation: a screen like the one shown below will be displayed on your computer.



A table containing As and Bs will appear on the left side of your screen; you can start completing your task as soon as the table is displayed on the screen. If you decide to contribute a table to project A, you have to count the number of A and report this number in the blank space labeled “Number of A”; if you decide to contribute to the project B, you have to count the number of B and report this number in the blank space labeled “Number of B”. You have to confirm your number by clicking on “OK”. If the number (of either A or B) you entered is correct and you clicked on “OK”, the next table will be displayed. In case the number you entered is not correct, a pop-up window will open right above the two green fields and will inform you that the number is wrong; you will have 3 trials for each table, after the third trial, no matter if the number is correct or not, the next table will be automatically displayed.

For each table, you can independently decide whether to contribute to either the project A or B. That is, if you start working for project B you can decide to switch to project A for next table. You can switch from one project to the other as many times as you want. However, note that, for each table, you can contribute to ONLY ONE of the two projects.

In part 2, you will have 3 minutes of time to complete as many tables as you can The time left is reported on the bottom part of the screen. The number of the tables and trials are also reported on the screen.

Earnings: you will be paid 2 Euros for this part, independent of how many tables (and for which project) you completed (them). It is very important that you try to solve as many tables as possible within the time given.

Information that you will receive: you will be informed about the number of tables you successfully completed for both project A and B. Please remember that if any type of communication is detected during the experiment, the session will be immediately concluded and nobody will be paid. If you have questions, please raise your hand and one of the experimenter will come to your desk to answer it privately.

Instructions (Part 3)

Math-I and GK-I treatments marked in yellow

Parts, rounds, and groups: Part 3 has 9 rounds. In part 3, you will be matched with 3 other participants in this room; the matching will depend on the performance in part 1. Specifically, participants in this room will be assigned to one of two categories:

3. *Best Performers:* this category will be formed by the 8 employees who correctly solved the highest number of additions [correctly answered the highest number of questions] in part 1;
4. *Worst Performers:* this category will be formed by the 8 employees who correctly solved the smallest number of additions [correctly answered the smallest number of questions] in part 1.

You will be matched with 3 other employees from the same category. For instance, if you are among the Best *Performer (Worst Performer)* the other 3 members of your group will be *Best Performers (Worst Performers)* as well.

Before part 3 will start, you will receive an envelope containing information about your performance in part 1 (*Best Performer* or *Worst Performer*) and consequently the type of group your will be working with (*Best Performers Firms* or *Worst Performers Firms*)

Once determined, the composition of a group [???] will be the same for all rounds.

The following instructions pertain to part 23.

Description and implementation of the task : You can think of yourself and the other members of your group as employees of a firm.

Task and implementation are the same as in part 2; the only difference is that in each round four tables at most can be completed. The time allotted for each round is 3 minutes, as in part 2.

<u>Payoffs</u>: Your earnings in each round depend on how many tables you and the other employees of the firm will successfully complete: a table will be counted only in case the number of either A or B which have been counted is correct. Moreover, your payoff will depend on how many tables you and the other members of your firm will decide to solve for the project A and B. Please, recall that completing a table for project A means that you correctly counted A, and completing a table for project B means that you correctly counted B.

For each successfully completed table you contribute to project B, you will get 50 ECUs no matter what other employees of the firm do. The payment for project A will be 80 ECUs multiplied by the MINIMUM number of tables a member of your firm (you included) has contributed to project A. That is, in order to determine the minimum contribution to project A (*min(A)*), you have to consider the number of tables solved for project A by each of the employees of the firm (you included) and pick the MINIMUM number. For example, if the number of completed tables for project A is 4, 3, 2, and 0, the smallest number is 0 and hence *min(A)=0*.

Formally, your earnings are determined by the following payoff function:

$$\pi = 80*[min(A)] + 50*[B] \qquad \text{(a)}$$

Please note that incorrectly or not complete tables will not be paid.

To help you understand better how your earnings are determined, below please see we report the payoff table for a particular case: the following payoff table is, in fact, **ONLY FOR THE CASE IN WHICH YOU COMPLETE ALL OF THE FOUR TABLES** IN THE ROUND. PLEASE, REMEMBER THAT THE FOLLOWING TABLE DOES NOT EXHAUST ALL POSSIBLE CASES, which can be derived by using the payoff function (a). Cases where not all 4 tables are completed will be discussed later on. All numbers in the non-shaded cells represent ECUs.

		Minimum contribution to project A: *min(A)* (number of tables)				
		4	**3**	**2**	**1**	**0**
Your contribution to project A (in brackets the number of tables you contributed to project B)	**4** (0)	320	240	160	80	0
	3 (1)		290	210	130	50
	2 (2)			260	180	100
	1 (3)				230	150
	0 (4)					200

For easy reference we also have provided you with a paper copy of the earnings table and function on your desk.

Focus initially on the first row (shadowed in grey) in the earning table; in that row, the possible MINIMUM contributions to project A are listed. *min(A)* ranging from 4 [the case where all employees of the firm contributed four tables to project A] to 0 [the case one or more members contributed zero tables to project A]. For example, if the number of tables completed for project A are: 4, 3, 2, and 2, then the minimum is 2; hence, you have to check the column identified by "**2**" in the first row for your potential earnings.

Focus next on the first column (shadowed in grey) in the above payoff table; in that column, the number of tables YOU can contribute to project A is reported in bold. In brackets, the number of tables you can contribute to project B is reported. For instance, you had successfully completed 4 tables for project A (and, as a consequence, 0 for project B): you would have had to check the row identified by "**4** (0)" in the column for your potential earnings.

Notice that the sum of tables contributed to project A and B is always four; that is because **ONLY** THE CASE OF FOUR TABLES SUCCESSFULLY COMPLETED IS REPORTED IN THE PAYOFF TABLE. You can determine your earnings in a round by crossing the relevant rows and columns. Consider the previous examples in which *min(A)*=2 and your contribution to project A was 4; your earnings are then determined at the intersection of row "**4** (0)" and column "**2**" yielding a payoff of 160. Please, notice that this payoff can be computed also by using the payoff function (a):

$$\pi = 80*[\mathit{min(A)}] + 50*[B]$$

$$|$$

$$= 80*[2] + 50*[0] = 160$$

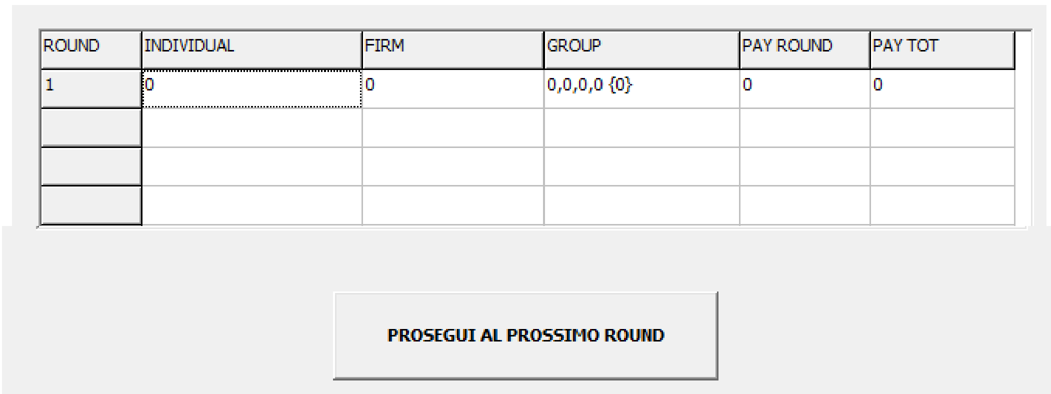
Consider the same example, but suppose you contributed 2 to project A, 1 to project B, you failed to complete one table, and the minimum is still 2. You cannot determine your earnings by using the payoff table, since you did not completed all of the four tables; hence, you have to use the payoff function (a):

$$\pi = 80*[min(A)] + 50*[B]$$

|

$$= 80*[2] + 50*[1] = 160 + 50 = 210$$

<u>Information that you will receive</u>: Each round, after all employees have finished their task or the time is over, the computer will display a screen like the one shown below. In the first column the number of the round is reported. The actual number of tables you completed for project B is reported in the second column, and the number of tables you completed for the project A is displayed on the third column. In the fourth column the number of tables each member of the group (you included) contributed to project A is displayed; in curly brackets *min(A)* is reported. Your payoff for each round, and your cumulated payoff is shown in the last two columns. When you are ready to start a new round, please click on “Go to next round” and a new table will be displayed.



At the end of the 10 rounds , information about that round will be displayed and a message on the screen will alert you that the experiment 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 in all the nine rounds of part 3 plus the flat payment for part 1 2, and 4. The conversion rate is 1 Euro for 200 ECUs.

To make sure that you have correctly understood how your earnings are computed, please answer to the questions in the anonymous questionnaire you have on your desk.

Chapter 6

Summary and Concluding Remarks

The present dissertation illustrates the potential of an experimentally-grounded approach to organizational coordination problems (e.g., Griffin and Kacmar, 1991; Camerer and Weber, 2007) and, more generally, organizational design. Even though coordination problems have been a subject of major concern for both experimental economists (for a review, Camerer, 2003) and organizational scholars (for a review, Okhuysen and Bechky, 2009), the two disciplines have only marginally informed each other. I investigated organizationally-relevant elements by means of minimum-(real)effort lab experiments, frequently appealed to well describe a host of real world situations (e.g., airplane take-off, chemical production, writing a chapter in a multiple-authors book). The main novelty of the present dissertation is the introduction of real, rather than chosen, effort in a coordination game of the minimum-effort type. The introduction of real-effort tasks in a lab setting goes in the direction of reducing the gap in abstraction between controlled experiments and actual organizations, thus increasing the external validity of the present results. The concern about reliability of chosen-effort experiments stems from recent evidence that chosen effort is not necessarily a good proxy for real effort (see Chapter 2 for a discussion). The new testbeds developed in this dissertation were used to study how different types of incentives, both monetary and non-monetary, affect coordination when extreme task interdependence is present. Specifically, Study I compares the effectiveness of group and individual incentives, while Study II and III explore the role of group-pride on coordination.

In *Study I*, an experiment where subjects had to sort and count coins worth 1, 2, 5, and 10 Euro cents within a given time interval was designed. The time interval was set so that task completion was not trivial and required quite some effort; moreover, since the time allotted was likely to be, at least initially, not sufficient for everyone, subjects were offered the opportunity to buy extra time to complete the task. There were two main payment conditions: in *Group* treatments, subjects were randomly assigned to groups of four and their earnings were a function of the worst counting performance of a group member; in *Individual* treatments, strategic interaction was stripped away by having subjects "work" alone and paying them according to their individual performance.

The observed behavior is in sharp contrast to the speedy downward drift to the Pareto-dominated equilibrium commonly observed in previous minimum-effort experiments, hence suggesting that chosen effort might not be a reliable measure of real effort. Indeed, after initially failing, almost 80 percent of the groups succeeded in coordinating on the efficient equilibrium. This behavior is inconsistent with adaptation but recalls "signaling" and "strategic teaching" dynamics (e.g., Camerer et al., 2002; 2003) observed in previous chosen effort weak-link experiments. It may be that subjects tried to perform at their best pushed by some sense of competitiveness or peer pressure that may exist in groups engaged in real-effort tasks. Shame of not being able to complete the assigned task not present in an abstract chosen-effort experiments may also account for the observed behavior.

The comparison between Group and Individual treatments serves to shed light on two issues: first, learning dynamics in performing the task at hand; second, relative effectiveness of group-based and individual-based incentive schemes. Effort in both *Group* and *Individual* treatments are largely indistinguishable in Study I. Since strategic uncertainty and interaction are stripped away in the *Individual* treatments, changes in the accuracy of performance over time are likely to depend upon task learning only. The similarity of behavioral patterns observed across treatments suggests that successful coordination was 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. Shame or peer pressure may account for the fact that extremely poor performers in the *Group* conditions improved their performance in the next round with a significantly higher frequency than in the *Individual* condition. This 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), and have important prescriptive implications for actual organizations.

In *Study II* subjects had to count the number of letters contained in a sequence of tables and in each period had to decide how to allocate their effort between two projects. The task was designed in such a way to recreate the prototypical features of the minimum-effort. After playing an initial part where they had to solve as many tables as they could within the allotted time, subjects were matched in groups of four based on the number of solved tables in the first part of the experiment: each group was formed either by best or worst performers, defined according to a median split. The feedback about relative performance in the first part and matching were varied to induce different levels of group pride. While each subject was informed about his/her type (i.e., high or low performer), and the matching protocol was common knowledge in the *Info* treatment, no information about both type and matching was provided in the *NoInfo* treatment, this being the only difference between the two treatments.

No differences in initial choices and, as a consequence, beliefs are detected between high and low performers' firms. Moreover, no difference in behavior and coordination between high and low performers was found in *NoInfo* treatment, while both effort and minimum effort for high performers was statistically higher in the *Info*. Indeed, high performers in the *Info* treatment managed to achieve efficient coordination (i.e., Pareto-dominant equilibrium) over time in half of the cases, while none of the other groups did so. This finding is primarily interesting for actual organizations since group pride may apply to all those groups that have some (real or imaginary) reason to be proud of themselves, and these are far more numerous than so-called elite groups *strictu sensu,* including successful work teams of any kind.

In *Study III* the main focus was on how different tasks used to group subjects affect group pride and coordination. In the initial part of the experiment, subjects were asked to either solve mathematical additions (*Math* treatments) or answer general knowledge questions (*GK* treatment); groups based on performance in the initial task were formed in the coordination part. As in the previous study, subjects were asked to count numbers in a sequence of tables in the coordination part. Following the protocol used in Study II, homogenous groups (high vs. low performers groups) were formed and feedback about type and matching were varied (*I* vs. *NI*). Both adding up numbers and counting letters are related to mathematical skills; hence, ability and performance in the two tasks is expected to be, and to be perceived as, correlated. In contrast, one should expect that there is no correlation —both actual and perceived— between performance in the general knowledge quiz and the counting task. *GK*-I treatment was meant to disentangle the relative contribution of group pride and strategic uncertainty reduction.[1] Thus, in case no

[1] Strategic uncertainty reduction hypothesis refers to the fact that high and low performers developed different beliefs about the feasibility of the efficient equilibria. While being grouped with high performers might generate positive expectations about everyone in the firm being able to

correlation between performance in the GK quiz and counting is detected, any difference in behavior between HP and LP in *GK-I* treatment has to be attributed to group pride and not to strategic uncertainty reduction.

In line with the hypothesis, correlation —both actual and perceived— between adding up numbers and counting letters was positive, while correlation between general knowledge and counting letters was negative. Thus, scoring high in the GK questionnaire has no, if not negative, effect on performance in the clerical task. As in Study II, only a negligible number of mistakes evenly distributed across periods was observed. The data suggest that performance in the math task can affect subsequent behavior when no information is provided; namely, high performers coordinated on higher effort levels as compared to low performers in *Math-NI*. Surprisingly enough, information about type and matching had, if anything, a negative impact on HP in the math treatment; thus, reduction of strategic uncertainty alone cannot account for the higher level of coordination achieved by high performers in Study II. In line with results in Study II, and contrary to evidence in *Math-I*, the data show a positive and significant effect over time of group pride on high performers in our GK sessions. More generally, Study II seems to suggest that grouping is not enough to induce group pride. Indeed, the task used to group employees does not necessarily need to be correlated with the main task, but most probably need to be perceived as valuable. Some support for this conjecture was found in the data.

complete all tables, low performers might think that the other employees within their firm are not able to successfully complete all tables, hence ruling out some equilibria from the strategy space.

Concluding remarks and future research

The experimentally-grounded approach to organizational design advocated in the present dissertation open the door to several unsettled questions and methodological caveat. Indeed, despite some creditable attempts to explore organizational issues in, and to increase external validity of, lab experiments, the gap in scope and abstraction between experimental and organizational evidence continue to exist. In the following, I list a few directions for future research in this area.

As discussed in Chapter 2, most of the organizational mechanisms implemented in the minimum-effort game (e.g., communication, leadership, financial incentives, etc.) helped avoiding or overcoming coordination failure otherwise commonly observed. Borrowing from organizational evidence can thus prove a useful source of ideas and can help finding new solutions to puzzling behaviors observed in the lab (e.g., coordination failure). However, we still know surprisingly little about the impact of a number of organizational theories and findings potentially relevant for coordination. A list of the areas still open to research includes: the effect of work rotation and interaction between experienced and inexperienced employees in a process of growth; the role of downsizing concerns in an endogenous process of growth; the relative efficiency of different communication conduits (e.g., face-to-face vs. chat); the influence of personality traits on coordination; the introduction of some degrees of conflict of interests between employees and leaders; the interaction between individual-based and group-based incentives. In particular, I intend to study in a future experiment how individual and group incentives interact when full information about individual effort is not available to the manager in charge of awarding the individual bonuses. A better understanding of these dynamics and the relative effectiveness of objective and subjective awarding procedure has important managerial implications for a better design of contracts and career plans.

The proposed shift from chosen to real effort, although problematic with some respects such as control over individual abilities and disutility from effort, seems to be an interesting way to introduce more realism in the lab. First, real effort captures an essential element of actual organizations: uncertainty about which equilibria are feasible and which are not. Indeed, in both actual organizations and real-effort experiments there are some cases in which the efficient outcome cannot be attained simply because some of the subjects involved in the production process are incapable, and not unwilling, to correctly perform the assigned task. The introduction of some degrees of uncertainty about the feasibility of the action space clearly points in the direction of a variety of organizational studies that have devoted a great deal of attention to the issue (for a review, see Okhuysen and Bechky, 2009), but it often comes at the cost of control. A better understanding of the effect of this additional uncertainty represents an interesting direction for future researches (Study III is a first attempt to control for this element; for a similar exercise, see also Hennig-Schmidt et al., 2004) Second, real effort can trigger heuristics developed in everyday experience that are not present in an abstract experiment (e.g., Ortmann and Gigerenzer, 2000 ; Harrison and List, 2004). For instance, the rule of working rather than shirk can be in place when performing a real task, but not when a cost of effort has to be chosen. A more precise understanding of the heuristics that real effort bring about constitute an interesting direction for future research.

Despite the great potentials, the process of blending experimental and organizational methodologies and insights is still undeveloped. Almost three decades ago, Schwenk (1982) exhorted a dialectic inquiry between experiments and strategic management in a continuous back and forth between theory, lab, and field. For the time being, the advocated dual process has remained largely unattended. Hence, the development of the proposed dual approach seems to be an important challenge with high stakes. Similarly to the design of markets and auctions (e.g., Roth, 2002), experiments could be tailored to specific organizational contexts to help advise managers on specific problems by first testing alternative solutions in the lab.

References

Abeler, J., Falk, A., Goette, L., and Huffman, D. (forth.). Reference Points and Effort Provision. *American Economic Review*.

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

Akerlof, G. A., and Kranton, R. E. (2000). Economics and Identity. *Quarterly Journal of Economics*, *115* (3), 715–753.

Akerlof, G. A., and Kranton, R. E. (2005). Identity and the Economics of Organizations. *Journal of Economic Perspective*, *19 (1)*, 9–32.

Altmann, S., Falk, A., and Wibral, M. (2008). Promotions and Incentives: The Case of Multi-Stage Elimination Tournaments. IZA DP No. 3835.

Argote, L. (1982). Input uncertainty and organizational coordination in hospital emergency units. *Administrative Science Quarterly, 27,* 420–34.

Awasthi, V., and Pratt, J. (1990). The effects of monetary incentives on effort and decision performance: the role of cognitive characteristics. *Accounting Review, 65,* 797-811.

Azar, O. H. (2009). Does relative thinking exist in mixed compensation schemes? Working paper, Ben-Gurion University of the Negev.

Ball, S., and Eckel, C. (1996). Buying Status: Experimental Evidence on Status in Negotiation. *Psychology and Marketing*, *13(4)*, 381-405.

Ball, S., Eckel, C., Grossman, P., and Zame P. (2001). Status in Markets. *The Quarterly Journal of Economics, 116(1)*, 161-188.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, *84*, 191–215.

Bangun, L., Chaudhuri, A., Prak, P., and Zhou, C. (2006). Common and almost common knowledge of credible assignments in a coordination game. *Economics Bulletin*, *3*, 1–10.

Barnard, C. I. (1938). *The Functions of the Executive*. Cambridge: Harvard University Press.

Battalio, R., Samuelson, L., and Van Huyck, J. (2001). Optimization incentives and coordination failure in laboratory stag-hunt games. *Econometrica*, *69,* 749–764.

Bechky, B. A. (2006). Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science*, *17*, 3-21.

Benabou, R., and Tirole, J. (2002). Self confidence and personal motivation. *Quarterly Journal of Economics*, *117(3),* 871-915.

Benabou, R., and Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, *70*, 489-520.

Berglas, S. (2006). How to keep A players productive. *Harvard Business Review*. *September*, 105-112.

Berglas, S., and Jones, E. E. (1978). Drug choice as a self-handicapping strategy in response to non-contingent success. *Journal of Personality and Social Psychology*, *36(4)*, 405-417.

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.

Bigley, G. A., and Roberts, K. H. (2001). The incident command system: Organizing for high reliability in complex and unpredictable environments. *Academy of Management Journal*, *44*, 1281-1299.

Blume, A. (2000). Coordination and Learning with a Partial Language. *Journal of Economic Theory*, *95*, 1-36.

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.

Blume, A., Kriss, P., and Weber, R. A. (in prep). Endogenous Costly Communication and Equilibrium Selection.

Boland, J. R., and Tenkasi, R. V. (1995). Perspective Making and Perspective Taking in Communities of Knowing. *Organization Science*, *6(4),* 350-372.

Bornstein, G., and Yaniv., I. (1998). Individual and Group Behavior in the Ultimatum Game: Are Groups More ‘Rational’ Players? *Experimental Economics*, *1(1),* 101–108.

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

Bosman, R, and van Winden, F. (2005). The Impact of Emotions in a Power to Take Game. *Journal of Economic Psychology, 26(3),* 407-429.

Bosman, R, Sutter, M., and van Winden, F. (2002). Emotional hazard in a power to take game. The Economic Journal, *112(476),* 147-169.

Bosman, R., and van Winden, F. (forth.). Global Risk, Investments and Emotions. *Economica*.

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*, *5(6),* 1223–1268.

Brandts, J., and McLeod, B. (1995). Equilibrium Selection in Experimental Games with Recommended Play. *Games and Economic Behavior, 11,* 36–63.

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

Brandts, J., Cooper, D. J., and Fatas, E. (2009). Stand By Me. Help, Heterogeneity and Commitment in Experimental Coordination Games. Mimeo.

Brown, R. (2000). Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology*, *30*, 745–778.

Brüggen, A., and Strobel, M. (2007). Real effort versus chosen effort in experiments. *Economics Letters*, *96(2)*, 232-236.

Bull, C., Schotter A., and Weigelt, K. (1987). Tournaments and Piece Rates: An Experimental Study. *Journal of Political Economy*, *95*, 1–33.

Butler, J. (2008). Trust, truth, status and identity: an experimental inquiry. Working Paper, available at http://www.eief.it/files/2008/11/trust_latest.pdf

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

Calsamiglia, C., Franke, J., and Rey-Biel, P. (2009). The Incentive Effects of Affirmative Action in a Real effort Tournament. Mimeo.

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

Camerer, C. F., and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19(1),* 7-42.

Camerer, C. F., and Weber, R. A. (2007). Experimental Organizational Economics. To appear in *Handbook of Organizational Economics*, eds. R. Gibbons and J. Roberts.

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

Carlile, P. R. (2002). A pragmatic view of knowledge and boundaries: boundary objects in new product development. *Organization Science, 13 (4),* 442-455.

Carpenter, J. P., Matthews, . H., and Schirm, J. (forth.) Tournaments and office politics evidence from a real effort experiment. *American Economic Review.*

Cason, T., Sheremeta, R., and Zhang, J. (2009). Reducing Efficiency through Communication in Competitive Coordination Games. Available at http://www.krannert.purdue.edu/faculty/cason/papers/Weak-Link-Comm.pdf

Cesarini, D. Sandewall, O., and Johannesson, M. (2006). Confidence interval estimation tasks and the economics of overconfidence. *Journal of Economic Behavior & Organization*, *61(3),* 453-470.

Charness, G., and Villeval, M. C. (2009). Cooperation Competition and risk attitudes: An intergenerational field and laboratory experiment. *American Economic Review*, *99(3),* 956–978.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, *97*, 1340-1352.

Chaudari, A., and Paichayontvijit, M. (2007). Credible assignments and Performance Bonuses in the Minimum Effort Coordination Game. Mimeo.

Chaudhuri, A., and Bangun, L. (2007). Credible assignments in the minimum effort coordination game. Mimeo.

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

Che, Y.-K., and Yoo, S.-W. (2001). Optimal incentives for teams. *American Economic Review*, *91 (3),* 525–541.

Chen R., and Chen. Y. (2009). The Potential of Social Identity for Equilibrium Selection. Working Paper available at http://www.si.umich.edu/~yanchen/papers/ID_EquilSelect.pdf

Cohen, M. D., and Bacdayan, P. (1994). Organizational routines are stored as procedural memory: Evidence form a laboratory study. *Organization Science*, ***5****(4)*, 554–568.

Cojuharenco, I., and Gozluklu, A. E. (2008). The Visible Hand of Leadership: Leaders.Attributions of Coordination Outcomes In Small and Large Subordinate Groups. Available at http://arieskenazi.com/tessina160308.pdf.

Collins, D., Ross, R. A., and Ross, T. L. (1989). Who Wants Participative Management? G*roup and Organization Studies*, *14(4)*, 422-445.

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

Cooper, D. J., and Kagel., J. H. (2005). Are Two Heads Better Than One? Team Versus Individual Play in Signaling Games. *American Economic Review*, *95(2),* 477–509.

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

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

Cowen, T., and Glazer, A. (2007). Esteem and ignorance. *Journal of Economic Behavior & Organization*, *63*, 373-383.

Crawford, V. P. (1995). Adaptive Dynamics in Coordination Games. *Econometrica, 63(1),* 103-143.

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

Deci, E. L., (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18,* 105–115.

Deci, E. L., Koestner, R., and Ryan, R. M. (1999). Meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125,* 627–68.

Devetag, G. (2003). Coordination and Information in Critical Mass Games: An Experimental Study. *Experimental Economics, 6*, 53-73.

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.

Devetag, G., and Warglien, M. (2003). Games and Phone Numbers: Do Short term Memory Bounds Affect Strategic Behavior? *Journal of Economic Psychology, 24,* 189-202.

Devetag, G., and Warglien, M. (2007). Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. *Games and Economic Behavior, 62(2),* 364-382.

Dickinson, D. L. (2001). The Carrot vs. the Stick in Work Team Motivation. *Experimental Economics*, *4(1)*, 107-124.

Dickinson, D. L., and Isaac, R.M. (1998). Absolute and Relative Rewards for Individuals in Team Production. *Managerial and Decision Economics*, *19*, 299-310.

Dickinson, D. L., and Villeval, M.C., (2008). Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, *63(1)*, 56-76.

Dipboye, R. L., and Flanagan, M. F. (1979). Are findings in the field more generalizable than in the laboratory? *American Psychologist, 34 (2),* 141-150.

Dobbins, G. H., Lane, I. M., and Steiner, D. D. (1988). A note on the role of laboratory methodologies in applied behavioral research: Don't throw the baby out with the bath water. *Journal of Organizational Behavior, 9*, 28 1-286.

Dohmen, T., and Falk, A. (forth.) Performance Pay and Multi-dimensional Sorting - Productivity, Preferences and Gender. *American Economic Review*.

Dougherty, D. (1992). Interpretative barriers to successful product innovation in large firms. *Organization Science*, *3(2),* 179-202.

Duffy, J., and Kornienko, T. (2005). Does Competition Affect Giving? An Experimental Study. Working Paper Department of Economics, University of Pittsburgh.

Dufwenberg, M., & Gneezy, U. (2005). Gender and coordination. In A. Rapoport, & R. Zwick (Eds.), *Experimental business research* (Vol. 3, pp. 253–262). Boston: Kluwer Academic.

Dugar, S. (2008). The effects of non-monetary sanctions and non-monetary rewards on coordination: Experimental evidence from the minimum action game. Unpublished manuscript. Available at: http://ssrn.com/abstract=897100 .

Eberlein. M., and Przemeck, J. (2006). Solidarity and Performance Differences. Mimeo.

Eckel, C., and Grossman, P. J. (2005). Managing Diversity by Creating Team Identity. *Journal of Economic Behavior and Organization*, *58(3)*, 371–92.

Ehrahrt, K-M, and Keser, C. (1999). Mobility and Cooperation: On the run. Mimeo.

Eisenberger, R., and Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologis,51,* 1153–66.

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

Erev, I., an Rapoport, A. (1998). Coordination, "Magic," and Reinforcement Learning in a Market Entry Game. *Games and Economic Behavior, 23(2),* 146-175.

Erev, I., Bornstein, G., and Galili, G. (1993). Constructive intergroup competition as a solution to the free rider problem: A field experiment. *Journal of Experimental Social Psychology, 29*,463-478.

Eriksson, T., Teyssier, S., and Villeval, M-C. (2006). Effort self-selection and the efficiency of tournaments. GATE Working Papers, 06-03.

Fahr, R., and Irlenbush, B. (2000). Fairness as a constraint on trust and reciprocity. *Economics Letters*, *66*, 275-282.

Falk, A., and Fehr, E. (2003). Why labour market experiment?. *Labour Economics*, *10*, 399-406.

Falk, A., and Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, *24 (1)*, 39-57.

Falk, A., Huffman, D., and Sunde, U. (2006). Self-confidence and search. IZA Discussion Paper, No 2525.

Faraj, S., and Xiao, Y. (2006). Coordination in fast-response organizations. *Management Science, 52,* 1155–1189.

Farrelll, J., and Rabin, M. (1996). Cheap Talk. *The Journal of Economic Perspectives, 10 (3),* 103-118.

Fatas, E., Neugebauer, T., and Perote, J. (2006). Within team competition in the minimum coordination game. 2006 ,*Pacific Economic Review, 11(2),* 247-266.

Feiler, L, and Camerer, C. (forth.) Code creation in endogenous merger experiments. *Economic Inquiry*.

Feri, F., Irlenbusch, B., and Sutter, M. (forth.). Efficiency Gains from Team-Based Coordination – Large-Scale Experimental Evidence. *American Economic Review*.

Foss, N. J. (2001). Leadership, Beliefs, and Coordination: An Explorative Discussion. *Industrial and Corporate Change*, *10*, 2, 357-388.

Franck, E., and Nüesch, S. (2008). The Effect of Talent Disparity on Team Performance in Soccer. Working Papers No. 0087, University of Zurich, Institute for Strategy and Business Economics (ISU).

Galbraith, J. R. (1973). *Designing Complex Organizations*. Reading, MA: Addison-Wesley.

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

Gneezy, U. (2002). Does high wage lead to high profits? An experimental study of reciprocity using real effort. Working Paper, University of Chicago GSB, Available at: http://www.chicagocdr.org/cdrpubs/pdf_index/cdr_519.pdf.

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.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments Gender differences. *The Quarterly Journal of Economics*, *August*, 1049-1074.

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

Goeree, J. K., Offerman, T. J., and Schram, A. (2005). Using first-price auctions to sell heterogeneous licenses. Working Paper, University of Amsterdam.

Goette, L., Huffman, D., and Meier, A. (2006). The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups. *American Economic Review*, *96* (2), 212–216.

Gordon, M. E., Slade, L. A., and Schmitt, N. (1986). The "science of the sophomore" revisited: From conjecture to empiricism. *Academy of Management Review, 11*, 191-207.

Griffin, R., and Kacmar, K. M. (1991) Laboratory Research in Management: Misconceptions and Missed Opportunities. *Journal of Organizational Behavior*, *12(4),* 301-311.

Gürtler, O., and Harbring, C. (2007). Feedback in tournaments under commitment problems: Theory and experimental evidence. IZA DP No. 3111.

Güth, W., Ploner, M., and Regner, T. (2008). Determinants of In-group Bias: Group Affiliation or Guilt-aversion? Jena Economic Research Paper, 2008-046.

Gurguc, Z. (2008). Democracy to promote coordination: An experimental approach. Mimeo.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2004). Diversity and Productivity in Production Teams. Mimeo.

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

Hargreaves Heap, S. P., and Zizzo, D. J. (2009). The Value of Groups. *American Economic Review*, *99(1),* 295–323.

Harrison, G. W., and List, J. A. (2004). Field Experiments. *Journal of Economic Literature, 42(4),* 1009-1055.

Harsanyi, J., and Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge: MIT Press.

Heat, C., and Staudenmayer, N. (2000). Coordination neglects: how lay theories of organizing complicate coordination in organizations. *Research in Organizational Behaviour*, *22*, 155-193.

Heinemann, F., Nagel, R., and Ockenfels, P. (2004). *Measuring strategic uncertainty in coordination games*. Mimeo.

Henning-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.

Hertwig, R., and Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences*, *24*, 383-451.

Heyman, J., and Ariely, D.(2004). Effort for Payment: A Tale of Two Markets. *Psychological Science, 15(11),* 787-793.

Hofstede, G. (1980). Motivation, leadership, and organization: Do American theories apply abroad? Organizational Dynamics, 9(1), 42-63.

Hogarth, R. M.(2005). The challenge of representative design in psychology and economics. Journal of Economic Methodology, 12(2), 253-263.

Hutchins, E. (1995). How a cockpit remember its speed. *Cognitive Science*, *19*, 265-288.

Irlenbusch, B., and Ruchala, G.K. (2008). Relative rewards within team-based compensation. *Labour Economics*, *15 (2)*, 141-167.

Ivanova-Stenzel, R., and Kübler, D. (2005). Courtesy and Idleness: Gender Differences in Team Work and Team Competition. Working Paper, SFB 649, Discussion Paper 2005-49.

James, W. (1890). *The Principles of Psychology*, Cleveland, OH: World Publishing.

Josephs, R., Larrick, R. P., Steele, C. M, and Nisbett, R. E. (1996). Protecting the self from the negative consequences of risky decisions. *Journal of Personality and Social Psychology, 62*, 26-37.

Kagel, J. H., and Roth, A. E. (2000). The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment, *Quarterly Journal of Economics*, 201-235.

Kareev, Y., and Avrahami, J. (2007). Choosing between adaptive agents: Some unexpected implications of level of scrutiny. *Psychological Science, 18(7),* 636–641.

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

Knez, M., and 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*, 194-216.

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

Kogan, S., Kwasnica, A. M, and Weber, R. A. (2008). Coordination in the presences of asset markets. Mimeo.

Konow, J. (2000). Fair shares Accountability and cognitive dissonance in allocation decision. *American Economic Review*, *90(4)*, 1072-1091.

Koszegi, B. (2006). Ego utility, overconfidence and task choice. *Journal of the European Economic Association*, *4(4),* 673–707.

Kube, S., Maréchal, M. A., and Puppe. C. (2006). Putting Reciprocity to Work - Positive versus Negative Responses in the Field. Working Paper, University of St. Gallen Department of Economics 2006-27.

Kube, S., Maréchal, M. A., and Puppe. C. (2008). The Currency of Reciprocity - Gift-Exchange in the Workplace. IEW - Working Papers No. 377, Institute for Empirical Research in Economics – IEW.

Kuhnen, C.M., and Tymula, A. (2009). Rank expectations, feedback and social hierarchies, Mimeo.

Kumru, C., and Vesterlund, L. (2005). The Effect of Status on Voluntary Contribution. Working paper, Department of Economics, University of Pittsburgh.

Lane, A. M., Jones, L., and Stevens, M. J. (2002). Coping with Failure: The Effects of Self-Esteem and Coping on Changes in Self-Efficacy*, Journal of Sport Behavior*, 25.

Larsson, R., and Finkelstein, S. (1999) Integrating Strategic, Organizational, and Human Resource Perspectives on Mergers and Acquisitions: A Case Survey of Synergy. *Organization Science, 10(1),* 1-26.

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

Lazear E. P., Malmendier, U., and Weber, R. A. (2006). Sorting in Experiments with Application to Social Preferences. Working paper.

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.

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

Levy-Garboua, L., Masclet, D., and Montmarquette, C. (2006). A Micro-Foundation for the Laffer Curve in a Real Effort Experiment. Working Paper, CIRANO 2006s-03.

Lewis R. E., and Heckman R. J., (2006). Talent management: a critical review. *Human Resource Management Review*, *16*, 139-154.

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

Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioural economics. *The Economic Journal 109,* F25–F34.

Malone, T. W. (1988). What is coordination theory? Working Paper No. 2051-88. Cambridge, MA: MIT Sloan School of Management.

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*., 357-370, October 07-10, 1990, Los Angeles, California, United States.

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.

Manzini, P., Sadrieh, A., and Vriend, N. J. (2009). On Smiles, Winks and Handshakes as Coordination Devices. *The Economic Journal*, *119*, 826-854,

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

Mc Kelvey, R. D., and Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior, 10,* 6-38.

McAllister, D. J., and Bigley, G. A. (2002). Work Context and the Definition of Self: How Organizational Care Influences Organization- Based Self-Esteem. *The Academy of Management Journal*, *45(5)*, 894-904.

McDaniel, T. M., and Rutstrom, E. E. (2002). Decision Making Costs and Problem Solving Performance, *Experimental Economics, 4*, 145-161.

Mintzberg, H. (1977). Policy as a field of management theory. *Academy of Management Review, 2*, 88-103.

Montmarquette, C., Rullière, J. L., Villeval, M. C., and Zeiliger, R. (2004). Redesigning Teams and Incentives in a Merger: An Experiment with Managers and Students. *Management Science*, *50(10)*, 1379-1389.

Munyan, L., and Camerer, C. (2005). Code creation in endogenous merger experiments. Mimeo.

Myung, N. (2008). Improving coordination and cooperation through competition. Mimeo.

Nalbantian, H. R., and Schotter, A. (1997) Productivity under Group Incentives: An Experimental Study. *American Economic Review, 87*, 314-41.

Nanda, A. (1999). Implementing organizational change. In M. Pina, E. Cunha, & C. A. Marques (Eds.), *Readings in organizational science: Organizational change in a changing Context*. Lisbon: ISPA.

Nelson, R. R., and Winter, S. G. (1982). *An Evolutionary Theory of Economic Change.* Oxford University Press.

Niederle, M., and Vesterlund, L. (2007). Do women shy away from competition? Do man Compete too much?. *The Quarterly Journal of Economics*, *August*, 1067-1101.

Okhuysen, G. A., and Bechky, B. A. (2009) Coordination in Organizations: An Integrative Perspective. *The Academy of Management Annals, 3(1),* 463-502.

Orrison, A., Schotter, A., and Weigelt, K. (2004). Multiperson Tournamants: An Experimental Examination. *Management Science*, *50*, 268-279.

Ortmann, A., and Gigerenzer, G. (2000). Reasoning in Economics and Psychology: Why Social Context Matters. In: Streit, M.E., Mummert, U. and Kiwit, D. eds. *Cognition, Rationality, and Institutions*. Berlin, Springer-Verlag.

Ottone, S., and Ponzano, F. (2008). How people perceive the welfare state. A real effort experiment. AL.EX Series, Paper No. 11.

Palacios-Huerta, I. (2003). Learning to open Monty Hall.s doors. *Experimental Economics, 6,* 235-251.

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

Pierce, J. L., Gardner, D. G., Dunham, R. B., and Cummings, L. L. (1993). Moderation by organization-based self- esteem of role condition-employee response relationships. *Academy of Management Journal*, *36*, 271-288.

Plott, C. R. (1982). Industrial organization theory and experimental economics. *Journal of Economic Literature, 20,* 1485–527.

Pokrny, K. (2008). Pay—but do not pay too much: An experimental study on the impact of incentives. *Journal of Economic Behavior & Organization, 66*, 251–264.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature, 37*, 7-63.

Radosveta Ivanova-Stenzel & Dorothea Kübler, 2005. "Courtesy and Idleness: Gender Differences in Team Work and Team Competition," IZA Discussion Papers 1768, Institute for the Study of Labor (IZA).

Ready, D. A., Hill, L. A., and Conger, J. A. (2008). Winning the race for talent in emerging markets. *Harvard Business Review*, *November*, 63-70.

Rick, S., Weber, R. A, and Camerer, C. F. (2007). Knowledge transfer in simple laboratory firms: the role of tacit vs. explicit knowledge. Mimeo.

Rick, S., Weber, R. A., and Camerer., C. F. (2007). The effects of organizational structure and codes on the performance of laboratory firms. Mimeo.

Rosenbaum, M. E., Moore, D. I., Cotton, J. L., Cook, M. S., Hieser, R. A., Shovar, M. N., and Gray, M. J. (1980). Group Productivity and Process: Pure and Mixed Reward Structures and Task Interdependence. *Journal of Personality and Social Psychology*, *39*, 626-642.

Roth, A. E. (1995a) Introduction to experimental economics. in Kagel, J., and Roth, A. (eds) *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University, 3–110.

Roth, A. E. (2002). The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics. *Econometrica*, *70(4)*, 1341-1378.

Rydval, O., and Ortmann, A. (2004). How financial incentives and cognitive abilities affect task performance in laboratory settings: an illustration. *Economics Letters, 85(3),* 315-320.

Salancik, G. (1977). Commitment and the Control of Organizational Behavior and Beliefs, in B. Staw and G. Salancik (eds.) *New Directions in Organizational Behavior*, Chicago: St. Clair Press.

Schein, H. E. (1985). *Organizational culture and leadership*. San Francisco: Jossey-Bass, 2nd print.

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

Schotter, A., and Weigelt, K. (1992) Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results., *Quarterly Journal of Economics*, *107*, 511-539.

Schram, A. (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology,12(2)*, 225-237.

Schwenk, C. R., (1982). Why Sacrifice Rigour for Relevance? A Proposal for Combining Laboratory and Field Research in Strategic Management. *Strategic Management Journal*, *3(3)*, 213-225.

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.

Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

Shih, M., Pittinsky, T. L., and Ambady, N. (1999). Stereotype Susceptibility: Identity Salience and Shifts in Quantitative Performance. *Psychological Science, 10 (1),* 81–84.

Simon, H. A. (1945). *Administrative behavior: A study of decision-making processes inadministrative organizations.* Thousand Oaks, CA: Free Press.

Simon, H. A. (1947). *Administrative Behavior*. New York: McMillan.

Sloof, R., and van Praag, M. (2007). Performance Measurement, Expectancy and Agency Theory: An Experimental Study. Working Paper, IZA DP No 3064.

Smith, V. L. (1982). Microeconomic systems as an experimental science. American *Economic Review, 72,* 923–55.

Sutter, M. (2006). Endogenous versus Exogenous Allocation of Prizes in Teams - Theory and Experimental Evidence. *Labour Economics*, *13(5),* 519-549.

Sutter, M. (2009). Individual behavior and group membership: Comment. *American Economic Review*, *99*, 2247-2257.

Sutter, M., and Weck-Hannemann, H. (2003) Taxation and the veil of ignorance a real effort experiment on the Laffer curve. *Public Choice, 115*, 217-240.

Tajfel, H., and Turner, J. (1979). An Integrative Theory of Intergroup Conflict. In Stephen Worchel and William Austin, eds., *The Social Psychology of Intergroup Relations*, Monterey, CA: Brooks/Cole.

Thompson, J. D. (1967). *Organizations in Action*. New York: McGraw-Hill.

Van Dijk, F., Sonnemans, J., and Van Winden, F. (2001). Incentive Systems 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 Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, *106*, 885–911.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1993). Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication. *Games and Economic Behavior*, *5*, 485–504.

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

Van Huyck, J. B., Gillette, A., and Battalio, R. C. (1992). Credible assignments in coordination games. *Games and Economic Behavior*, *4*, 606–626.

Wageman, R. (1995): Interdependence and Group Effectiveness. *Administrative Science Quarterly*, *40*, 145-180.

Weaver, M.J. (1999). Beyond the ropes: Guidelines for selection experiential training. *Corporate University Review*.

Weber, R. A. (2000). Organizational coordination: A game-theoretic view. Mimeo.

Weber, R. A. (2005). Managing Growth to Achieve Efficient Coordination: Theory and Experimental Evidence. Mimeo.

Weber, R. A., and Camerer, C. (2003). Cultural Conflict and Merger Failure: an Experimental Approach. *Management Science*, *49(4)*, 400-415.

Weber, R. A., and Rick., S. (forth.). Meaningful learning and transfer of learning in games played repeatedly without feedback. *Games and Economic Behavior*.

Weber, R., Camerer, C., Rottenstreich, Y., and Knez, K. (2001). The Illusion Of Leadership: Misattribution Of Causes In Coordination Games. *Organization Science*, *12*, 582-598.

Weber, R., Rick, S., and Camerer, C. (2004). The effects of organizational structure and codes on the performance of laboratory 'firms'. Mimeo.

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

Weick, K. E. (1969). Laboratory Organizations and Unnoticed Causes. *Administrative Science Quarterly*, *14(2),* 294-303.

Weick, K. E. (1977). Laboratory Experimentation with Organizations: A Reappraisal. *The Academy of Management Review*, *2(1),* 123-128.

Weick, K. E. (1993). The Collapse of Sensemaking in Organizations: The Mann Gulch Disaster. *Administrative Science Quarterly*, *December*,628-652.

Weick, K. E., and Gilfillan, D. P. (1971). Fate of arbitrary traditions in a laboratory microculture. *Journal of Personality and Social Psychology*, *17*, 179-191.

Weick, K. E., and Roberts, K. H. (1993). Collective Mind in Organizations: Heedful Interrelating on Flight Decks. *Administrative Science Quarterly, 38*, 357-381.

Weigelt, K., Dukerich, J., and Schotter, A. (1989). Reactions to Discrimination in an Incentive Pay Compensation Scheme: A Game-Theoretic Approach. *Organizational Behavior and Human Decision Processes, 44*, 26-44.

Xi Kuang, J., Weber, R. A., and Dana, J. (2007). How effective is advice from interested parties? An experimental test using a pure coordination game. *Journal of Economic Behavior and Organization, 62(4),* 591-604.