HAL open science

Redesigning Teams and Incentives in a merger. An Experiment with Managers and Students

Claude Montmarquette, Jean-Louis Rullière, Marie Claire Villeval, Romain Zeiliger

▶ **To cite this version:**

Claude Montmarquette, Jean-Louis Rullière, Marie Claire Villeval, Romain Zeiliger. Redesigning Teams and Incentives in a merger. An Experiment with Managers and Students. Management Science, 2004, 50 (10), pp.1379-1389. halshs-00161723

HAL Id: halshs-00161723

https://shs.hal.science/halshs-00161723v1

Submitted on 17 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Redesigning Teams and Incentives in a Merger:

An Experiment with Managers and Students

Claude Montmarquette[a], Jean-Louis Rullière[b],
Marie-Claire Villeval[bc], Romain Zeiliger[b]

Published in Management Science (2004), 50(10), 1379-1389

[a] CIRANO and University of Montreal, 2020 University Street, Montreal, (Quebec), Canada, H3A 2A5. Tel: (514)985-4015, Fax: (514)985-4039, Claude.Montmarquette@cirano.qc.ca

[b]GATE, CNRS and University of Lumiere Lyon 2, 93 chemin des Mouilles, 69130 Ecully, France. Tel: +33 472 86 60 60, Fax: +33 472 86 60 90, rulliere@gate.cnrs.fr, villeval@gate.cnrs.fr, zeiliger@gate.cnrs.fr

[c] IZA Research Fellow, Bonn

We are greatly indebted to AVENTIS Pharma for its collaboration in this research. We are also grateful to Laurent Volpi, Muriel Prat and Jean-Benoît Grégoire-Rousseau for their help in conducting the research. Our thanks to David Dickinson and Edward P. Lazear for their useful comments and suggestions. The final version of the paper has benefited from comments and criticisms provided by the associate editor and the referees. Financial support from the Region of Rhône-Alpes and of the LUBC3E (Laboratoire Universitaire Bell) is acknowledged.

Abstract: After a merger, company officials face the challenge of making compensation schemes uniform and of redesigning teams with managers from companies with different incentives, work habits and recruiting methods. In this paper, we investigate the relationship between executive pay and performance after a merger by dissociating the respective influence of shifts, which occur in both compensation incentives and team composition. The results of a real task experiment conducted with managers within a large pharmaceutical company not only show that changes in compensation incentives affect performance but also suggest that the sorting effect of incentives in the previous companies impact cooperation and efficiency after the merger. Replicating this experiment with students showed differences in strategy rather than in substance between the two groups of subjects with managers appearing performance driven while students are more cost driven.

JEL-Code: C81, C92, J33, M52

Keywords: Executive and team-based compensation, subject pool effects, real task experiment, incentives, sorting, mergers.

.

1. Introduction

There is strong evidence of the existence of heterogeneity among the compensation packages applied by firms within the same industry. It is not surprising to find that after a merger, difficulties can arise because of the different compensation policies of the newly merged firms and that new consolidated policies need to be designed. Furthermore, downsizing and the reorganization of production entail a reshuffling of teams, which affect executives from the companies involved in the merger. In order to promote social cohesion, mergers usually lead to the harmonization of company statutes so that all executives are paid according to the same compensation schemes. But within new teams comprised of executives of the merged companies, performance also depends on the willingness of individuals to cooperate with others. This willingness to cooperate may be affected by the heterogeneity of past compensation practices, work habits and non-market interactions.

Many pitfalls can hamper an empirical analysis of the relationship between new executive pay packages and executive performance after a merger. The impact of new compensation packages may differ from one employee to another depending on the short, medium and long-term influence of preceding modes of compensation (see Nalbantian and Schotter, 1997). Thus, assessing the impact of new compensation schemes on executive performance after a merger requires controlling for the possibility of the long-term impacts of the compensation packages used before the merger. Furthermore, unbiased estimates of the relationship between pay and performance require disentangling the effects of shifts in direct incentives from the effects of the emergence of a new group culture founded on a variety of previous corporate cultures (Kreps, 1990). Individuals coming from a variety of previous cultures, with different norms of fairness and social comparison, could be

expected to behave differently in the new company. Previous cultures can affect the efficiency of a new unified compensation policy particularly in the short run. Experimental methods can help in circumventing part of these difficulties through the comparison of various treatments in a controlled environment. This point has been successfully made in the context of a merger by Weber and Camerer (2003) who show that performance decreases following the merger of two laboratory firms because of failure in coordination.

In this paper, we design an experiment to analyze the relationship between executive compensation schemes and performance after a merger. Our laboratory experiment was conducted with 36 managers of a large pharmaceutical company created by the recent merger of two companies, one French and one German. It was replicated with 72 students in Lyon and Montreal. Thus, our paper also contributes to the literature on subject pool effects by comparing decisions made by student-subjects and manager-subjects. Studies by Dyer, Kagel and Levin (1989), and Carpenter, Burks and Verhoogen (2003) have underlined differences (risk attitudes, fairness) and similarities (winner's curse) across subject pools. Hannan, Kagel and Moser (2002) show that MBA effort levels are greater than those of undergraduate students and Fehr and List (2002) observe that CEO's exhibit more trustful behavior. How can one explain that in some games expertise seems to influence behavior whereas in others it has no significant influence? Cooper, Kagel, Lo and Gu (1999) emphasize the relationship between context and expertise. Expertise can improve the relative efficiency of manager-subjects in experiments only when they are able to recognize the similarity between the laboratory context and their field experience. Here, we could expect that conducting the experiment within the company context and

performing an abstract task which reproduces characteristics of executive teamwork, should induce a greater level of performance from managers. Our paper is also innovative since it examines these effects in a team context with individual and team rewards.

The task required of participants consisted of searching for the highest value of a multiple-peaked function in a two-dimensional space. This task has a cognitive component since intense concentration is required because of uncertainty and time pressure and there is a monetary cost linked to the chosen speed of progression. This real-task experiment adds to the limited number of experimental papers (Dickinson, 1999, Sillamaa, 1999, van Dijk, Sonnemans and van Winden, 2001, Falk and Ichino, 2003, Gneezy, Niederle and Rustichini, 2003), which study rewards and behavior in a real work setting. Many experiments, which require subjects to choose an effort level not related to a real effort confirm that monetary incentives do matter, but assume the equivalence between intention of contribution and effort, and between disutility of effort and money. Our experiment also differs from naturally occurring experiments (for example, Lazear, 2000) or field experiments (Erev, Bornstein and Galili, 1993) in which subjects perform a task in a real work environment. In our study, manager-subjects undertake tasks, which reproduce aspects of a manager's job under a familiar structure of incentives. They make their decisions in an artificial environment of anonymous interactions, according to instructions using neutral wording and without field referents.

Our design involves two parts: in the first part, teams are homogeneous and are paid according to the rules in effect before the merger of the pharmaceutical firms. In the second part, the teams remain homogenous or are formed randomly with participants originating

from each of the two merged companies or from both schools; all are paid according to the rules in use after the merger. These treatments are aimed at disentangling the effect of the shift in team composition and the impact of the shift in incentive schemes.

Our main results indicate that there is a pure effect on performance of the shift in incentives after the merger. They show that the past matters in as much as some managers reduce their effort when they are potentially mixed with managers from the other incoming firm. This may be the result of sorting effect of previous incentives schemes: paying executives under different rules has probably contributed to the creation of attitudes towards cooperation in teams. Lastly, we find evidence that manager-subjects and student-subjects differ more in strategy than in substance, with managers being oriented more to the maximization of performance while students focus more on cost minimization.

In Section 2, the design of the experiment is outlined. Section 3 presents the experimental procedures. The empirical results are discussed in Section 4. In section 5, we conclude.

2. Experimental Design

The experiment consists of two parts. It reconstitutes in the first part the pay structure of the firms before they merge and in the second part, the pay structure prevailing in the merger. The nature of the task to be performed during the experiment remains the same, thus allowing an analysis of the consequences on performance of both changes in the payment structure and of team composition. In this section, we present the design of the task, the structure of payment schemes, the treatments and information conditions.

2.1. Task design and behavioral heuristics

One original aspect of this experiment lies in the design of a task, which mimics some aspects of the content of a manager's job (concentration, variability, adjustment of means to

targets and ability to cope with uncertainty under time pressure). The challenge was to be able to discriminate the impact of a cognitive effort on the outcome from that of ability. The task consists in searching for the highest value of an increasing function built in a two-dimensional space defined vertically by height (H) and horizontally by distance (D) from the origin, with $H \in [0,100], D \in [0,300]$ and with $H^{Max} = f(D)$. It is common knowledge that the curve can have single or multiple peaks. When the period starts, the box in which the curve will appear is fully black. During a one-minute period, the subject progressively uncovers the curve on his computer screen starting at the origin, by clicking a button.[i] The subject progresses by discrete steps on the horizontal axis. She can stop her progression at any moment. The outcome achieved by a subject in a period is given by the maximum height reached on the curve, which depends notably on the number of moves.

The monetary cost associated with the task is represented by the choice of the speed of progression, i.e. the work pace. Two buttons are available to the subjects: a “1-step button” used to uncover the curves at a regular speed and is cost free; a “2-step button” that doubles the speed but which costs 0.4 point. The subject can switch speeds at will. This design allows a control over the subject’s tasks cost and makes possible an analysis of efficiency. It is impossible to reach maximum height during the one-minute period allowed by using the regular speed only, with the exception of one of the 13 randomly occurring curves. The subjects do not know this information. A subject's outcome depends on the use of two-step moves and on chance since there is uncertainty about the distances from the origins to the peaks. Therefore, the cognitive dimension of the task relates to the uncertainty about the shape of the curve, to time pressure and to the decision to use the 1-step or 2-step button.

Unlike van Dijk, Sonnemans and van Winden's (2001) real-task experiment, ours involves no algorithm enabling the discovery of a peak at minimum cost while under time constrains; no one benefits from previous learning. At each point on the curve, the subject cannot infer from the already uncovered part of the curve the slope of the curve at the next point. The subject cannot discover one single algorithm but must build heuristic grounded on a continuous trade-off between the speed to choose and the time left. For example, if a subject has already used many extra cost moves and the curve remains flat, paying each additional 2-step move requires a continuous trade-off between its marginal cost and its expected marginal revenue. Compared to a traditional experiment where effort values are chosen from a payoff table with a unique trade-off between cost and outcome, our task involves an unknown optimal way of reaching the maximum height at a minimum cost since subjects do not know the locations of peaks. The property of duality is not checked: the minimization of cost does not correspond to the maximization of the objective. The multiplicity of heuristics and uncertainty concerning the location of peaks prevent any theoretical prediction. [ii]

In short, there are two aspects to the task chosen in this experiment: the decision to use the costless or costly way to progress and the speed of progression, where both can be measured. However, the cognitive problem of how to reach a higher altitude cannot be measured. For the rest of the paper, the altitude reached is the observable outcome of the cognitive effort. This output and the cost of progression are the variables to be explained.

2.2. Payment schemes

Another point of originality lies in the design of compensation schemes that combine fixed pay and performance-pay, the latter involving both individual and team incentives. The

game involves teams consisting of three subjects who have to uncover the same curve. Subjects are not allowed to communicate with teammates and are not informed of their simultaneous progression. Individual earnings from a task in a given period are calculated from the sum of three elements whose amount and relative proportion depend on the stage of the game and on the treatment. Specifically, earnings are defined as: $\pi_{\alpha i} = F_\alpha + I_\alpha + T_\alpha$, with $\alpha = \{X, Y\}$. X and Y correspond to the two firms before they merge, and for simplicity X and Y subjects keep their respective labels after the merger in order to track their origin.

F_α is a fixed-wage earned by subject i when her individual outcome reaches a first threshold, $H_1^{\min}$, defined by the height reached. This threshold can always be achieved with no cost steps in the time allowed, but this information is not given to subjects. An employer would consider a distance achieved below this benchmark as professional misconduct. I_α is an individual bonus earned if i's outcome reaches a second threshold, $H_2^{\min}$ with $H_2^{\min} > H_1^{\min}$. T_α is a team reward obtained when the sum of individual outcomes by the team of 3 subjects reaches a third threshold $H_3^{\min}$ with $H_3^{\min} > 3H_2^{\min} > 3H_1^{\min}$. A subject may earn this reward even though she does not reach an output level greater than the level giving her the fixed wage or the individual bonus. This creates an incentive to free ride.

At each repetition of the game, a new curve is randomly drawn, whose shape determines the extent of uncertainty faced by the subjects. The analysis of performance and costs must control for the degree of difficulty of the curves. Because of the structure of the compensation package, the difficulty of a curve depends on the location of the various thresholds. An index of difficulty is calculated as : $d = (D_1)^2 + (D_2 - D_1)^2 + (D_3 - D_2)$ with

D_1 being the abscissa at the origin of the first threshold, D_2 the abscissa of the second threshold and D_3 the abscissa of the maximum height. The more distant the first threshold is from the origin and the greater the distance between the first and the second thresholds, the more difficult it becomes to reach additional rewards.

2.3. Experimental treatments and information conditions

The experiment aims at identifying the separate influences of changes in incentives and in team composition. To measure the impact of changes in payment schemes, a session was designed having two parts of 10 periods each, with a random order of presentation of 13 payoff curves. In the first part, used as a benchmark, we reproduce initial payment schemes that were used before the merger; in the second part the payment scheme in use after the merger is applied. In the first part, we team X and Y subjects separately, each playing under the payment scheme used in their initial company. Members of X teams may receive a fixed wage, an individual bonus and a team reward. Earnings for members of Y teams are derived from a fixed wage and a team reward only (see Table 1). The proportion of potential total earnings from the fixed wage is higher for Y subjects than for X subjects, but the same performance is required from all subjects to trigger their fixed payment and team reward. In the second part of the session, the payment scheme is the one used after the merger and is the same for all subjects. It includes a fixed wage, an individual bonus and a team reward. Compared to the first part, Y subjects may now receive an individual bonus and X subjects have seen an increase in their fixed wage. In avoiding a decrease in the absolute level of any pay component of earnings between the two parts of the session, the maximum earnings all subjects can obtain is now 120 ECU instead of 100.

To measure the impact of team composition on performance, a new matching treatment was introduced in two sessions. This treatment is referred as the mixed-treatment, where teams of 3 randomly associated subjects are formed potentially involving both X and Y subjects. The fixed-treatment used in the first part of all sessions, where X and Y subjects are teamed exclusively with X and Y subjects, respectively, is maintained in the second part in one session. The fixed-treatment serves as a benchmark against which the effect of team composition after the merger is tested. A stranger matching protocol is always used. In the second part of the session, only team composition is changed and not the size of the teams.[iii]

[Insert Table 1 about here]

All subjects, except those playing the fixed treatment, knew of the existence of two categories of subjects in equal numbers in the room, but they were kept unaware of the meaning of labels X and Y. Subjects learned their own identity by reading the instruction sheet and were notified that they would keep the same identity throughout the session. The instruction sheet for the first 10 periods mentioned that they were matched with two other subjects belonging to the same category as themselves and that the composition of groups would within the same category change with each new period. They were informed that they would never know the identity and the payoff of their successive teammates. Subjects knew the description of the task to be performed and the payoff structure applicable to their category. The participants were not given information about the payoff structure of the other category of subjects; were aware that the same task was to be achieved by the three members of their group and had no current information about the progression of their teammates on the curve. During play, their screen indicates the time currently left, the

cumulated cost of their 2-step decisions and the height of the different thresholds reached. At the end of each period, a historic table gave each subject feedback on their own outcome, the outcome (cumulative height) achieved by their team, total cost, amount of compensation obtained, by source, and total payoff net of costs.

In the second part of the session, subjects were informed about the payoff structure of the first part for each of the two categories. The new payoff structure common to all subjects was explained. In one session, an equal number of subjects from both categories continued to play under the fixed treatment for periods 11 to 20. In two sessions, subjects drawn in equal numbers from X and Y were informed that during the remaining 10 periods they could be matched randomly with members of the other category.[iv]

3. Experimental Procedures

The experiment was first conducted with managers of the pharmaceutical company about two years after the merger and it was later replicated with students in engineering and business from Lyon and Montreal.

36 volunteer executives were recruited by email with average annual earnings of 69,000 euros, consisting of 18 managers from each incoming firm. To limit uncontrolled peer group effects, sessions were designed such that the participants represented a large diversity of departments and came from different geographical locations.[v] The experiment was conducted at the company headquarter in Paris on the same day to limit the dissemination of information. In the first two sessions, 12 participants (6X, 6Y) were subjected to the mixed-treatment protocol. In a third session, 12 participants played under the fixed-treatment protocol. On average, a session lasted 75 minutes including practice periods. Transactions were conducted in Experimental Currency Units, with ECU convertible to

Euros at the rate 150 ECU = 4.5 €. A show-up fee of 8 € was added. On average, a subject earned 51.45 € (S.D.=3.75). Subjects were paid a few days later with vouchers.

Upon arrival, each manager had to register and was invited to draw a ticket from an envelope to assign him or her a computer. In fact, company specific envelopes were presented to subjects according to their originating company, but subjects were unaware of this allocation rule. Subjects discovered a set of written instructions for the first part of the session under their keyboard. As the payment schemes differed among X and Y participants, the experimenter did not read the instructions aloud.[vi] Instructions were phrased in neutral terms with queries answered in private. Three practice rounds were run before the experiment began. At the end of the first part, the game stopped and further instructions for the second part were distributed, without any questions allowed.

This experiment was replicated with 72 student subjects, in the experimental laboratories of GATE-Lyon and at the LUBC3E-CIRANO in Montreal. Three sessions were organized over two days. The two mixed-treatment sessions were played on-line, the REGATE software enabling the on-line reshuffling of groups between Montreal and Lyon subjects. During each mixed-treatment session, 12 French subjects and 12 Canadian subjects interacted as subjects X and Y without being informed that they were playing with subjects abroad. The third session (fixed-treatment) involved 12 students at each location. Sessions were conducted under the same conditions as sessions with manager-subjects, except that the ECU were convertible at the rate 150 ECU = 1 Euro for the French subjects and at the rate 150 ECU = 1.5 Canadian Dollars for the Canadian subjects. A show-up fee of 4 Euros or 6 Canadian Dollars was added. On average, a subject earned 12.89 € (S.D.=0.80 €). Subjects were immediately paid in cash in a separate room.

4. Experimental results

4.1. Overall statistics

In Figures 1 to 3, we present histograms of average output achieved, average rate of free-riders and average return to output level plotted against treatment, subject pool and the group of periods (1-10 or 11-20). Figure 1 shows that the reaction to incentives differs according to the subject pools: the average output level is higher for managers than for students throughout the experiment. For the managers, output increases in the second part of the experiment under the merged company's compensation package. This increase is particularly noticeable in the fixed treatment sessions, where the only change introduced relates to the increase in monetary incentives; when both incentives and team composition are changed, the increase in managers' average output is not significant.

[Insert Figures 1, 2, and 3 about here]

These different output levels may be related to differences in the proportion of free riders in the two pools. Because of the team reward in the payment scheme, our experiment can be connected with the Voluntary Contributions Mechanism in public good games. It is interesting to verify whether the co-existence of a fixed wage, an individual bonus and a team reward in the same compensation scheme changes the behavior towards cooperation as typically observed in these games. Transposing the strict definition of free riding in the VCM to our experiment means that a free rider is here defined in the following way. In part 1, X is qualified as a free rider in period t when her output is ≤ 60, i.e. when the subject does not contribute to the team outcome beyond the level that triggers her individual bonus. The same condition applies for part 2. In part 1, Y is qualified as a free-rider when her

output in period t is ≤ 40, i.e. when she does not contribute to the team outcome beyond the level that secures her fixed wage, and in part 2 when her output level is ≤ 60. In a typical public good game, whereas each subject has a dominant strategy not to contribute, the level of contribution is initially positive and declines with repetition (see Ledyard, 1995). Figure 2 shows the rate of free riders by subject-pool and by treatment for parts 1 and 2. It indicates that there is a low proportion of free-riders compared to traditional results of public good games, probably because of both the coexistence of individual and collective payments and the impossibility of calculating the marginal per capita return of investing in the team outcome.[vii] The proportion of free riders is lower among managers than among students in both parts.[viii] This proportion is lower when the threshold that triggers the greater individual reward is low. It increases in part 2 in both subject pools, particularly in the mixed sessions, and this cannot be explained by a restart effect at the 11th period of the game as will be shown later. This is in line with the declining cooperation over time usually found in public good experiments and could be compared with the differences observed between stranger and partner-matching protocols (see Keser and van Winden, 2000). Figure 3 displays the average return to output level, i.e. the average output achieved divided by the average cost level. On average, managers achieve a lower level of efficiency than students. The former perform a greater output but at a higher cost. In part 2, the average level of efficiency rises slightly in both subject pools.

However, overall statistics are uncontrolled for time, difficulty of the curves and individual effects. Regression analyses controlling for these dimensions are thus required to identify the determinants of two endogenous variables: output levels and costs. In Table 2, we present the definition and descriptive statistics of variables used in these regressions.

[Insert Table2 about here]

Exogenous variables are the period, the category of subjects, the mode of compensation, the composition of groups and an index of difficulty for each curve. The "lagged output achieved of the other group members" variable assesses whether subjects modulate their outputs to what their teammates did in the previous period. Interaction variables involving the Y subjects are created to test whether X and Y subjects behave differently. Coefficient estimates of the variable "mode of compensation in part 1" report the decisions of the X subjects in part 1 relative to their decisions under the mode of compensation in part 2 (element of the constant term). With the coefficients of the "mixed session" variable, we further distinguish the decisions of the X subjects in the mixed sessions of part 2 relative to the fixed sessions of part 2. The decisions of the Y subjects in part 1 are the sum of the coefficients of the variables "Y subject", "mode of compensation in part 1" and "Y subject and mode of compensation in part 1". This last variable is needed as the modes of compensation differ in part 1 between X and Y subjects. Summing-up, coefficient estimates of variables "Y subject", "mixed session" and "Y subject and mixed session" give the decisions of Y subjects in the mixed sessions in part 2. The coefficient of the "Y subject" variable shows the decisions of "Y subjects" in the fixed sessions in part 2.

The index of difficulty and the period variables enter regressions with interacting variables and nonlinear forms. Instead of including a dummy for each period to control for time effects, we distinguish the variable "logarithm of the period" from the "periods 1 to 20" variable in order to separate two potentially adverse effects. While the former may account for elements such as boredom that could push output level down over time, the latter may account for learning of the task that could increase output over time. Lastly, demographic

variables such as gender and age were entered to control for their potential impact but since none proved to be significant, they were subsequently removed.

4.2 Econometric results

Column 1 of Tables 3 and 3a display the results for the students and managers, respectively, on the subjects' output achieved. Column 2 of both tables report the results on the determinants of cost levels. The cost incurred corresponds to the number of occasion the 2-steps button was used to perform the task. All econometric estimates are obtained with a linear one-way random effects model.[ix]

[Insert Tables 3 and 3a about here]

For the students considered as a benchmark, few variables significantly influence outputs, except periods and the difficulty of the task. In contrast, the cost levels are influenced by more elements. A change in monetary incentives in the second part of the experiment makes students react by increasing their cost levels but not their output. An increase in the output levels of their teammates during the preceding period incites them to increase their costs, particularly for the X subjects. Lastly, costs decline over time and the relationship between the difficulty of the curves and the costs indicates a reverse U-shape.

In contrast, we observe for managers a significant and substantial increase in outputs by both X and Y subjects in part 2 relative to part 1. The change in incentives after the merger increases outputs by almost 12 points. The change in the composition of teams exerts no significant influence on output. Note however the negative but statistically insignificant coefficient of the interaction variable "Y subject and Mixed": Y subjects, knowing that they may be interacting with X subjects, have a tendency to lower their output levels. Despite

the stranger-matching protocol, Y subjects are also influenced by the behavior of their teammates in the preceding period: an increase in teammate outputs motivates subjects to increase their own output. It should also be noted that, if we observe a negative first period effect on output achieved, there is no restart effect at the 11th period. The difficulty of the curves and the number of periods affect outputs in a nonlinear way.

The change in composition of the teams in part 2 has no significant impact on effort decisions, however, it exerts a determinant impact on the cost levels chosen by the subjects. Ceteris paribus, X subjects increase their costs in a mixed session by 6.23 units relative to a fixed session, while Y subjects substantially reduce theirs by 8.38 units when they know they may be teamed with X subjects. [x] The relationship between the difficulty of the curves and the costs indicates a reverse U-shape as with the students. Lastly, there is a positive first period effect on cost levels, but costs decline more linearly as the experiment evolves.

4.3 Discussion

This section will focus on the reactions to the shifts in incentives, in team composition and in the difficulty of the task. Our results provide clear evidence of differences in reactions to a change in monetary incentives between the subject-pools. The new compensation scheme in the second part does increase the output levels of X and Y managers. In contrast, the same monetary incentives, controlling for other variables, incite students to increase not their outputs but their costs to accomplish the different tasks imposed on them. Among managers, Y participants substantially decrease their costs in part 2. Another difference is the reaction to lagged output levels of the other group members. Y managers increase their

outputs if the output levels of their previous teammates increase. Student participants modify their cost levels to the previous output levels of their group members.

These results show that the change in incentives influences both subject-pools, with managers appearing to be objective driven while students are cost driven. This is not that surprising as managers in their professional life are evaluated, remunerated and promoted by meeting their objectives whatever the cost they must incur to do so (long working hours for example). For most students, a large part of the return of their academic effort is from lowering their cost (time devoted to studying a specific matter) in order to obtain good grades. Note that we were able to observe this kind of results because we used a real costly task. If the student-subjects appear to be more money maximizing than managers in the laboratory, this cannot be attributed to the differences in the opportunity costs to participate, since this was taken into account in the conversion rates.

The changing composition of teams also influences the behavior of subjects. Within the same category (X or Y), most subjects are influenced by the behavior of their preceding teammates although teams are reshuffled each period. This might suggest that the subjects refer to their category as a whole. However, the linear one-way effects models recognized that subjects are heterogeneous; our results suggest that subject-pools and categories are also heterogeneous. Both X and Y student-subjects are influenced by the output of the subjects they were previously teamed with but they are not influenced by the merging between categories in the second part of the experiment; thus, they do not refer to a specific category but to the whole set of subjects. By contrast, after the merger, the decisions of the category of managers who are more sensitive to the outputs of previous teammates (the Y managers) change behavior knowing that they may be interacting with subjects from the

other category; they become less cooperative and they substantially reduce their costs and to a lesser degree their outputs. On the contrary, X managers who are not influenced by the behavior of previous teammates react positively to the merging of the categories in the same teams. Groups of reference are not the same across subject-pools.

In traditional experimental economics literature, reaction to others' decision is usually characterized by a reciprocity concept (see Fehr and Falk, 2002). Since our experiment is run with randomly re-matched subjects at each period, this can also be explained by learning and conditional cooperation: subjects learn the behavior of other subjects in the same category or in the same room and constantly update their beliefs. But more puzzling is the reaction to the mixing of categories after the merger. Three explanations could be evoked. A first explanation might be that the merger changes preferences by introducing new incentives. This explanation cannot help since we measure the impact of a shift in team composition other things equal, i.e. we control for changes in incentives. The psychological concept of "in-group/out-group" (Tajfel, Flament, Billig and Bundy, 1971) might affect the cooperative behavior of participants. This explanation cannot directly help here since students and managers do not react in the same way. If managers were able to transfer their experience of the merger into the laboratory whereas the students were not able to do so, this psychological effect might play a role for the managers. But it does not explain why X and Y managers did not react similarly. A more convincing explanation refers to a sorting effect of incentives as documented by Lazear (2000). The incoming companies that merged might differ in recruiting different profiles of managers, then developing different cultures, because of their various incentive schemes. The pre-existence of an individual bonus may contribute to focusing on one's own performance instead on the other's output.

The reaction to the difficulty of the task is similar for managers and students. The U-shape curve suggests that more difficult tasks may actually elicit, to some extent, more output. This job challenge effect is present even in the later stages of the experiment (see the "index of difficulty and period" crossed variable), whereas the logarithm of period exerts a negative effect on the output level (that could be interpreted as boredom or tiredness). This result is consistent with the psychological literature showing that challenging goals lead to higher performance than easy goals (Locke, Saari, Shaw and Latham, 1981). The job challenge may even be at the initiative of the subject herself (for example if she uses targets like reaching the second peak). The relationship between the difficulty of the curves and the costs supported by all participants indicates a reverse U-shape. If the task is too difficult, subjects increase their outputs but without resorting to costly 2-step moves. This result reinforces our preceding analysis: an increased difficulty does not discourage output under the condition that subjects can save on their costs. Lastly, there is a positive first period effect on the cost levels (significant for the managers), but costs decline more linearly as the experiment evolves. This is possibly due to a learning effect on the task, on the other's behavior, and on the best moment to use the costly moves. Students appear to learn more than managers as they play the game since they not only decrease their costs but also increase their outputs (see variable "Periods 1 to 20").[xi]

5. Summary and conclusion

Executive behavior with respect to performance, motivation and cooperation is a major element in the success or failure of a merger between companies. Traditionally, economists have suggested looking for an adapted compensation policy to facilitate cooperation within

groups of individuals coming from different corporate cultures. The aim of this paper is to check whether a harmonization of compensation packages is sufficient to motivate all managers to cooperate to the same extent. A laboratory experiment has been run involving managers of two large pharmaceutical companies that recently went through a merger. The experimental design has introduced various compensation schemes, including an incentive scheme combining individual and team incentives that were implemented in the context of a real task. As in most mergers, these manager-subjects have experienced the redesigning of both compensation schemes and team composition in their newly merged company. The protocol reproduced the pre- and post-merger situation in terms of both compensation and team composition. To complement this experiment, a replication with a subject pool consisting of French and Canadian students was conducted that can serve as a benchmark.

The results show that financial incentives do work in improving output achieved, in accordance with standard results (Prendergast, 1999). However, the unified incentives are not entirely sufficient to create cooperation among heterogeneous groups, as already experimentally observed (Meidinger, Rulliere and Villeval, 2003). Past matters. In contrast with Nalbantian and Schotter (1997), it matters more in terms of shifting team composition than in terms of incentives, since the change from the two pre-merger incentive schemes to the unified one increases output in the same proportion. Individuals coming from different corporate cultures, likely with different fairness norms and social comparison behavior, tend to react differently in the mixed treatment part of our experiment. This may result notably from a sorting effect, attributable to various selection policies in the originating firms: companies with different incentive policies probably attract different types of managers. This suggests that shifting team composition may limit, at least in the short run,

the efficiency of a new unified compensation policy, if not taken into account. Merging cultures requires more time than merging incentives and deserves special attention. Our experiment also shows that introducing complex tasks is not necessarily detrimental to performance and cooperation. The concept of job challenge is perhaps more important to soliciting greater effort among employees than is usually suggested in current literature.

Results from the student-subject pool differ in strategy more than in substance, allowing confirmation of the external validity of laboratory experiments. In contrast to the managers, students react to an increase in monetary incentives by accepting more costs to complete a given task rather than increasing their output levels. They are cost driven whereas managers appear to be objective-driven. This means that the managers are also more cooperative and free ride less than the student-subjects. Our results corroborate the interpretation of Cooper et al. (1999) in that when they are able to recognize the similarity between the laboratory context and their field experience, manager-subjects may choose different strategic options than inexperienced subjects. Moreover, it may indicate that if students are more inclined to minimize costs than experts in the laboratory, when one observes the existence of other-regarding preferences in traditional experiments involving student-subjects, one may deduce that this deviation from the equilibrium is likely more developed in real settings.

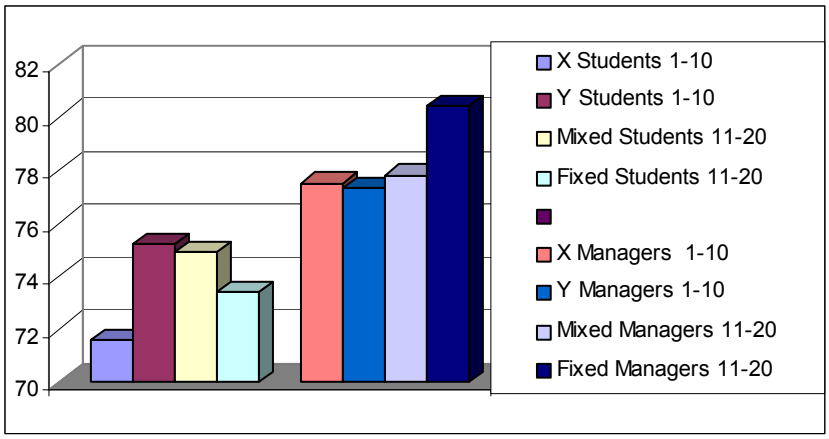

Figure 1. Average output by subject pool and by treatment

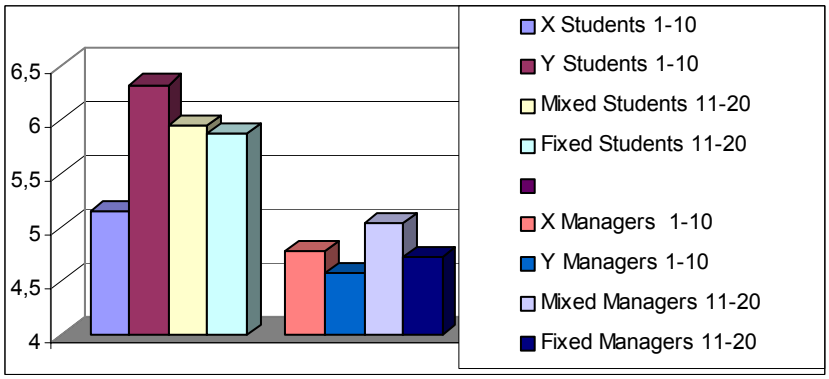

Figure 2. Average proportion of free riders by subject pool and by treatment

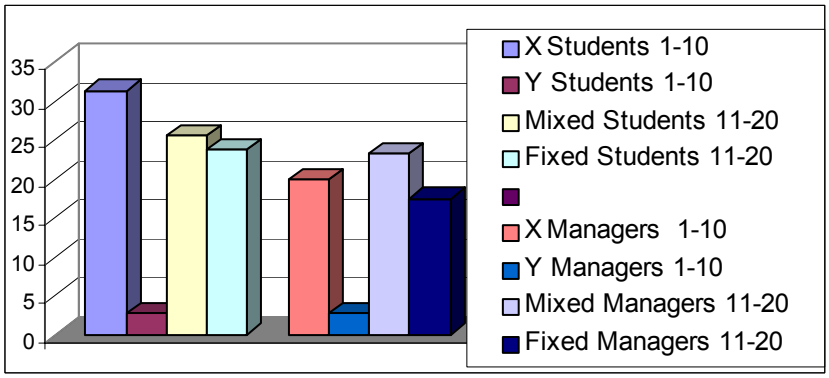

Figure 3. Average return to output by subject pool and by treatment

Note: The average indices of difficulty of the curves are the following: In the student-sessions, periods 1 to 10: 22 949; periods 11-20 in the mixed treatment: 24 047; periods 11-20 in the fixed sessions: 21 635. In the manager-sessions, periods 1 to 10: 22 949; periods 11-20 in the mixed treatment: 22 339; periods 11-20 in the fixed sessions: 21 635.

Table 1-- Payment schemes in ECU

Group Composition	Part 1 – All treatments		Part 2 – Fixed treatment		Part 2 – Mixed treatment
	Teams of X subjects	Teams of Y subjects	Teams of X subjects	Teams of Y subjects	Teams of X and Y subjects
Height reached					
$H_i < 40$	0	0	0	0	0
$40 \le H_i < 60$	40	60	60	60	60
$60 \le H_i \le 100$	60	60	80	80	80
$\sum_{1}^{3} H_i \ge 240$	100	100	120	120	120

Note: For the teams of X subjects the table should be read as follows. In the first part of the experiment, if a X subject realizes an outcome lower than the first threshold of 40, he receives no payoff. If he reaches a height between the first (40) and the second threshold (60), he receives only a fixed wage of 40 (F_X). If the subject reaches the second threshold, he receives a payoff of 60, consisting of the sum of the fixed wage and the individual bonus ($F_X + I_X$). If the subject's team reaches a cumulated height of 240, the subject receives a payoff of 100, corresponding to the sum of the fixed wage, the individual bonus and the team reward ($F_X + I_X + T_X$). Part 2 and the teams of Y subjects should be interpreted in a similar manner.

Table 2--Variables and descriptive statistics

Variable	Definition	Mean (Standard deviation)	
Endogenous variables		**Managers**	**Students**
Output achieved	Output levels reached by the subjects	78.04 (20.96)	73.9 (20.680)
Cost	Costs of outputs supported by the subjects	16.60 (13.19)	12.75 (11.39)
Exogenous Variables			
1st period effect	1 if the 1st period; 0 otherwise	0.05 0.95	0.05 0.95
11th period effect	1 if the 11th period ; 0 otherwise	0.05 0.95	0.05 0.95
Y subject	1 if the individual is a Y subject; 0 otherwise (X subject)	0.50 0.50	0.50 0.50
Mixed session	1 if X and Y subjects can interact; 0 otherwise	0.33 0.67	0.33 0.67
Mixed session and Y subject	1 if Y is involved in a mixed session; 0 otherwise	0.16 0.84	0.16 0.84
Mode of compensation in part 1	1 if part 1 (periods 1 to 10); 0 if part 2 (periods 11 to 20)	0.50 0.50	0.50 0.50
Mode of compensation in part 1 and Y subject	1 if a Y subject in part 1; 0 otherwise	0.25 0.25	0.25 0.25
Lagged output levels of the other members of the group	Output of the other members of the group in the preceding period	154.88* (32.07)	146.49* (34.67)
Lagged output levels of the other members of the group and Y	Output of the other members of the group in the preceding period and Y	155.15** (32.37)	147.58** (35.04)
Index of difficulty	Index of difficulty of the curve/100	225.27 (143.10)	230.96 (143.36)
Index of difficulty squared	Index of difficulty squared	71798.18 (72847.8)	73882.02 (72.466.04)
Index of difficulty and period	Interaction of index of difficulty and period	2498.6 (2482.7)	2581.17 (2512.12)
Period	Period number from 1 to 20	10.5	10.5
Logarithm of period	Logarithm of period	2.12	2.12

* 1st and 11th periods excluded. ** For Y subjects only and 1st and 11th periods excluded.

Table 3--Students: determinants of output levels and costs

Variable	Output Panel Random*		Cost Panel Random**	
	Coefficient	*t-ratio*	*Coefficient*	*t-ratio*
Constant	**135.9[a]**	**26.5**	**11.3[a]**	**4.21**
1st period	**- 43.09[a]**	**-10.9**	0.671	0.362
11th period	-4.858	-1.46	0.387	0.303
Y subject	3.045	0.858	-1.9608	-0.831
Mixed	-0.7015	-0.324	-1.322	-1.24
Y and Mixed	1.908	0.625	1.912	1.27
Mode of compensation in part 1	-2.740	-1.02	**-5.27[a]**	**-4.10**
Y and mode of compensation in part 1	2.896	1.14	1.410	1.14
Output levels of other group members	-0.01724	-1.17	**0.02892[a]**	**4.18**
Y and output levels of other group members	-0.01738	-1.23	**-0.01186[b]**	**-1.79**
Index of difficulty	**-0.3076[a]**	**-24.2**	**0.03958[a]**	**6.63**
Index of difficulty squared	**0.0004981[a]**	**21.6**	**-0.00003859[a]**	**-3.57**
Index of difficulty and period	**0.001846[a]**	**2.73**	**0.0007233[a]**	**2.28**
Periods 1 to 20	**1.0409[a]**	**2.50**	**-0.6201[a]**	**-3.17**
Logarithm of period	**-18.20[a]**	**-7.35**	0.4660	0.968
Adjusted R^2	0.3569		0.163	
Number of observations	1440		1440	

* Linear one-way random effects model. Lagrange multiplier test versus OLS = 433.71 (1 df, prob value = 0.00000).** Linear one-way random effects model. Lagrange multiplier test versus OLS = 2761.10 (1 df, prob value = .000000)

a) significant at 5% level; b) significant at 10% level.

Table 3a--Managers: determinants of output levels and costs

Variable	Output Panel Random*		Cost Panel Random**	
	Coefficient	*t-ratio*	*Coefficient*	*t-ratio*
Constant	**125.32**[a]	**13.3**	6.197	1.13
1st period	- 11.80	-1.59	**10.25**[a]	**2.42**
11th period	-1.539	-0.289	0.967	0.318
Y subject	-3.030	-0.526	**8.06**[a]	**2.16**
Mixed	1.500	0.414	**6.23**[a]	**2.83**
Y and Mixed	-7.194	-1.41	**-14.61**[a]	**-4.69**
Mode of compensation in part 1	**-11.791**[a]	**-2.47**	-1.805	-0.653
Y and mode of compensation in part 1	-4.559	-1.04	**-10.09**[a]	**-3.88**
Lagged output levels of other group members	0.03469	1.251	0.02427	1.53
Y and output levels of other group members	**0.05420**[a]	**2.33**	0.0195	1.37
Index of difficulty	**-0.2132**[a]	**-9.41**	**0.05503**[a]	**4.27**
Index of difficulty squared	**0.0003381**[a]	**8.37**	**-0.00007559**[a]	**-3.29**
Index of difficulty and period	**0.003394**[a]	**2.82**	**0.001122**[b]	**1.64**
Periods 1 to 20	**-1.354**[b]	**-1.83**	**-1.037**[a]	**-2.47**
Logarithm of period	**-7.768**[b]	**-1.76**	2.799	1.11
Adjusted R^2	0.1562		0.1792	
Number of observations	720		720	

* Linear one-way random effects model. Lagrange multiplier test versus OLS = 12.06 (1 df, prob value = .000515).
** Linear one-way random effects model. Lagrange multiplier test versus OLS = 114.41 (1 df, prob value = .000000).
a) significant at 5% level; b) significant at 10% level.

References

Camerer, C.F., Knez, M., 1994. Creating "expectational assets" in the laboratory: "weakest-link" coordination games. *Strategic Management Journal* 15, 101-119.

Carpenter, J., Burks, S., Verhoogen, E., 2003. Comparing Students to Workers: The Effects of Social Framing on Behavior in Distribution Games Mimeo, Middlebury College.

Charness, G., Frechette, G., Kagel, J. (2001). "How Robust is Laboratory Gift Exchange?", Working Paper http://ssrn.com/abstract=292854.

Cooper, D.J., Kagel, J.H., Lo, W., Gu, Q.L., 1999. Gaming against managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. *American Economic Review* 89(4), 781-804.

Dickinson, D.L., 1999. An Experimental Examination of Labor Supply and Work Intensities. *Journal of Labor Economics* 17 (4), 608-638.

Dyer, D., Kagel, J.H., Levin, D., 1989. A Comparison of Naive and Experienced Bidders in Common Value Offer Auctions: A Laboratory Analysis. *The Economic Journal* 99, 108-115.

Erev, I., Bornstein, G., Galili, R., 1993. Constructive Intergroup Competition as Solution to the Free Rider Problem: A Field Experiment. *Journal of Experimental Social Psychology* 29, 463-478.

Falk, A., Ichino, A., 2003. Clean Evidence on Peer Pressure. IZA Discussion Paper 732, Bonn.

Fehr, E., Falk, A., 2002. Psychological Foundations of Incentives. *European Economic Review* 46, 687-724.

Fehr, E., List, J.A., 2002. The Hidden Costs and Returns of Incentives – Trust and Trustworthiness among CEOs. University of Zurich, Working Paper n°14.

Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, August, 1049-1074.

Hannan, R.L., Kagel, J.H., Moser, D.V., 2002. Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior. *Journal of Labor Economics* 20(4), 923-951.

Kreps, D., 1990. Corporate Culture and Economic Theory. In J. Alt and K. Schepsle (Eds.), *Perspective on Positive Political Economy*, Cambridge, Cambridge University Press, 90-143.

Lazear, E.P., 2000. Performance Pay and Productivity. *American Economic Review* 90 (5), 1346-1361.

Ledyard, J.O., 1995. Public Goods: A Survey of Experimental Research, in J.H. Kagel and A.E. Roth (Eds.), *Handbook of Experimental Economics*, Princeton, Princeton University Press, 111-194.

Locke, E.A., Saari, L.M., Shaw, K.N., Latham, G.P., 1981. Goal Setting and Task Performance: 1969-1980. *Psychological Bulletin* 90 (1), 125-152.

Meidinger, C., Rulliere, J.L., Villeval, M.C., 2003. Does Team-Based Compensation Give Rise to Problems when Agents Vary in Their Ability? *Experimental Economics* 6, 253-272.

Nalbantian, H.R., Schotter, A., 1997. Productivity under Group Incentives: An Experimental Study. *American Economic Review* 87, 314-341.

Prendergast, C., 1999. The Provision of Incentives in Firms. *Journal of Economic Literature* XXXVII (1), 7-63.

Sillamaa, M.A., 1999. How Work Effort Responds to Wage Taxation: An Experimental Test of a Zero Top Marginal Tax Rate. *Journal of Public Economics* 73, 125-134.

Tajfel, H., Flament, C., Billig, M.G., Bundy, R.P., 1971**.** Social Categorization and Intergroup Behavior. *European Journal of Social Psychology* 1, 147-175.

Van Dijk, F., Sonnemans, J., van Winden, F., 2001. Incentives Systems in a Real Effort

Experiment. *European Economic Review* 45, 187-214.

Weber, R.A., Camerer, C.F., 2003. An Experimental Approach to the Study Culture. *Management Science*. In Press.

Notes

[i] The experiment was computerized using the REGATE program. All curves used in this experiment and the instructions can be found at the following web site: http://platypus.gate.ec-lyon.fr/groupware/g_gate/Raventis.htm.

[ii] Once the curve is known, it is possible to find the best strategy to reach the highest peak at the lowest cost. A referee's suggestion to compare the subject's behavior with this optimal strategy appears worthwhile to pursue to study the learning process and the cognitive effort of the participants.

[iii] The negative impact on efficiency of the increasing size of teams in a merger has been documented by Camerer and Knez (1994).

[iv] It would have been interesting to test whether belonging to the majority category or the minority category of a team would influence individual behavior within teams. However, this would have required collecting a far greater number of observations. For this reason we did not inform subjects about the detailed composition of their teams.

[v] Overall, participants came from 4 different sites and 22 departments.

[vi] Reading instructions aloud guarantees that rules are common knowledge. However the section of instructions related to different payment schemes of the X and Y subjects must remain unknown until the end of the first part of the session. Reading aloud only other sections of the instructions would have focused undue attention on the question of compensation.

[vii] This is evocative of the result found by Charness, Frechette, and Kagel (2001) that in a gift-exchange game, there is less reciprocity when the payoffs for different actions are made clearer.

[viii] The observation that managers free-ride less than students is reminiscent of the Hannan et al. (2002) result that MBA students (which are more like managers) do more gift-exchange than other students.

[ix] Let E_{it} measure individual *i*'s output level in period *t*, explained by a vector of observable variables z_{it}, the corresponding parameter vector δ, a random individual component η_i and a random variable ε_{it}: $E_{it} = z_{it}\delta + \varepsilon_{it} + \eta_i$, $i = 1,\ldots,n$, $t = 1,\ldots,T$ with the usual assumptions, $\varepsilon_{it} \sim N(0,1)$, $\eta_{it} \sim N(0,\sigma^2)$, $\sigma_{\varepsilon\eta} = 0$.

[x] From the calculation: [8.06 – (8.06 + 6.23 – 14.61)]. The first term is the coefficient of the variable for Y participants in part 2 when groups are fixed. The second term represents the value of the coefficients for Y participants in part 2 when groups are mixed.

[xi] Despite their different strategies, the similarity of students' and managers' average net payoffs is striking. However, with more output achieved by the managers, the average surplus of a hypothetical firm (height times a constant minus wage costs) is higher when employing managers-subjects rather than students-subjects. The net payoffs are higher under the Y payment scheme than under the X payment scheme before the merger. After the merger, net payoffs are higher when teams, whether composed of students or managers, are homogeneous. In contrast, the hypothetical firm's performance is greater under the X payment scheme before the merger and with mixed teams after the merger.