HAL
archives-ouvertes.fr

Coordination in Teams : A Real Effort-task Experiment with Informal Punishment

Radu Vranceanu, Fouad El Ouardighi, Delphine Dubart

▶ **To cite this version:**

Radu Vranceanu, Fouad El Ouardighi, Delphine Dubart. Coordination in Teams : A Real Effort-task Experiment with Informal Punishment. ESSEC Working paper. Document de Recherche ESSEC / Centre de recherche de l'ESSEC. ISSN : 1291-9616. WP 1310. 2013. <hal-00857364>

HAL Id: hal-00857364

https://hal-essec.archives-ouvertes.fr/hal-00857364

Submitted on 3 Sep 2013

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.

Coordination in Teams: A Real Effort-task Experiment with Informal Punishment

Research Center
ESSEC Working Paper 1310

2013

Radu Vranceanu
Fouad El Ouardighi
Delphine Dubart

ESSEC
BUSINESS SCHOOL

August 23, 2013

Coordination in Teams: A Real Effort-Task Experiment with Informal Punishment

Radu Vranceanu[*], Fouad El Ouardighi[†], Delphine Dubart[‡]

Abstract

This paper reports the results from a real-effort team production experiment, where best performers can impose either tacit or explicit sanctions on their less-performing partners. The behavior of the best performer in the team differs from one condition to another. When explicit sanctions are not allowed, good performers reduce their effort in response to the advantageous difference in scores; when they can impose sanctions, their change in effort is no longer related to the difference in scores. To some extent, a mechanism of explicit sanctions allows good performers to focus on their own performance. Not sanctioning an opponent who under-performs, what we refer to as forgiveness, prompts the latter to improve his performance, but applying the sanction has a stronger effect.

Keywords: Real-effort task, Team production, Punishment, Forgiveness, Tacit vs. explicit sanction.
JEL Classification: M52; C92; D03

[*]ESSEC Business School, Cergy, France, PB 50105, 95021 Cergy, France E-mail: vranceanu@essec.fr

[†]ESSEC Business School, Cergy, France, PB 50105, 95021 Cergy, France, E-mail: elouardighi@essec.fr

[‡]ESSEC Business School, Cergy, France, PB 50105, 95021 Cergy, France. E-mail: dubart@essec.fr

1 Introduction

In general, the organization of economic activity within firms involves teams of individuals who jointly produce goods and services. In many cases the manager cannot observe the contribution of every individual to the final output. Hence in practice many organizations will reward employees proportional to total output. Economists are rather skeptical about the merits of such compensation schemes: if each individual relying on the efforts of his partners reduces his own effort, in equilibrium, all follow the same strategy and output melts down. Substantial theoretical literature has investigated contracting mechanisms that can contain free-riding behavior in team production (inter alia, Alchian and Demsetz, 1972; Holmstrom, 1982; McAfee and McMillan, 1991; Itoh, 1991,1992; Legros and Matthews, 1993; Che and Yoo, 2001). One standard result is that free-riding might be alleviated if the *group* develops internal monitoring processes and can *punish* defectors by taking actions that will diminsh defectors's payoff.

This conclusion seems to be corroborated by experimental research on a particular context-free coordination problem known as the *public good game* [1] Following the pioneering work by Fehr and Gächter (2000, 2002), many experiments analyzed the merits of punishment in repeated public good games. In these experiments, punishment is usually expressed by the ability to decrease the payoff of any player, and in general it is costly for the punisher.[2] The key result is that allowing for punishment does help containing free-riding behavior. Lab experiments also pointed out that subjects differ in their punishment strategies. For instance, Camerer and Fehr (2006) acknowledge existence of a large proportion of "strong reciprocators", i.e. individuals that show a combination of altruistic rewarding and altruistic punishment, defined as the propensity to impose sanctions on others for norm violations. Such strong reciprocators will bear the cost of punishing even if they gain no individual benefit.[3]

[1] In a typical public good game, a group of N players is asked to contribute to a common pool of money, each getting back $m\%$ of the total contribution. The non-trivial situation requires $(1/N) < m < 1$; in this case, a player does not gain enough to individually contribute one unit to the public good ($m < 1$), but everybody would gain from jointly contributing the same one unit ($Nm > 1$). The dominant strategy of the rational agent is to free-ride.

[2] See for instance: Anderson and Putterman, 2006; Carpenter, 2006, 2007; Carpenter et al., 2009; Casari and Luini, 2009; Herrmann et al. 2008; Masclet et al., 2003; Masclet and Villeval, 2008; Nikiforiakis and Normann, 2008; Nikiforiakis, 2008, 2010. See also Chauduri (2011) for a survey.

[3] Falk et al. (2005) and Casari and Luini (2009) show that sanctions can be sometimes driven by spiteful behavior, being imposed by defectors on players who cooperate.

On the other hand, acknowledging that team production involves subtle and complex human interactions, other experimentalists developed analyses that simulate a genuine production process. In general these experiments measure the participants' effort on the input side, and the group's performance on the output side. One can distinguish between "chosen effort", where participants can select the "effort level" from an increasing scale of costs[4] , and "real effort", where participants are asked to perform a well-defined tedious task.[5]

However, in contrast with literature on context-free experiments (public good, prisoner dilemma), the literature on punishment in real-effort experiments is extremely scarce. This paper aims to fill this gap, by studying the consequences of punishment in a real-effort team production experiment. More in detail, we aim to compare team performance under two distinct types of punishment applied to defectors: outright sanctions and voluntary effort reduction (tit-for-tat).

Falk and Ichino (2006) used a real effort experiment to study the effect of peer pressure in team production.[6] In their conclusion, they hinted that cooperation in real-effort joint production experiments might further increase if group members can directly impose sanctions on defectors (they did not test it). This assumption is consistent with the above mentioned theoretical papers, and from evidence provided by the public good experiments. Nonetheless, if we agree that real-effort experiments contribute to develop of strong ties between subjects, negative emotions related to the sanction might hamper effort, what psychologists refer to as the "motivation crowding-out" effect (see for a survey Frey and Jegen, 2001). Symmetrically, not imposing a sanction when one has the option to do so might bring about a "crowding-in" effect where positive emotions create the motivation to increase effort. The potentially detrimental effect of sanctions and, less frequently, the stimulating effect of refraining from imposing sanctions, have already been emphasized by experimental economists who studied other games, such as the principal/agent game (Dickinson and Villeval, 2008; Kirstein, 2008), the trust game (Fehr and List, 2004; Fehr and Rockenbach,

[4] See for instance Meidinger, Rullière and Villeval (2003) or Brands and Cooper (2006).

[5] See: van Dijk, Sonnemans and van Winden (2001), Falk and Fehr (2003); Brüggen and Strobel (2007), Vandegrift and Yavas (2011), Eriksson, Poulsen and Villeval (2009), Mohnen, Pokorny and Sliwka (2008), Falk and Ichino (2006), Kirstein (2008). Brüggen and Strobel (2008) bring some support to the equivalence between real effort and menu cost experiements.

[6] Two individuals that observe each other were asked to put letters into envelopes.

2003; Houser et al., 2008) or the ultimatum game (Gneezy, 2003).[7]

Furthermore, the response of a defector to a sanction can depend on the nature of the sanction. Falk et al. (2005, p. 2017) called sanctions *informal* when "they are not imposed by formal, legal bodies, but by private parties who punish other people's observed behaviors". We will further distinguish between explicit and tacit sanctions. An explicit sanction is directly observable and identified as such; in the business world, fines, written critics, dismissals and other disciplinary measures belong to this category. Tacit sanctions are insidious costs imposed on the punished person such as social exclusion from the group, assignment to useless tasks, constant taunts, etc.

Our analysis builds on an experiment developed by Mohnen, Pokorny and Sliwka (2008). In their study, subjects are matched in pairs at the beginning of the experiment. During one round that lasts for eight minutes, the computer displays successive blocs of figures. For each bloc, subjects were asked to count how many times the figure seven occurs and to accurately report the number.[8] They alternatively can push a "time-out button" and earn a predetermined amount of money while resting. The task must be performed for seven distinct rounds.[9] The individual reward from the task is proportional to the joint production of right answers. Mohnen et al. (2008) compared the average output when the performance of each partner was provided after each round (transparent condition), to the average output when information was provided only at the end of the experiment. One important result was to show that in the transparent condition, a low performance player would increase his effort from one round to another, and, more importantly, a high performance player would reduce his effort.

We extend their experiment by adding one original treatment where subjects who outperform in the task have the costless option to impose a fine on their partner. Those who underperform cannot impose penalties on the other. At any round, we refer as the "good performer" ("bad performer") to the player who provides the highest (lowest) score of good answers (roles can thus

[7] See also El Ouardighi (2013) for a dynamic game theoretical analysis where team production is influenced by behavioral patterns such as the propensity to forgive and the propensity to develop ties with his partner.

[8] Counting 7s is an appealing task for experimentalists because it does not require any special ability, it yields a low intrinsic reward, and learning how to execute it is limited. On the negative side, this task exposes subjects to boredom and fatigue stemming from its repetitive nature.

[9] The first and the seventh round being assigned to an individual task.

change from one round to another). In case of equal performance, no penalty can be imposed. Not punishing a player who underperforms will be defined as *forgiveness*. To contain fatigue and boredom specific to such a routine task, in our experiment we limit the number of active rounds to four and the duration of each round to four minutes.

Comparing performance in the no-sanction treatment with performance in the with-sanction treatment provides interesting insights. Mohnen et al. (2008) argued that good (bad) performers tend to reduce (increase) because of inequality aversion: since a good performer earns the same amount as the poor performer but deploys a higher effort, the former would reduce his effort and the latter would increase it to equalize net payoffs. However, "tit-for-tat" might provide an alternative explanation.[10] In this experiment, good performers might reduce their effort only to punish poor performers; this choice can be seen as a tacit sanction. Unlike the explicit sanction, this tacit sanction is costly for a good performer. The empirical evidence provided by our experiment tends to support the "tit-for-tat" explanation: when the explicit sanction is possible, good performers no longer reduce their effort from one round to another, while they do so in the no explicit sanction treatment. An additional interesting question is why some good performers forgive defectors in the with sanction treatment, and what are the consequences of forgiveness on average performance.

Through this experiment we will be able to address three main research questions with non-negligible managerial implications:

1. *Does introduction of an explicit sanction mechanism in a no-sanction environment contribute to improve the overall performance of the team?*

2. *Does introduction of an explicit sanction mechanism in a no-sanction environment modify the retaliation strategy of best performers? And what is the response of the poor performers?*

3. *When an explicit sanction mechanism is at work, should best performers punish or forgive the worst performers?*

[10] As shown by Axelrod (1984), the "tit-for-tat" strategy defined by the sequence "cooperate at the first move, than do what the other player does", supports the emergence of cooperation in repeated games (see also Che and Yoo, 2001).

The paper is organized as follows: The next section introduces the experimental design. Section 3 presents the results. The last section presents our conclusions and provides some managerial implications.

2 Experimental design

All subjects were recruited from the student population of the ESSEC Business School, who answered to a call for paid decision experiments. This group of students is relatively homogenous in terms of computing and intellectual abilities. Five sessions were organized, with a total of 118 subjects (see Table 1).

The real effort experiment consists of a tedious computer task in which subjects are asked to count the number of 7s in blocks of random numbers (Mohnen et al., 2008; Pokorny, 2008). At the beginning of the experiment, subjects are matched in pairs at random. The team composition is kept the same across rounds. Interaction is anonymous, subjects do not know who are their partners. They play behind the computer screen; instructions and data collection are computerized, the programme is developed with z-Tree (Fischbacher, 2007). Pairs must execute the real effort task for four consecutive identical rounds. Each round lasts for 4 minutes.[11] In a round, at each point in time, subjects have the choice between two options: they can choose to work on the counting task and receive a payout based on the number of correct answers (the individual compensation will depend on the joint production) or they can choose to push a time-out button and, for 20 seconds, the screen is blocked, but they are still paid 6 ECU. This second option can be viewed as an opportunity cost of working.[12]

Three distinct treatments have been implemented (different subjects). At the beginning of each treatment, subjects are informed about the structure of the game: task, payoffs, information, and number of rounds.

(**T1**). This is a piece-rate treatment, used as a benchmark, where subjects are rewarded on

[11] A fifth round was also included, where players must perform the task individually.

[12] The time-out button is desactivated 20 seconds before the end of the round (this is the average time needed to count 7s in the last block of numbers).

an individual basis according to the compensation function:

$$Z_i = 10N_i + 6k_i, \tag{1}$$

where Z_i is the total payoff a player i, N_i is the number of correct answers provided and k_i is the number of times the time-out button was pressed. For instance, a player that brings 8 correct answers and pushes twice the time-out button will be entitled to $Z = 10 \times 8 + 2 \times 6 = 92$ ECU.

At the end of each round, players are informed about their own performance: number of blocks worked on, number of correct answers, number of times the time-off button was used, as well as about their gain in ECU.

(**T2**). In the joint production no-sanction treatment, a player's reward is a linear function in the sum of correct answers provided by the two players; if an individual pushes the time-out button for k times, he also gets the connected payoff. Denoting by N_j the number of correct answers provided by his partner j, the compensation function of an individual i is:

$$Z_i = 10\frac{N_i + N_j}{2} + 6k_i. \tag{2}$$

At the end of each round, they learn their own performance (number of correct answers, number of blocks worked on, number of times the time-off button was used), as well as the number of correct answers provided by the partner. They also learn their own ECU payoff for the round.

The parameters of the game were chosen such that free-riding (i.e., pushing the time-out button) be the dominant strategy. An individual who makes a normal effort to count the 7s can provide a right answer in 20 seconds on average. If the other player does the same, both would earn 10 ECUs. But if one pushes the time-out button (blocks the screen and rests for 20 seconds) while the other works and gives a right answer, the player who free rides gets 11 ECUs (and the other gets 5ECUs). For sure, 11 ECU while resting is better than 10ECU *and* the disutility form executing the boring task. If both players free ride and push the time-out button, they get 6 ECUs. This is a typical prisoner dilemma setting, where, in a one-shot game, both players have an incentive to play the selfish strategy. We can remark that, in this variant of the game, cooperators (i.e. those who deliver a high effort level) can punish defectors only by reducing their own effort level at the next round, i.e., play tit-for-tat.

(**T3**). In the joint production with possible sanction treatment, the reward from the task is similar to T2, but we allow for punishment opportunities. At the end of each round, players learn their own performance (number of blocks worked on, number of correct answers, number of times the time-off button was pressed), as well as the number of correct answers provided by the partner. Then, the player who out-performed (i.e., provided more correct answers than the other) is asked if he wants to impose a penalty on his partner. If he answers "yes", he can wipe-off to the other player an amount p, with p between 1 and 30 ECUs without any cost for himself.[13] Notice that we adopt this costless sanction design on purpose, in order to focus on those cases where the best performer in the team does not sanction the poor performer. The compensation function of individual i is thus:

$$Z_i = 10\frac{N_i + N_j}{2} + 6k_i - p_i, \tag{3}$$

where $p_i = 0$ if $N_i > N_j$, and $p_i \in [0; 30]$ if $N_i < N_j$. Remark that p_i diminishes the payoff of i, but is decided by player j.

Before moving to the next round, both players are informed about their ECU payoff. In addition, the player that under-performed is informed whether he was sanctioned (and by how much) or not.

The key treatments are T2 and T3. The purpose of T1 to provide a more complete information about the behavior of the players when submitted to this specific real task.

Table 1 presents the distribution of subjects with respect to sessions and treatments:

	date	T1	T2	T3
Session 1	12/06//2012	no	no	22 subjects
Session 2	25/10/2012	no	no	20 subjects
Session 3	08/11/2012	no	no	26 subjects
Session 4	15/11/2012	no	26 subjects	no
Session 5	21/03/2013	24 subjects	no	no

Table 1. Distribution of subjects with respect to sessions and treatments

[13] The penalty reduces the payoff of the punished subject and has no effect on the punisher; it is not a transfer from the punished to the punisher.

T1. Individual production, piece-rate payment, T2. Team production: no sanction possible, T3. Team production: sanction possible

On average, subjects earned 12.86 euros.[14]

3 Results

3.1 Comparisons between average performance

Table 2 and Figure 1 provide data on individual's "performance" as measured by the average number of correct answers per player, for each treatment and per round. The last column provides the all round mean performance value.

	R1	R2	R3	R4	All round av.
T1	7.7	8.7	10.3	10.8	9.4
T2	8.6	9.3	10.8	10.5	9.8
T3	7.2	9.1	10.8	11.3	9.6

Table 2. Average number of correct answers per subject

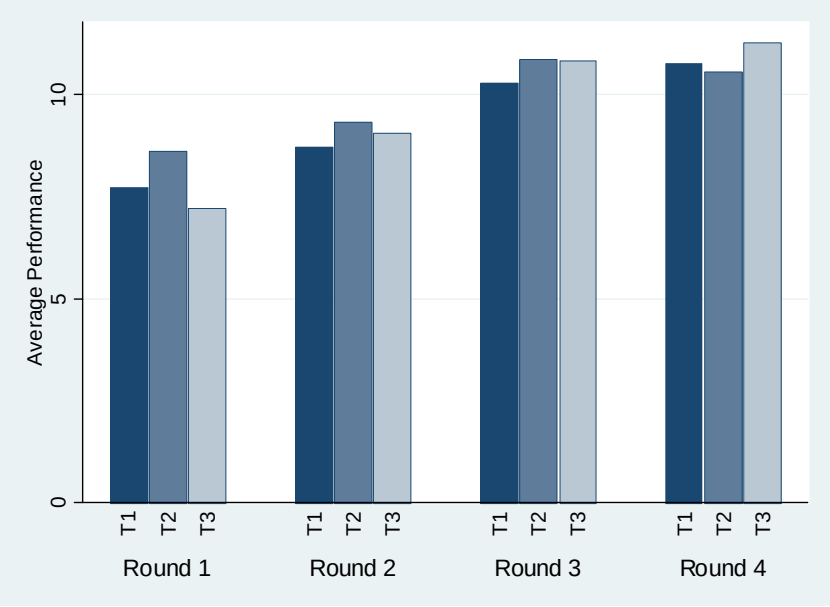

Figure 1: Average score per individual, over treatements and rounds

[14] The average gain was 12.52 euros for T1, 13.24 euros for T2 and 12.84 euros for T3.

A *Kruskal-Wallis* non-parametric test confirms that there is no significant difference in average performance between the three treatments (overall and for each round).

The correspondence between the performance in T1 on the one hand and T2 and T3 on the other hand suggest that in such a "real effort" experiment subjects adopt a cooperative behavior even if "pure rationality" would prompt them to choose the "free-riding" strategy.

A less trivial result is the correspondence between average performance in T2 and T3. Contrary to what intuition would suggest, the punishment option does not raise the overall average performance.

However, a similar *Kruskal-Wallis* test shows that there is a significant difference between the first and the last round performance in T2 and T3: over time subjects either "warm-up", or increase their effort level, or both. In particular, in the with-sanction treatment, the initial effort level is lower than in the no-sanction treatment, but the over rounds increase in performance is stronger than in the no-sanction treatment.

As shown in Figures 2 and 3, the behavior over time (rounds) is also different between good and bad performers in T2 as compared to T3; recall that the good (bad) performer at a given round is the player with the highest (lowest) score among the two players (we exclude observations where scores are identical). In general, good performers will improve their performance from one round to another; while the change is not significant in T2, it is quite large in T3 (Figure 2). Bad performers will also improve their performance, and the progress they made is stronger in T3 as well (Figure 3).

Some clues about the causes of improving performance can be obtained by checking the average number of time-outs per player, to be interpreted as a "pure" decision to free-ride.

	R1	R2	R3	R4	All round av.
T1	0.50	0.71	0.33	0.13	0.42
T2	0.85	0.58	0.38	0.35	0.54
T3	0.63	0.29	0.31	0.38	0.40

Table 3. Average number of time-outs per subject

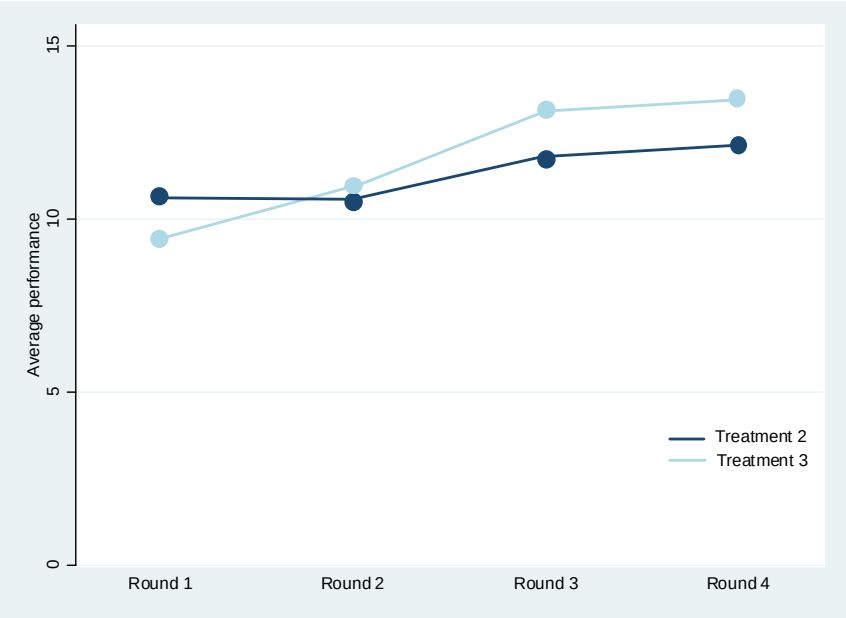

Figure 2: Average score of good performers, T2 vs. T3

In general the number of time-outs tends to decline over the rounds, although there is no clear pattern. As expected, the number of time-outs decline in a significant way in the individual treatment (T1). Some of the improvement in performance can thus be attributed to higher effort. Comparing the all round average, the number of time-outs is larger in the no sanction treatment (T2) than in the with sanction treatment (T3), indicating that the absence of sanctions would lead to additional free riding. However, if in T2 the number of time-outs declines from round 2 to round 4, in T3, the number of time-outs is raising.[15]

Turning now to T3 and focusing on the sanction, Table 4 presents the number of possible sanctions (counting all situations where the performance of one player was better than the performance of the other, and excluding all similar performances), the number of actual sanctions applied, the average amount and the extremes.

[15] A quick check of individual data show that time-outs are not specific to subjects who got a sanction at the previous round (no retaliation effect).

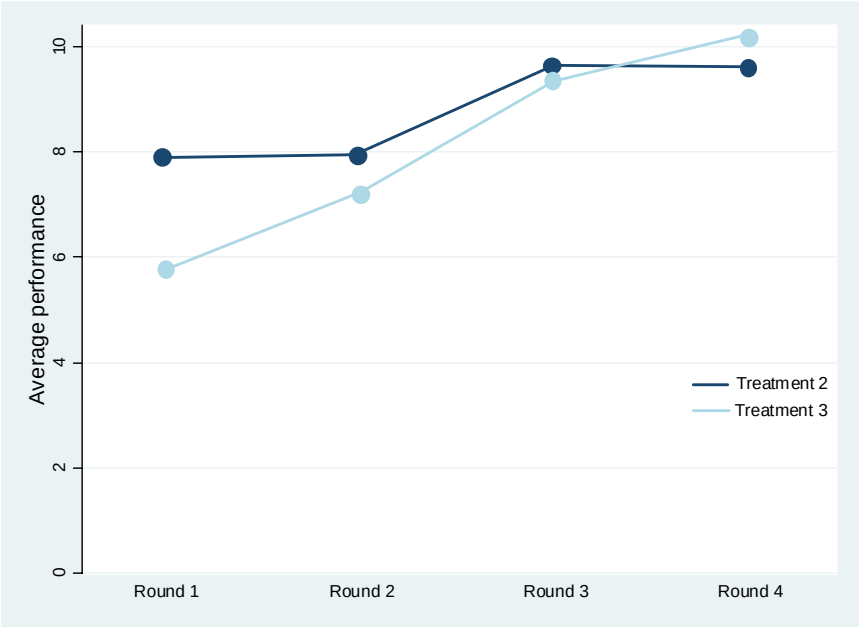

Figure 3: Average score of poor performers, T2 vs. T3

	R1	R2	R3	R4	All round av.
Nb. of possible sanctions	30	30	30	31	121
Nb. of actual sanctions	9	12	7	10	38
Average amount	11.67	13.91	16.71	19	15.23
Highest	20	30	30	30	30
Lowest	1	1	1	1	1

Table 4. Treatment 3: The sanction statistics

Sanctions were applied in about 31% of the cases and forgiveness was chosen in the other 69%; the average amount of the sanction raises steadily from round 1 to round 4.

Notice that in 10 out of 31 situations, good performers apply the sanction at the last round, when strategic considerations (reputation, motivation) no longer apply. This outcome could be justified if they behave as "strong reciprocators" (Camerer and Fehr, 2006), but also with the fairness explanation provided by Mohnen et al. (2008). It can be also driven by plain vengeance, a motive not so far from spiteful behavior revealed by Falk et al. (2005) and Casari and Luini (2009). Unfortunately our experiment cannot bring a definitive conclusion for this "end of game"

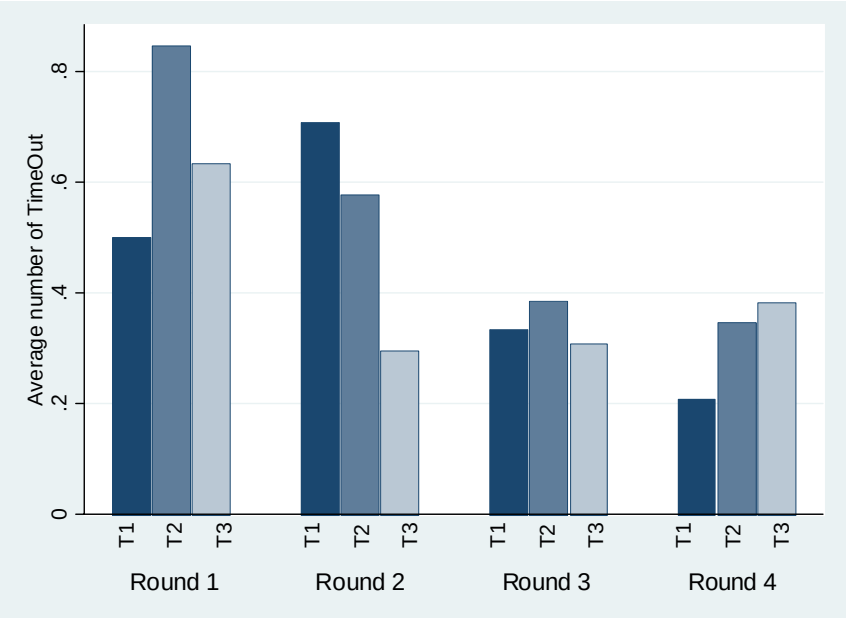

Figure 4: Average number of time-outs per individual, over treatements and rounds

behavior.

We are interested in particular in the possibility that the same individual adopts a systematic "punish" or "don't punish" strategy. By systematic, we understand a situation where individuals who performed better more than once, applied the same strategy. As shown in Table 5, 9 subjects performed better than their partner during two out of the four rounds, 7 subjects performed better than the partner during three rounds, and 18 subjects performed better over the four rounds. The strategy "systematically forgive" was applied 17 times (last line), the strategy "systematically punish" was applied 6 times, and no systematic strategy occurred in 11 cases.

	Nb.	Systematically forgive	Systematically punish	Do both
Do better two times	9	7	1	1
Do better three times	7	3	1	3
Do better four times	18	7	4	7
Overall	34	17	6	11

Table 5. Treatment 3: Distribution of punishment strategies

3.2 Regression analysis

Regression analysis allows us to push the analysis a step further. Let us introduce firstly the basic notations:

N_{it} the score of a player i at round t; N_{jt} is the score of his partner, at the same round.

$DIFSCORE_{it} = N_{it} - N_{jt}$; if i is the best performer in the team of two, the variable is positive, if i is the worst performer in the team, the variable is negative;

$DELTAN_{it} = N_{it} - N_{it-1}$ is the score change from one round to another for individual i;

$DUMG_{it}$ takes the value 1 if $N_{it} > N_{jt}$, and 0 if else;

$DUMB_{it}$ takes the value 1 if $N_{it} < N_{jt}$, and 0 if else;

$MSANC_{it}$ is the amount of the sanction [0, 1,... 30] (0= no sanction);

$DUMS_{it}$ takes the value 1 if individual i applies a sanction (whatever N_i, N_j), 0 if else;

$ROUND \in [1, 2, 3, 4]$.

We follow Mohnen et al. (2008) and estimate a "Performance equation" for Treatment 2 (no-sanction). The dependent variable is the score change, $DELTAN$, depending on whether the subject was the poor (index B) or the best performer (index G) in the team.[16]

$$\begin{aligned} DELTAN_{it} &= C + \alpha_G DUMG_{it-1} \times DIFSCORE_{it-1} + \\ &\quad + \alpha_B DUMB_{it-1} \times DIFSCORE_{it-1} + \alpha_R ROUND + u_i + e_{it}, \end{aligned} \tag{4}$$

where u_i is the individual error term, and e_{it} is the standard error term.

The ROUND variable would capture any warm-up / learning effect ($\alpha_R > 0$) or boredom / fatigue effect ($\alpha_R < 0$) . Notice that for poor performers the $DIFFSCORE_{it-1}$ variable is a negative number. Hence a negative coefficient α_B would suggests that a poor performance is associated with an improvement in the performance of the poor performers. Table 6 presents our estimates of Eq (4):

[16] Besides the score of right answers, Mohnen et al. (2008) also used as a proxy for performance the number of blocs counted. For the sake of parsimony we do not follow their approach; we just checked that numbers of blocs counted and score are highly correlated variables (there is no correlation coefficient below 0.75 for any of the rounds 1 to 4 in any treatment).

	Coef.	z	Pr>|z|
$DUMG_{it-1} \times DIFSCORE_{it-1}$	$\mathbf{-0.137}$	-2.34	0.019
$DUMB_{it-1} \times DIFSCORE_{it-1}$	-0.088	-1.56	0.118
$ROUND$	-0.039	-0.11	0.913
C	1.110	1.07	0.284
Nb. Obs	78		
Nb. Subj.	26		
Nb. Rounds	3		
R-sq. overall	0.11		
Prob $< \chi^2$	0.025		

Table 6. Performance regression, T2. Random effects regression model

Mohnen et al. (2008) found out that, in a similar no-sanction team production experiment, good performers tend to reduce their effort in a significant way ($\alpha_G = -0.987$, significant at 1%), and poor preforms tend to increase it ($\alpha_B = -0.139$, significant at 10%).

According to our estimates, a good score is associated to a reduction in the score of the good performers from one round to another. Indeed, the coefficient α_G is negative and statistically significant. With respect to poor performers, the sign of the coefficient α_B is indeed negative but it is not statistically significant; we cannot state that poor performers respond to the difference in score in the no sanction treatment.

We turn now to the analysis of our original with-sanction treatment (T3).

The overall performance equation is:

$$\begin{aligned} DELTAN_{it} &= C + \alpha_G DUMG_{it-1} \times DIFSCORE_{it-1} + \\ &+ \alpha_B DUMB_{it-1} \times DIFSCORE_{it-1} + \alpha_S MSANC_{it-1} + \alpha_R ROUND + u_i + e_{it} \end{aligned} \quad (5)$$

where u_i is the individual error term, and e_{it} is the standard error term. The estimated equation is:

	Model 1			Model 2		
	Coef.	z	Pr>|z|	Coef.	z	Pr>|z|
$DUMG_{it-1} \times DIFSCORE_{it-1}$	0.056	0.97	0.33	0.056	0.98	0.32
$DUMB_{it-1} \times DIFSCORE_{it-1}$	**−0.249**	−3.76	0.00	**−0.276**	−4.76	0.00
$MSANC_{it-1}$	0.0211	0.84	0.40	–	–	–
ROUND	**−0.705**	4.48	0.00	**−0.703**	−4.48	0.00
C	**2.945**	5.76	0.00	**2.938**	5.75	0.00
Nb. Obs	204			204		
Nb. Subj.	68			68		
Nb. Rounds	3			3		
R-sq. overall	0.18			0.18		
Prob $< \chi^2$	0.00			0.00		

Table 7. Performance equation, T3. Random effects regression model.

We observe that best performers *do not* systematically reduce their performance in response to a difference in scores (the coefficient α_G is not statistically significant), while, according to data in Table 6, they reduce their performance from one round to another in the no-sanction treatment (T2). Furthermore, poor performers *do* improve in a significant way their scores from one round to another while they do not improve in a systematic way their performance in T2 (Table 7).

The amount of the sanction is not statistically significant (Model 1), which hints that applying sanctions might not be an efficient method to motivate subjects to work harder. On the other hand, the sanction can be correlated with the difference in scores, so the latter variable can absorb the impact of the sanction. So, an interesting complementary test would consist in studying the improvement in performance *only of those subjects who under-performed*, considering the impact of the amount of the sanction at time $t-1$ (this amount can be 0 or positive).

$$DELTAN_{jt} = C + \mu_S MSANC_{jt-1} + u_j + e_{jt}. \tag{6}$$

The estimated equations is:

	Coef.	z	Pr>|z|
$MSANCT_{jt-1}$	**0.053**	2.07	0.039
$ROUND$	**−0.694**	−2.66	0.008
C	**3.629**	4.44	0.000
Nb. Obs	90		
Nb. Subj.	41		
Round (av.)	2.2		
R-sq. overall	0.11		
Prob $< \chi^2$	0.004		

Table 8. Performance of poor performers. RE regression

Results are interesting; on the one hand, the higher the sanction, the higher the improvement in performance; the maximum sanction (of 30) would add 1.59 score points to the no-sanction change.[17] Punishment helps motivating poor performers. On the other hand, no sanction is still consistent with a positive increase in performance; even at Round 4, setting MSANC to zero still brings about a positive change in performance. In this respect, forgiveness might not be the best strategy to raise performance, but is neither a performance deteriorating strategy as would suggest the theory of purely opportunistic behavior (according to which an unpunished cheater continues to cheat).

At the end of the experiment we addressed a short survey asking poor performers at round 2 what did they feel about receiving (or not receiving) a sanction. We focus on the second round because there were two more rounds left and strategic considerations further apply. We realize that the sample is very small, so these answers should be taken with extreme caution. Nonetheless, answers corroborate the regression results. The twelve participants who received a sanction were offered two sentences: "*I realized that it was necessary to increase my effort if I want to improve my gains*" or "*I was somehow frustrated, which prompted me to reduce my efforts for the next*

[17] It is a plausible assumption that "reasonable" sanctions would bring about a positive incentive effect, while too strong sanctions would entail negative sentiments prompting receivers to reduce their effort. We thus also tested a quadratic form, including the term $MSANC^2$. While the coefficient is negative, it is not statistically significant.

period". Answers can be provided on a scale from 1 to 5, going from "total disagreement" (1) to "full agreement" (5). The average score to the first sentence was 3.75 (9 out of 12 respondents agree or strongly agree), and as low as 1.5 to the second (11 respondents disagree or strongly disagree).

There were also 18 participants who could receive a sanction but did not receive one. They were also offered two sentences aiming to gauge their feelings on being forgiven: "*A sort of trust, that I had to confirm by working harder*" and "*A sort of comfort; I had in front of me a tolerant person so I could maintain or reduce efforts at the next round*". 16 out of 18 respondents agreed and fully agreed with the first statement, only 3 fully agreed or agreed with the second (average 2).

A last interesting topic is the decision to apply the sanction, over the four rounds ($t = 1, 2, 3, 4$). We estimate a sanction equation where the dependent variable is the amount of the sanction or a the sanction/no sanction dummy. We will thus build an unbalanced panel including *only observations for players that outperformed*. On average, the same player outperformed the others in 2.8 rounds. Thus, over 68 players, we have 44 players that outperformed at least once. The "sanction equation" can be written:

$$MSANC_{it} = C + b_D DIFSCORE_{it} + b_R ROUND + u_i + e_{it}, \tag{7}$$

where u_i is the individual error term, and e_{it} is the standard error term.

The estimated equation is:

	Coef.	z	Pr>|z|
$DIFSCORE$	**1.558**	5.21	0.000
$ROUND$	0.727	1.46	0.145
C	−2.668	−1.39	0.164
Nb. Obs	121		
Nb. Subj.	44		
Round (av.)	2.8		
R-sq. overall	0.10		
Prob $< \chi^2$	0.000		

Table 9. The determinants of the sanction. RE regression.

As expected, the amount of the sanction appears to be related to the difference in scores; a one point difference in scores raises the sanction by 1.6 ECUs. While these results suggests that the amount of the sanction is related to the difference in scores, these data cannot tell us much about the deep motivation to punish the low performer. Probably, the incentive motive plays an important role, validated by the fact that punished persons do raise their efforts. However, the non negligible number of sanctions issued at the last round (Table 4), shows that other motives such as strong reciprocation, aligning earnings with effort (fairness) or plain vengeance cannot be ruled out.

4 Conclusion

This paper aims to contribute to the literature on cooperation in teams by extending the real effort-task experiment introduced by Mohnen et al. (2008) to allow for explicit informal punishment. In what can be seen as an original treatment, at each round high performers have the option to freely impose a sanction on low performers. There is also a piece-rate, individual production treatment and a joint production, no-sanction treatment. We can summarize our main findings, as answers to the three questions raised in the Introduction.

1. In this real-effort experiment cooperation in teams appears to be quite strong, and this

outcome is observed in both the no-sanction and the with-sanction treatments. On average, performance in team production, as measured by the all-round score of good answers, is not lower than performance in individual piece-rate production. In particular, introduction of an explicit sanction mechanism does not bring about a significant improvement in the team's performance, compared to individual production or to the no-sanction setup.

2. The introduction of a mechanism of explicit sanction modifies behaviors of good and poor performers. In our experiment, in the no-sanction condition, good performers respond to the difference in score by reducing their effort, but poor performers do not systematically raise their effort level. We interpret the good performers behavior as illustrative of a tit-for-tat retaliation strategy intended to restore cooperation. Indeed, a lower effort is tantamount to a sanction on the poor performers, though at a cost for the good performers themselves. This retaliation explanation is backed by results from our third treatment. When explicit sanctions are allowed, good performers no longer reduce their performance in a systematic way, they even show a weak tendency to improve on it. In other words, they no longer need to resort to the tit-for-tat strategy to punish the poor performers, because they now have a more efficient tool. This explanation is also consistent with the behavior of the poor performers who, under the threat of the sanction, do improve their performance in a substantial way from one round to another.

Remark that tit-for-tat is most efficient in two-person teams; the larger the team, the smaller the consequence of individual effort reduction on the poor performers; in this context, the explicit sanction – if it can be applied without costs on *all* poor performers – would be a more efficient device to achieve cooperation.

3. A poor performer who does not receive a sanction will, on average, improve his performance, which suggests that forgiveness can be a meaningful choice. However, optimism on this finding is limited by evidence showing that those who receive a sanction tend to work even harder. Of course, better morale and a better group dynamic might provide additional arguments in favor of the forgiveness option, but our experiment does not allow to address this important question.

In addition to these essential findings, we observed that the amount of the sanction is related to the difference in performance, a large difference justifying a large sanction. Less intuitively,

some subjects will impose sanctions at the last round, showing that forgiveness is not a generalized behavior even among young subjects belonging to the same group. The experiment leaves unanswered whether this end-of-game behavior is driven by fairness concerns, strong reciprocation, or pure spite.

Results from such a simple experiment shed light on some important managerial issues, especially when deciding about the nature of interactions that should be established among team members on the workplace. Tit-for-tat appears to be a powerful coordination device; introducing explicit sanctions would not increase the overall performance of the team, but would allow best players to focus on improving their performance over time, and use sanctions or forgiveness as a way to motivate poor performers. When the success of a team project depends on the performance of a group of outstanding people backed by followers (R&D, innovation-driven growth, etc.), then internal explicit sanctions can help protecting the performance of the leading group against followers' free riding.

Further research might analyze a longer interaction, to test for the long-run benefits of forgiveness such as better morale or a better group dynamic; it might also be interesting to compare the relative merits of tit-for-tat and explicit sanction in larger teams. Bigoni et al. (2013) have shown that white collars and students present different responses to punishment in a context-free coordination game; it would be interesting to test our real effort punishment game with other types of subjects. Finally, another extension of this study would consider the case of costly sanctions; it might be interesting to determine at what cost of the sanction subjects become indifferent between punishing by reducing effort and punishing by applying the sanction.

Acknowledgments. The authors would like to thank participants to the *ESA European Conference* in Koeln, 12-15 September 2012, Nona Pepito, Gorkem Celik, and Oana Peia for their suggestions and remarks that helped them to improve this paper.

References

Anderson, Christopher M. and Louis Putterman, 2006, Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism, *Games and Economic Behavior*, 54, 1, pp. 1–24.

Andreoni, James, William T. Harbaugh and Lise Vesterlund, 2003, The carrot or the stick: Rewards, punishments and cooperation, *American Economic Review*, 93, pp. 893–902.

Alchian, Armen and Harold Demsetz, 1972, Production, information costs, and economic organization, *American Economic Review*, 62, 5, pp. 777–795.

Axelrod, Robert, 1984, *The Evolution of Cooperation*, Basic books, New York.

Bigoni, Maria, Gabriele Camera and Marco Casari, 2013, Strategies of cooperation and punishment among students and clerical workers. Forthcoming, *Journal of Economic Behavior and Organization*.

Brandts, Jordi and David J. Cooper, 2006, A change would do you good. An experimental study on how to overcome coordination failure in organizations, *American Economic Review*, 96, 3, pp. 669–693.

Brüggen, Alexander and Martin Strobel, 2007, Real effort versus chosen effort in experiments, *Economics Letters*, 96, 2, pp. 232–236.

Carpenter, Jeffrey P., 2007a, Punishing free-riders: How group size affects mutual monitoring and the provision of public goods, *Games and Economic Behavior*, 60, 1, pp. 31-51,

Carpenter, Jeffrey P. 2007b, The demand for punishment, *Journal of Economic Behavior and Organization*, 62, pp. 522–542

Carpenter, Jeffrey P., Samuel Bowles, Herbert Gintis and Sung-Ha Hwang, 2009, Strong reciprocity and team production: Theory and evidence, *Journal of Economic Behavior and Organization*, 71, 2, pp. 221–232.

Casari, Marco and Luigi Luini, 2009, Cooperation under alternative punishment institutions: An experiment, *Journal of Economic Behavior and Organization*, 71, 2, pp. 273–282.

Camerer, Colin F., and Ernst Fehr, 2006, When does "economic man" dominate social behavior?, *Science*, 311, 5757, pp. 47–52.

Chaudhuri, Ananish, 2011, Sustaining cooperation in laboratory public goods experiments: A slective survey of the litterature, *Experimental Economics*, 14, pp. 47–83.

van Dijk, Frans, Joep Sonnemans, Frans van Winden, 2001, Incentive systems in a real effort experiment, *European Economic Review* 45, pp.187–214.

El Ouardighi, Fouad, 2013, Attachment and forgiveness in a behavioral, dynamic non-cooperative game, In: El Ouardighi, F., and Kogan, K. (eds.): *Dynamic Methods and Applications, Essays in Honor of Charles S. Tapiero*, Springer International Series in Operations Research and Management Science, New York.

Eriksson, Tor, Anders Poulsen and Marie-Claire Villeval, 2009, Feedback and incentives: Experimental evidence, *Labour Economics*, 16, 6, pp. 679–688.

Falk, Armin and Ernst Fehr, 2003, Why labour market experiments?, *Labour Economics*, 10 4, pp. 399–406.

Falk, Armin, Ernst Fehr and Urs Fischbacher, 2005, Driving forces behind informal sanctions. *Econometrica,* 73, 6, pp. 2017–2030.

Falk, Armin and Andrea Ichino, 2006, Clean evidence on peer effects, *Journal of Labor Economics*, 24, 1, pp. 39–57.

Fehr, Ernst and Simon Gächter, 2002, Altruistic punishment in humans, *Nature*, 415, 10, pp. 137–140.

Fehr, Ernst and Simon Gächter, 2000, Cooperation and punishment in public goods experiments, *American Economic Review*, 90, 4, pp. 980–994.

Fehr, Ernst and Bettina Rockenbach, 2003, Detrimental effects of sanctions on human altruism, *Nature,* 422, pp. 137–140

Fehr, Ernst and John List, 2004, The hidden costs and returns of incentives - Trust and trustworthiness among CEOs, *Journal of the European Econonmic Association*, 2, pp. 743–771.

Fischbacher, Urs, 2007, z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* 10, 2, pp. 171–178.

Frey, Bruno S., and Reto Jegen, 2001, Motivation crowding theory, *Journal of Economic Surveys*, 15, 5, pp. 589–611.

Gneezy, Uri, 2003, The W effect of incentives. Mimeo, Chicago.

Gneezy, Uri and Aldo Rustichini, 2000, Pay enough or don't pay at all, *Quarterly Journal of Economics*, 115, pp. 791–810.

Herrmann, Benedikt, Christian Thöni, Simon Gächter, 2008, Antisocial punishment across societies, *Science*, 319, pp. 1362-1364.

Houser, Daniel, Erte Xiao, Kevin McCabe and Vernon Smith, 2008, When punishment fails: Research on sanctions, intentions and non-cooperation, *Games and Economic Behavior*, 62, 2, pp. 509–532

Holmstrom, Bengt, 1982, Moral hazard in teams, *Bell Journal of Economics*, 13, 2, pp. 324–340.

Itoh, Hideshi, 1991, Incentives to help in multi-agent situations, *Econometrica*, 59, pp. 611–636.

Kirstein, Annette, 2008, Bonus and malus in principal agent relations with fixed pay and real effort, *Fixed Wage*, 60, pp. 280–303.

Legros, Patrick, and Steven A. Matthews, 1993, Efficient and nearly efficient partnerships, *Review of Economic Studies*, 60, pp. 599–611.

Lin, Chung-cheng and C.C. Yang, 2006. Fine enough or don't fine at all., *Journal of Economic Behavior and Organization*, 59, pp. 195–213

Masclet, David, Charles Noussair, Steven Tucker and Marie-Claire Villeval, 2003, Monetary and non-monetary punishment in the voluntary contributions mechanism, *American Economic Review*, 93, 1, pp. 366–380.

McAfee, Preston, R. and John McMillan, 1991, Optimal contracts for teams, *International Economic Review*, 32, pp. 561-577.

Meidinger, Claude, Jean-Louis Rullière and Marie-Claire Villeval, 2003, Does team-based compensation give rise to problems when agents vary in their ability?, *Experimental Economics*, 6, pp. 253–272.

Mohnen, Alwine, Kathrin Pokorny and Dirk Sliwka, 2008, Transparency, inequity aversion, and the dynamics of peer pressure in teams: Theory and evidence, *Journal of Labor Economics*, 26, 4, pp. 693-720.

Nikiforakis, Nikos and Hans-Theo Normann, 2008, A comparative statics analysis of punishment in public-good experiment. *Experimental Economics,* 11, pp. 358–369.

Nikiforakis, Nikos, 2008, Punishment and counter-punishment in public good games: Can we really govern ourselves?, *Journal of Public Economics,* 92, (1-2), pp. 91–112.

Nikiforakis, Nikos, 2010, Feedback, punishment and cooperation in public good experiments, *Games and Economic Behavior*, 68, 2, pp. 689–702.

Vandegrift, Donald and Abdullah Yavas, 2011, An experimental test of behavior under team production, *Managerial and Decision Economics*, 32, 1, pp. 35–55.

APPENDIX

INSTRUCTIONS FOR THE WITH-SANCTION TREATEMENT[18]

Slide 1.

Good morning

Thank you for participating to this experiment.

Please read these instructions carefully and, should you have any question, raise your hand and call the administrator.

Communication between participants is forbidden. Please turn off cellular phones.

Slide 2

You will be required to perform an effort task jointly with a partner, during 4 identical rounds of the same experiment.

Each round lasts for 4 minutes; the clock starts when you will open the first active screen, and stops after 4 minutes. During a round, the remaining time is displayed in red characters, in the upper right corner of the screen (in seconds).

A euro payoff will be delivered at the end of the experiment. The payment is connected to performance in the task, according to a rule known by everybody.

Partners will be matched in pairs at random. Your partner will not change from one round to another. His identity will not be revealed to you.

During each round the computer displays one after the other blocs of figures (0 to 9) in six lines and 30 columns. Your task is to accurately count how many times the figure 7 appears in a bloc and then report this number in a box. The answer is considered correct if it corresponds to the right number of 7s in the bloc, with a tolerated error margin of 1. For instance, if the right number of 7s is 30, answers 29, 30 and 31 will be considered as correct.

Once the counted number of 7s is recorded in the box, you must press the "validate" button, to save it. At the click, a new random block of figures is automatically generated and the effort task can continue.

[18] **The no-sanction treatment and the piece-rate treatments are just simpler versions of the with-sanction treatment.**

At any moment you can take a break by pressing the button “Take a brake”. The break stops the counting task for 20 seconds; a screen with ESSEC logo appears. If the round stops in less than 20s, breaks are no longer possible.

At the end of each round the computer will display the total number of right answers that you have provided, and the total number of right answers provided by your partner.

At the end of each round, before moving to the next round, the player who provided the largest number of correct answers can, if he wants so, impose a fine on his partner. The decision belongs to him, it is not compulsory to impose the fine. In the case players have provided an identical number of right answers, no sanction is possible.

Slide 3. The example slide

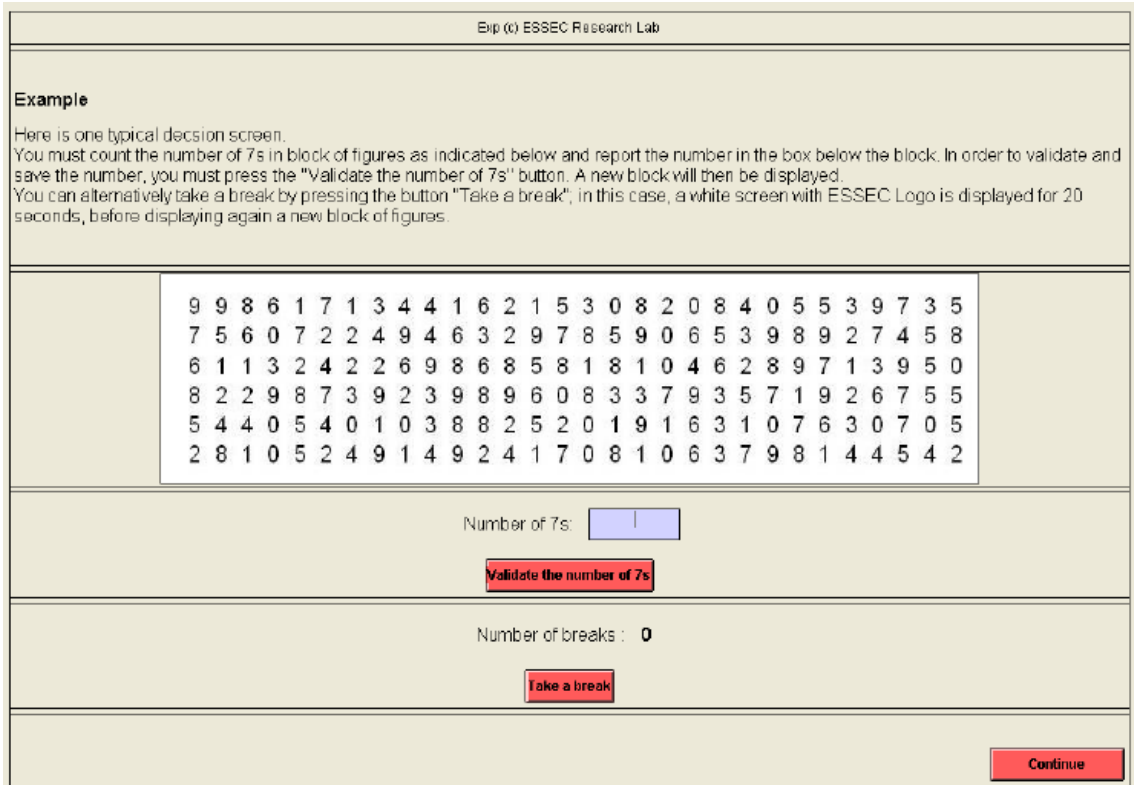

Figure 5:

Slide 4. Compensation rule

For each round, the payoff of one player is made up of three elements:

1. the compensation related to the effort task

2. a gain provided when taking a break

3. less the penalty (if any)

1. For each player, the ECU compensation related to the effort task is equal to half the total number of correct answers provided by the team during the round, times 10.

For instance, if player 1 brought 8 right answers and player 2 brought 5 correct answers, the gain for each one related to the effort task is 0.5x(8+5)x10=65 ECUs

2. For each 20 s break, you get 6 ECU, whatever your compensation for the effort task.

3. At the end of each round, before the next round starts, the player who brought the largest number of correct answers can, if he wants so, apply a fine to his partner, for an amount between 1 and 30 ECUs. The gain of the partner is reduced by that amount. No sanction is possible if players provide the same number of right answers.

At the end of the experiment, the total amount in ECU will be converted into Euros at the exchange rate 100 ECU = 2.5 Euros.

Slide 5. Check questions.

In order to make sure that you have understood well the rules of the game, please answer these questions:

Case 1.

During the round you got 4 right answers and your partner got 2 right answers. You took two breaks.

Your gain in ECU is:

(a) 0.5*(4+2)*10+2*6 ; (b) (4+2)*10 + 2* 6 ; (c) I do not know

Case 2

At the end of the round you got 4 right answers and your partner got 2 right answers.

Can you impose a fine on your partner?

(a) Yes (b) No (c) Don't know

Can your partner impose a fine on you?

(a) Yes; (b) No; (c) Don't know

Case 3

At the end of the round you chose to impose a fine on the partner. The amount of the fine can be:

(a) Between 1 and 10; (b) Between 1 and 30; (c) I do not know.

Slide 6.

Correct answers are:

Case 1

During the round you got 4 right answers and your partner got 2 right answers. You took two breaks.

The right answer is: you have half of the total points times 10, and the compensation for the breaks (2 x 6 ECUS), that is a total of 0.5(4+2)*10 + 2x6

Case 2

You got 4 right answers and your partner got 2 right answers.

Yes, you can impose a fine on him. Attention, this is an option; you do not need to impose a fine.

No, he cannot impose a fine on you

Case 3

At the end of the round you chose to impose a fine on the partner. The amount of the fine can be between 1 and 30. The payoff of the partner will be reduced by this amount.

Slide 7.

If you have any question, please rise your hand and address it to the administrator.

If you are sure to have understood the rules of the game, you can press the button below in order to launch the experiment.

The experiment starts when all subjects have pressed the button.

Slide 8. (main decision screen)

Similar to "Example" in Slide 3 (but without the text on top of the screen; and the timer on the right upper corner).

Slide 9. Results on task

Your performance: Number of counted blocks[ ], Number of right answers[NRA1], Number of breaks [ ]

The performance of your partner: Number of right answers [NRA2].

[If NRA1>NRA2 the computer displays] Do you want to impose a fine on the partner ?

Yes // No

If you click YES the computer displays "chose the amount of the fine" [A=1 to 30]

Validate

Slide 10. Payoff for the round

Your partner has imposed a fine on you (did not impose a fine on you)

The amount of the fine is: [A]

Your payoff for the round is: [ ]

Slide 11. – At the end of the experiment

Thank you for having participated to this experiment.

The total gain for the experiment is [ ]

ESSEC Business School
Avenue Bernard Hirsch
BP 50105
95021 Cergy-Pontoise Cedex
France

Tél. +33 (0)1 34 43 30 00
Fax +33 (0)1 34 43 30 01

www.essec.fr

ESSEC Executive Education
CNIT BP 230
92053 Paris-La Défense
France

Tél. +33 (0)1 46 92 49 00
Fax +33 (0)1 46 92 49 90

http://formation.essec.fr

ESSEC Business School
Singapore Campus
100 Victoria Street
National Library Building # 13-02
Singapore 188064

essecasia@essec.fr
Tél. +65 6884 9780
Fax +65 6884 9781

www.essec.edu

Informations

Alison Bougi
+33 (0)1 34 43 33 58
bougi@essec.fr
www.essec.fr
research.center@essec.fr

ISSN 1291-9616

ESSEC BUSINESS SCHOOL

CCI À vos côtés, précisément. AACSB ACCREDITED EFMD EQUIS ACCREDITED