# Creativity and Incentives

Gary Charness* and Daniela Grieco+

**Abstract:** Creativity is a complex and multi-dimensional phenomenon with tremendous economic importance. A crucial question for economists and for firms is the interplay of incentives and creativity. We present experiments where subjects face creativity tasks where, in one case, ex-ante goals and constraints are imposed on their answers ("closed" tasks), and in the other case no restrictions apply ("open" tasks). The effect of tournament incentives on creativity is then tested. Our experimental findings provide striking evidence that financial incentives in the form of tournament competition affect creativity in "closed" (constrained) tasks, but do not facilitate creativity in "open" (unconstrained) tasks, whereas being ranked relative to one's peers is an effective non-monetary incentive with both types of tasks. We develop a structural model that allows for subjects' heterogeneity in being affected by the openness of the task, and then use the structural model to not only estimate creative output in tournaments but also to predict creative output in two counterfactual incentives schemes: piece rate and target bonus.

**Keywords:** creativity; experiment; structural estimation; incentives; constraints.

**JEL codes:** C91, D03, O39

* *UCSB*

+ *Bocconi University*

**Acknowledgments:** For valuable comments we thank Doug Bernheim, Anat Bracha, Christiane Bradler, Francesco Bripi, Alessandra Casella, Vince Crawford, Stefano DellaVigna, Catherine Eckel, Ernst Fehr, Drew Fudenberg, Uri Gneezy, Peter Hammond, Botond Koszegi, Nicola Lacetera, Christine Lauritzen, John List, Ulrike Malmendier, Meg Meyer, Susanne Neckermann, Charlie Plott, Simon Quinn, Antonio Rangel, Luca Rigotti, Kelly Shue, Bruno Strulovici, Nora Szech, Verena Utikal, Emanuel Vespa, Marie Claire Villeval, Alistair Wilson, and participants at seminars at the UCSB Psychology Department, UCSD, Columbia University, the University of Pittsburgh, ZEW Mannheim, the 2014 UCSB Conference on Behavioral and Experimental Economics, the 2014 Spring School on Experimental Economics at UCSD, the 2014 Experimental Methods and Policy Conference in Curacao, the University of Oxford, the 2014 Stanford Institute for Theoretical Economics, the 2014 CESifo conference in Munich, the University of Lyon (GATE), and the 2015 FSU Workshop on Experimental Economics and Entrepreneurship. We thank Madeline Kardos, Tyler Martin, and Zeke Wald for valuable research assistance.

# 1. Introduction

Creativity is a vital input into the success of a society, contributing in economic, social, and aesthetic dimensions. According to Henri Poincaré, creativity represents the "ability to unite pre-existing elements in new combinations that are useful". Creativity implies (a) a *combination* of existing things that should be (b) *recognized in its utility by peers* (Mumford, 2003).

Historically, prizes have been used to stimulate many discoveries. These include Archimedes' method for measuring the volume of the king's crown, the canning process to preserve food needed by Napoleon's troops, the invention of margarine that was triggered by Napoleon III, who offered a prize to any chemist who would develop a cheap butter substitute to feed France's armies, and the smallpox vaccine that was developed in pursuit of a financial prize offered by the English Parliament. In the same vein, the patent system was developed with the aim of providing a strong incentive to produce novel ideas and products without the gains from doing so being appropriated by other entities.

Creativity extends into expressive and performance activities such as art, music, dance, and writing. It may be the case that artists and even academic researchers do not need financial incentives to produce creative art or research, respectively: perhaps ideas arrive at their own rate, independently of direct incentives.[1,2] On the other hand, environmental factors might be crucial in stimulating creativity: "All human societies contain inventive people. It's just that some environments provide more starting materials, and more favorable conditions for utilizing inventions, than do other environments" (Diamond, 1997, p.

[1] Kremer (1993)'s model assumes that each person's chance of inventing something is not affected by interaction with others. Even if individual research productivity is independent of population, total research output is shown to increase in population due to the non-rivalry in technology.

[2] In fact, until 2007, French artists could benefit from a form of subsidy that was reserved to the so-called "*intermittent du spectacle*" (an arts and entertainment industry worker who receives payments and benefits during periods of unemployment), that was aimed to sustain French culture and that has been criticized for being unable to promote quality.

408). There is also the notion from social psychology (see for example the seminal work by Deci and Ryan, 1985) that extrinsic reward can crowd out intrinsic motivation, so that providing financial rewards can be counterproductive.[3] Thus, the effect of economic incentives on creativity is not clear *ex ante*. Empirical evidence is clearly needed, and there is very little work linking creativity and incentives. Since it is difficult to perform clean tests using field data, controlled experiments provide a promising avenue for exploring this issue.

We wish to emphasize that there are many different conceptualizations of creativity, and that the way that we parse creativity is just one approach. We consider a particular dimension of closed creativity tasks versus open creativity tasks, depending on how well-defined is the task at hand. In creativity problem-finding research (e.g. Runco, 1994), scholars examine the degree to which the problem has been formulated before the creator begins the process. With closed creativity tasks, there is a specific and delineated goal. Examples could be finding a way to decrease the size of a computer or developing a new drug for a specific purpose.[4] Instead, open creativity tasks could be painting an abstract painting, representing unfettered thinking outside the box without any obvious underlying *ex-ante* goal or direction.

Our first contribution is to provide an experimental test of the effects of incentives on creativity. In our between-subject design, we vary whether a task is "open" or "closed" and randomly assign participants across sessions to either receive a flat payment for completing the task (with their performance ranked relative to peers or not ranked, according to the treatment) or be paid according to tournament incentives.

---

[3] Hennessey and Amabile (2010) review the creativity literature in social psychology and state (p. 581) that historically: "High levels of extrinsic motivation were thought to preclude high levels of intrinsic motivation; as extrinsic motivators and constraints were imposed, intrinsic motivation (and creativity) would necessarily decrease."

[4] An interesting case is represented by InnoCentive, a crowdsourcing company based in Massachusetts that accepts commissioned research and development problems in a number of fields like engineering, computer science, mathematics, and so on: these problems are framed as "challenge problems" for anyone to solve, and the best solutions that meet the challenge criteria receive cash awards.

Our results indicate that monetary incentives are effective in stimulating creativity when and only when *ex-ante* goals are specifically set and the nature of the task is more well-defined ("closed" tasks). So when society has a clear objective in view, it does appear useful to reward creativity that helps to achieve this objective. On the other hand, incentives for performance with respect to open creativity tasks provide no benefit in our setting.[5] To the best of our knowledge, there has been no previous evidence regarding the relative benefit of incentives depending on the type of creativity task involved. We also find that peer ranking, which has been previously found to lead to increased effort, works as a form of non-monetary incentive that fosters creativity in *both* types of tasks.

Our second contribution consists of presenting a structural model that derives individual optimal creativity effort under four monetary payment schemes: 1) flat payment (with and without peer ranking), 2) tournament incentives based on a performance ranking within a group of peers, 3) piece-rate incentives where individuals are paid according to their "creative output", and 4) target bonus where only individuals who reach a given threshold of creativity output are paid. The heart of the model is in viewing individual's creativity effort as being driven not only by idiosyncratic factors, but also by peer ranking and financial incentives. Although peer ranking motivates people, regardless of the type of task, to exert effort in order to avoid a poor creative performance relative to one's peers, the effectiveness of financial incentives in stimulating higher creative output crucially depends on the precision of the task definition and goals: we show that people are uncertain as to what to do when the task is open, so that incentives are less effective. Giving a well-defined structure to the task lowers this uncertainty and allows scope for incentives to effect behavior.

[5] However, an exception applies to ambiguity-averse people, who tend to otherwise avoid the less-defined open-creativity tasks. Even though ambiguity-averse individuals might be very creative in such tasks, the uncertainty surrounding them might cause them to simply not take them on. Incentives could potentially overcome this reluctance, so that ambiguity-averse people might effectively be influenced by extrinsic rewards in this context.

Our third contribution is to structurally estimate the model, allowing us to account for individual heterogeneity in how the openness of the task affects the effectiveness of incentives. The estimation is carried out using both experimental data on flat payment and tournament, as well as simulated data on the counterfactual incentive schemes (piece rate and bonus).

The remainder of this paper is organized as follows. We discuss related literature in Section 2, and illustrate the experimental design in Section 3. Section 4 shows the experimental results. The structural model and estimation are presented in Section 5. Section 6 provides a discussion and Section 7 concludes.

## 2. Related literature

### 2.1. Definitions and dimensions of creativity

Until the middle of the 20th century, creativity was studied as a minor topic within a number of various disciplines such as psychology, sociology, and cognitive science. The turning point for the emergence of creativity as a separate sphere of study can be traced back to the seminal works of Guilford (1950) and Torrance (1962, 1974, 1989), who attempted to measure creativity from a psychometric perspective. The Torrance test of creative thinking compares "convergent" to "divergent" thinking and is still a reference tool for measuring creativity.

For our purpose, creativity can be defined as "the production of novel and useful ideas in any domain" (e.g. Stein, 1974; Woodman et al., 1993). In contrast, innovation represents the successful implementation of creative ideas within an organization; creativity by individuals and teams is therefore the starting point for innovation. Given the necessity of generating creative ideas repeatedly, firms have traditionally relied on an internal staff of professional inventors in R&D labs (Schulze and Hoegl, 2008). More recently, many organizations have turned to employee suggestion schemes (Ohly et al., 2006) or to outsourcing of creative ideas in an attempt to get fresh hints (Surowiecki, 2004).

Galenson (2004)'s research on creativity identified two creative methods or styles: conceptual and experimental. The former relates to the generation of a new idea (a kind of deductive process), the latter is a new combination of existing items (an inductive or synthetic process that relies on experience). Convergent tasks call for a single correct response, whereas divergent tasks involve potentially producing many different correct answers (Hudson, 1966; Runco, 2006 and 2007). Although creativity tasks are usually categorized as either convergent or divergent, most creative problems contain elements of both (Nielsen et al. 2008).

As emphasized by Unsworth (2001), in general researchers propose a continuum ranging from closed to open problems (Getzels and Csikszentmihalyi, 1967; Wakefield, 1991; Eysenck, 2003): “a true closed problem is one that is presented to the participant, when the method for solving the problem is known […]; open problems occur when the participant is required to find, invent, or discover the problems” (Unsworth, 2001, p.290). In Torrance’s terminology, the former case suggests convergent thinking, whereas the latter suggests divergent thinking. In the perspective of Dual Process Theory (e.g. Stanovich and West, 2000; Kahneman, 2011), closed tasks – characterized by specific directions to follow – could exert a signal to use the cognitive system to proceed rationally, slowly and according to logical standards. Open tasks, on the contrary, cause one to proceed in a much more unplanned and unaware manner. In a similar flavor, Unsworth (2001) argues that closed problems require responsive and contributory creativity, whereas open problems require expected and pro-active creativity. Dillon (1982) argues that most artistic endeavors generally represent open problems; responses to a suggestion scheme illustrate outcomes of organizational open problems.

In the context of business innovation, a closed task might be represented by the enhancement of a technology process that solve problems, eliminate flaws or reduce costs, or by the refinement of an existing product. On the other hand, an open task might imply the development of a new idea, product or process without a specific aim or problem to solve.

Our creativity tasks have some overlap with previous definitions of conceptual versus experimental creativity: the open task requires the development of something new, without restrictions on what can be done; the closed task involves combining a set of elements, which also act as constraints. In the open task there are no specific instructions to follow on how to get to a “creative output”, so that the goal to reach is perceived as vague; on the contrary, in closed tasks the precise definition of the constraints makes the goal easier to grasp. In this aspect, the more a task is open, the closer it is to divergent thinking, since more possible interpretations of the goal to reach are available. An interesting example of incentives to "open" vs. "closed" creativity tasks in the realm of academic life-sciences funding is represented by Azoulay et al. (2011)'s study of the careers of investigators of two health institutes: the former gives wide freedom to explore, tolerates early failure, rewards long-term success, whereas the latter gives investigators multiple sources of constraints, imposing short review cycles, predefined deliverables, and renewal policies that do not forgive failure.

### 2.2. Motivation and incentives to creativity

A big question underlies involvement in the creative process. *Why* do people engage in creative activity? Motivations might depend on internal sources, such as a need for self-actualization or simply the joy one receives from being creative. In general, skills like tenacity, self-discipline and perseverance are important traits for success in life (Heckman and Rubinstein, 2001); since intrinsic motivation enhances self-efficacy (Walls and Little, 2005), individuals with higher levels of such skills are expected to exert greater effort and be more engaged in creative tasks.

Alternatively, creative behavior might be a response to an external demand perhaps reflecting a job description, an experimental requirement, or environmental needs. Both intrinsic and extrinsic motivations appear to play roles as determinants of creative behavior. A number of studies show the importance of high intrinsic motivation consisting of the excitement and challenge of engaging in a creative activity. On the other hand, there is little agreement among scholars on the effectiveness of financial incentives (and, more generally, rewards and extrinsic motivations) on creative performance.

Despite the conventional wisdom in economics, financial incentives are not always helpful and may even be counterproductive. Deci and Ryan (1985) report an experiment in which children's intrinsic motivation to engage in an activity is undermined by financial rewards. Similarly, Gneezy and Rustichini (2000) show that paying only a small wage for charitable work can lead to lower productivity than relying completely on intrinsic motivation and paying nothing. Paying an excessive amount can also lead to poor outcomes due to a sense of pressure, as suggested by the aforementioned results in Ariely et al. (2009). Literature on reference earnings (e.g. Pfeffer and Langton, 1993) has shown that increasing stakes with high degree of dispersion among workers may lead to lower satisfaction and productivity.

Amabile (1989, 1996)'s seminal studies both on children and adults show that crowding out can occur in the presence of monetary incentives, which seem to undermine intrinsic motivation and affect creative performance negatively. According to Kohn (1993): "It is simply not possible to bribe people to be creative" (p. 294). In the same vein, Hennessey and Amabile (1998) conclude "the preponderance of the evidence demonstrates that working for reward, under circumstances that are likely to occur naturally in classrooms and workplaces every day, can be damaging to both intrinsic interest and creativity" (p. 675).

Nevertheless, some empirical research shows positive effects of rewards on creativity (Eisenberger *et al*, 1998; Eisenberger and Rhoades, 2001), although these results seem to be driven by very specific contexts or derived under experimental conditions not fully consonant with the methods of modern experimental economics. The use of reward has been possibly confounded with the presence of cues indicating the appropriateness and desirability of a creative performance (Winston and Baker, 1985). In addition, many studies in psychology use the *promise* of a reward (aimed at establishing reward expectancy).[6]

Collins and Amabile (1999) show that rewarding children's creativity can be successful if combined with intensive cognitive training designed to encourage a focus on the assigned task rather than on the reward. Crucially for our investigation, financial incentives are shown to lead to enhanced performance when the pattern of solution is clear and straightforward, i.e. in what Collins and Amabile (1999) call "algorithmic" task (such as, for instance, making a collage after being told precisely how to make a creative one). Closed tasks may be viewed as those in which the path to a solution is apparent and people have learned how to generate solutions. McCullers (1978) highlights that incentives increase

[6] However, it is possible that the credibility of these promises was undermined by an expectation of deception.

performance when this involves making “simple, routine, unchanging responses,” (p. 14) but that incentives are less effective in situations that depend on flexibility, conceptual and perceptual openness. McGraw (1978) identifies two conditions under which incentives will even have a detrimental effect on performance: “first, when the task is interesting enough for subjects that the offer of incentives is a superfluous source of motivation; second, when the solution to the task is open-ended enough that the steps leading to a solution are not immediately obvious” (p. 34).

Erat and Krishnan (2012) develop a model to examine the relationship between problem specification, award structure, and breadth of solution space for a firm that manages a contest for outside agents working on the solution of open-ended problems: the model predicts that, as the problem becomes better specified, the searchers will perceive limited risk that their evaluation of solution quality does not match the principal’s evaluation, and thus even small prizes can induce search. Both Collins and Amabile (1999)’s and Erat and Krishnan (2012)’s contributions suggest that, whereas monetary incentives might promote creativity in closed tasks, directly incentivizing open creativity would be ineffective or even counter-productive. The prediction that incentives should hurt idea generation seems consistent with the assumption that – in open tasks – creativity relies primarily on random variations in the search process. However, in closed tasks, structure (and not randomness) is the key to creativity (Goldenberg et al., 1999), so that idea generation appears as a more algorithmic task, likely to be enhanced by incentives.

### 2.3. Literature in experimental economics

Many real-effort tasks have been used in experimental economics in recent years (for reviews, see Charness and Kuhn, 2011 and Alexseev, Charness, and Gneezy, 2017). Some of these involve solving a puzzle with a specific and clear insight that may not be immediately obvious and that is supposed to lead to the (unique) solution of the task: in our terminology, this is an example of a fully-closed task. Rütstrom and Williams (2000) used a Tower of Hanoi puzzle, which involves three rods and a number of disks of different sizes that can slide onto any rod. The puzzle starts with the disks in a neat stack in ascending order of size on one rod, the smallest at the top, thus making a conical shape. The objective of the puzzle is to move the entire stack to another rod, following some simple rules. Ariely et al. (2009) used the "Packing Quarters" game: participants are asked to fit nine metal pieces of quarter circles into a frame within a given time. To fit all nine, the pieces must be packed in a particular way.[7] Toubia (2006) recruited subjects at an anti-war walkout and asks them to generate ideas on a specific problem: "How can the impact of the UN Security Council be increased?". This can still be considered a closed task, since there is the constraint of reaching a specific goal, but less so than the Tower of Hanoi puzzle. Toubia finds that incentives lead participants to try harder to generate ideas and to give up less easily.

There are several recent experimental papers (some of which are contemporaneous to our experiments) that consider aspects of incentives on creativity. Overall, the results are rather mixed. Chen et al. (2012) examine whether the efficacy of either individual- or group-based creativity-contingent incentives depends on the form (piece-rate or tournament) they take: individual intragroup tournament pay increases individual efforts, but is not effective in enhancing the creativity of group solutions relative to individual piece-rate pay: reward systems result to be more likely to promote group creativity through collaborative efforts rather than independent individual efforts. Eckartz et al. (2012), in a within-subject design, ask subjects to form words out of letters under three incentive schemes: a flat fee, a linear payment and a tournament; they also use two control tasks (the Raven test and adding numbers). There was no real effect of any incentives on performance. They also find no effect of gender on tournament entry, in contrast to Niederle and Vesterlund (2007).

Bradler et al. (2016) compare the effects of financial incentives on performance on a routine task and a creative task. The routine task is the slider task (Gill and Prowse, 2012), while the creative task is the "Unusual Uses task" (Guilford, 1967; Torrance, 1968) – subjects are asked to name as many different and unusual uses as possible for a routine object. The payoffs are structured as a tournament prize for above-

[7] Here very large financial incentives led to poorer performance than did more modest stakes.

average effort. They find that tournament incentives work well and have similar effect sizes in these tasks, and estimate that concern for relative rank accounts for about one-fourth of this effect.

Erat and Gneezy (2015) examine the effect of piece-rate and competitive incentives (as well as two different time limits) on performance on a task involving a rebus. This is "a puzzle made with words and/or pictures with a hidden and non-obvious solution," so that there is a unique correct response. Even though financial incentives lead to greater effort (time spent on the rebus), incentives do not improve the creative output relative to the case in which participants are not offered any external monetary incentives for creativity; moreover, the type of incentives matter, and competitive incentives reduce creativity relative to piece-rate incentives. This finding might reflect that the specific "problem-solving" task chosen has a unique correct solution that leaves limited room for creative answers and, as the authors suggest, might generate pressure.

Laske and Schroeder (2016) study the effect of incentives on different dimensions of creative work, with incentives either for quantity alone or in combination with usability or novelty. They compare performance to that in a baseline treatment with fixed pay. Incentivizing quantity alone or quantity in combination with novelty results in an increase in quantity and novelty, but decreases the average quality compared to the baseline. Combining incentives for quality and quantity does not significantly affect any of the dimensions of creativity.

Our design differs from each of these studies in that we test for the effect of incentives with two different forms of creativity tasks that differ only in that one is somewhat more closed than the other. We are unaware of any study that considers the effect of peer ranking and financial incentives on two different types of creativity tasks. Our experimental tasks do not have a unique and correct solution, so may reflect a richer form of creativity, and they allow a full range of open-ended personal expression.

## 3. The experiment

Our experiment involves asking individuals to perform a task in a creative manner. The experiment has a 2x3 design, consisting of two real-effort tasks (closed vs. open) and three treatments (Flat payment and

Tournament incentives, plus the No-ranking control). Each participant was assigned to only one of the six conditions. The relative creativity of each participant is evaluated by peers (with the exception of the No-ranking treatment) and by external judges - blind to treatments and conditions - in line with Poincaré's definition emphasizing that the "new combination" should be recognized in its utility by peers and with Amabile's notion of "social consensus". Instructions sheets are enclosed in Appendix A. We provided no guidance concerning evaluating creativity; when asked, we simply stated that this was for each participant to gauge.

### 3.1. Tasks

We capture closed creativity tasks by using "combination" tasks, and open creativity tasks by asking for the development of a totally new product or vision. Subjects had 20 minutes to complete the chosen task. While these tasks are consistent with our notion of open versus closed creativity, we certainly do not claim that these specific tasks are fully representative of all dimensions of creativity. In fact, as mentioned above, we chose our tasks in part to be a modest difference on the open-versus-closed dimension.

*Closed task*

In the closed condition, people were asked to choose from the following questions: [8]

1. "Choose a combination of words to create an interesting story." The words supplied are: house, zero, forgive, curve, relevance, cow, tree, planet, ring, send. Participants were told to use these words along with any other combination of words that they wished.
2. "Starting from the number 27, obtain the number 6 by using at least two different numerical operations. Possible answers include: (27:3) – 3 = 6, or [(27 + 3): 2 – 12]! = 6."

[8] We gave subjects the possibility of choosing the task because we wanted them to be more likely to face an endeavor with which they were comfortable. Although restricting to one task only has the advantage of making the comparison among subjects easier for judges, our aim was not testing subjects' ability, but their talent to create. Since people do differ in their attitude towards verbal and math tasks, we wanted to avoid poor creative performance due to subjects being negatively disposed towards one compulsory task. In the analysis, we control for the task chosen and show that it does not affect our results, see Table 2 (column 4) and Table 3 (column 3).

In both cases, subjects had to be creative but faced constraints: although they could write the story they liked and use the mathematical operations they preferred, the task compelled them to use a given set of words or to start/end with a given number, respectively.

*Open task*

In the open condition, people were asked to choose from the following questions:

1. "If you had the talent to invent things just by thinking of them, what would you create?"
2. "Imagine and describe a town, city, or society in the future."

Unlike the closed task, no constraints of any form were imposed in this case.[9]

With the exception of the No-ranking treatment, participants were told that the creativity of their output would be ranked in relation to that of the other four people in the group. People in another mutually-anonymous five-person group (in order to avoid strategic effects on the evaluations) performed this ranking.

**3.2. Treatments**

*No-ranking control*

We paid people a flat amount of $9 (plus the $5 show-up fee) for completing the response. The tasks were as described above. Tasks were the same as in the other treatments. We paid people a flat amount of $9 (plus the $5 show-up fee) for completing the response. No ranking of the creativity of their work

[9] We acknowledge that we did not offer an open mathematical option (due to our own limited imagination) in the open task. While we suspect that this did not have a large impact on the effectiveness of peer ranking and tournament incentives on creative performance (see Table 2, column 4), a reviewer makes the excellent point that since we did not include an open mathematical option, we cannot rule out that our results reflect variation in how people respond to incentives when they have a mathematical option.

occurred. As above, we had ex-post evaluation by external judges, who were blind to treatments and conditions.

*Pure ranking treatment*

Here people were told that the five individuals in another group would anonymously rank the creativity of their work, although their pay for their output was a flat $9. People have been shown to care about their rank *per se*, as in Charness et al. (2014). Since peer ranking may well also affect subjects' creative performance, we isolate its pure effect relative to the no-ranking control treatment.

*Tournament treatment*

Here we paid people on the basis of the assessments made. People were told that the five individuals in another group would anonymously rank the creativity of their work. In each group, the person with the best ranking received $15, the second-best received $12, the third-best received $9, the fourth-best received $6, and the worst received $3; these payments were made in addition to the standard $5 payment for showing up on time for the experiment). We note that these are relatively "soft" tournament-style incentives, with a marginal change in earnings of only $3 per ranking. The average earnings were the same as in the other treatments.

One may wonder why we paid subjects according to how they are ranked relative to their peers. One practical consideration is that we wished to pay them at the time of the session and having raters come to the session would have led to considerably longer sessions and would have also made it impossible for one to rate creativity in more than one treatment (since they would no longer be blind to the treatment). Second, a major advantage of having evaluations by peers is that they are most likely to be attuned to what is perceived to be creative in the relevant reference group (recall that "creativity should be recognized in its utility by peers"). Of course, to perform rankings across sessions, it was necessary to later have the responses evaluated by external judges, who were blind to treatments and conditions. As we shall see, the correlation across rankings by students and raters was high.

**3.3. Questionnaire**

In our questionnaire (presented after completing the task), we requested demographic information and also asked subjects to answer two incentivized questions on risk and ambiguity attitude (Gneezy and Potters, 1997; Charness and Gneezy, 2010): each individual is endowed with 100 units and could invest any portion in a risky asset that had a 50% chance of success and paid 2.5 times the amount invested if successful and nothing if unsuccessful; the individual retains whatever units were not invested. This procedure provides a measure of risk aversion for each individual: the higher the investment, the less risk averse is the individual. The question on ambiguity attitude is identical except that we did not tell people the probability that the investment would be successful (which was 25%, requiring two successful coin flips).

The questionnaire also included 10 questions on creative and cognitive style and sensation-seeking attitude, based on Nielsen at al. (2008)'s questions on creative style and on Zuckerman et al. (1964)'s questions on sensation-seeking attitude, as well as seven questions on demographic features: gender, age, major, number of siblings, birth-order, right or left-handed, married/divorced/unmarried parents plus other six questions on past involvement in creative activities, as in Hocevar (1980). The full questionnaire is shown in Appendix B.

**3.4. Procedures**

The experiments were conducted at the University of California, Santa Barbara. There were 18 sessions, with a total of 328 participants. There were 138 people who faced the closed task, with 53 in the pure-ranking condition, 45 in the tournament condition, and 40 in the no-ranking condition; there were 190 people in the open task, with 70 in the pure-ranking condition, 69 in the tournament condition and in the 51 in the no-ranking condition.[10] The subjects were undergraduate students (42% from Social Sciences, 42% from STEM disciplines and 16% from Humanities), with 56% females. We employed a between-subjects

[10] The number of people attending the experiment was 330, but we end up with 328 answers because: (a) one subject in Session 6 did not give us back the sheet containing his answer; (b) two subjects in Session 14 did not show up, (d) Session 17 had one extra subject, with a group of six people instead of five (but there was no group-ranking).

design (no one participated in more than one session). Participants were paid a $5 show-up fee, plus their earnings from the experiment. Written instructions were distributed to the participants and were also read aloud by the experimenter. All subjects completed a final questionnaire containing demographic information, personality details, and the two incentivized questions measuring risk and ambiguity aversion. The sessions took one hour, with average earnings of $15.

## 4. Experimental results

### 4.1. Creativity evaluation

As mentioned, in the Flat-payment and the Tournament-incentive treatments people in one group evaluated and ranked the individual responses from people in another group: subjects received no specific criteria to follow but their own taste for creativity and made a general ranking, no matter the task chosen. Rankings exhibit a fair degree of consistency (Cronbach's alpha = .536; Cronbach's alpha computes the inter-item correlations or co-variances for all pairs of judges' evaluations). To make comparisons across treatments, we had two external judges – blind to treatments – assess all of the answers on a 1-10 scale. As with peer ranking, external raters received no indication of any specific criteria to be followed and no guidance regarding comparing across the two tasks. The judges' evaluations exhibited a good degree of consistency (Cronbach's alpha = .713; this was .758 and .690, respectively, for closed and open tasks separately.[11,12] Our creativity score is the average of the two independent evaluations and is highly correlated with the group ranking (in the cases of the Pure-ranking and Tournament treatments: Spearman correlation test, with coefficient = .518, $p$ = 0.000).[13]

[11] It is interesting to note that consistency between judges is higher with closed tasks, perhaps because it is easier for them to identify the goal pursued.

[12] We also calculated consistency using Interclass Correlation Coefficients and obtained very similar values (details available upon request).

[13] Throughout the paper, we round all $p$-values to three decimal places.

In addition to subjective evaluations, we introduce two objective measurements of creativity effort that reflect the number of words used in the answers to the verbal task and the number of operations used in the answers in the math task: these measures unequivocally capture the "size" of the creative output and are summarized by a variable labeled "count". With closed creativity, participants used an average of 191 words (SD = 92.18) or 70 mathematical operations (SD = 70.40); with open creativity, answers had an average of 211 words (SD = 84.61).

For the closed task, we had two other judges (different from the ones who assigned the creativity score) classify the answers according to the two-fold taxonomy shown in Appendix C. For the verbal task, the judges used this taxonomy to identify the specific meaning according to which each of the words the subjects had used, and to assign a score reflecting the degree of originality of the meaning they selected.[14] For the math task, the judges assigned a score that reflected the complexity of each operation used.[15] In the case of the closed verbal creativity task, participants obtained an average taxonomy score of 19.66 (std. dev. = 12.13); for the closed math creativity task, the taxonomy score was on average 37.76 (std. dev. = 61.08).

### 4.2 Descriptive statistics

Figure 1 provides a graphic summary of the creative scores in each task (open vs. closed) in the three treatments.

[Figure 1 about here.]

[14] Meanings are ordered according to WordReference.com's ranking in use frequency: for each word, the score increases in the originality of the meaning used.

[15] Operations are grouped and ordered according to the school level in which they are typically taught: the subject earns the score corresponding to the maximum level she reaches, no matter the number of operations in each set.

Figures 2a and 2b show the distributions of creativity scores by comparing creativity scores with and without ranking (Figure 2a) and the distributions of creativity scores by comparing creativity scores with and without tournament incentives (Figure 2b): the ranking appears to flatten the creativity distribution (only in the case of the closed task) and shift it on the right (for both closed and open tasks), whereas the presence of tournament incentives shifts the distribution further to the left, but only for closed tasks.

[Figures 2a and 2b about here.]

Table 1 reports summary statistics according to the task and the treatment. It shows a dramatic and distinctive ranking effect on average creativity with both types of tasks, and a strong affect from tournament incentives with the closed task.

[Table 1 about here.]

**4.3. Role of ranking**

Creativity scores are significantly higher when subjects are peer-ranked, even when there is only a flat payment. We compare the No-ranking control and the Pure-ranking treatments, and find that this is true for both closed and open tasks. For the closed task, the average level of creativity score increases significantly with ranking from 3.625 to 5.075 (Wilcoxon rank-sum test on individuals, $Z = -4.603$, $p = 0.001$). With the open task, the average level of creativity score increases significantly with ranking from 4.068 to 5.150 (Wilcoxon rank-sum test on individuals, $Z = -3.880$, $p = 0.001$). Being ranked by peers serves as a form of non-monetary incentive that fosters creativity no matter the degree of openness of the task.

This effect from purely being ranked is the largest effect of all. It is quite clear that simply being ranked (even anonymously) is a powerful motivational force with regards to creativity. Our result on the role of ranking is in line with previous experiments (e.g., Charness et al., 2014).[16] However, to the best of our knowledge, ours the first study to find such an effect of financial incentives on creativity.[17] Section 4.6 below reports decile regressions and Wald tests showing that this effect is present across deciles.

### 4.4. Role of tournament incentives

The introduction of tournament incentives has a positive effect on the level of creativity when the task is characterized by the presence of *ex-ante* goals and constraints. In the closed condition, participants whose pay depended upon their ranking[18] are more creative than subjects who receive only a ranking: the average creativity score increases with tournament incentives from 5.075 to 5.909 and this difference is significant (Wilcoxon rank-sum test on individuals, $Z = -2.673$, $p = 0.007$).[19] There is no significant difference between the creativity scores for incentivized and non-incentivized open tasks or that of non-incentivized closed creativity.[20,21]

---

[16] However, we note that Charness et al. (2014) found that status-seeking also had negative effects (e.g., sabotage).

[17] Erat and Gneezy (2015) find that competition is generally ineffective in stimulating creativity: since they use a unique-solution task, their result may reflect the pressure generated by having subjects compete for reaching the only correct solution. Eckartz et al. (2012) use a scrabble-type task, finding that incentives have very small effects; differences in performance are predominantly related to individual skills. On the contrary, Bradler et al. (2016) provide evidence that routine as well as creative task performance increase significantly under the tournament scheme, whereas unconditional gift triggers higher effort only in tasks while creative performance is not affected.

[18] Our payoff structure is a relatively soft tournament scheme. We might expect to find even stronger results with sharper marginal differences in payoffs.

[19] All statistical tests are two-tailed unless otherwise specified.

[20] The respective test statistics are $Z = 0.532$, 0.298 and -0.195, with $p$-values 0.594, 0.832 and 0.845.

[21] The closed condition gives subjects a choice between a math task and a verbal task, while the open condition only offers a choice of verbal tasks. In principle, this could lead to higher creativity scores in the closed condition, since

It is worth noting that, whereas tournament incentives matter in determining the creativity score, the taxonomy score is not improved either by tournament incentives or by peer ranking.[22] On the contrary, subjects' effort (captured by the "count", i.e. number of words or math operations) increases with tournament incentives and with peer ranking, but only in the case of verbal tasks. Our general interpretation of these mixed results is that it is the subjective component of creativity evaluation – difficult to capture by means of objective criteria such as the ones we introduced – that is fostered by tournament incentives and by peer ranking. Other explanations can be related to a lack of "universal" effectiveness of these kinds of incentives in promoting the quality of the creativity effort instead of the creativity effort itself.

In the open condition, the average creativity score with tournament incentives is not significantly different from that without incentives: the average level of creativity score is 5.079 and 5.150, respectively (Wilcoxon rank-sum test on individual averages, with $Z = 0.532$, $p = 0.594$). Note also that the creativity scores in both parts of the open condition are nearly the same as the score in the closed condition without

people who are more math-oriented might score higher on a creativity task that is mathematical. However, in fact the scores on the math task were not higher than those on the verbal task. Note that there is no difference in the average creativity score across open and closed non-incentivized conditions.

[22] In the verbal closed task, the average taxonomy score is 24.88 with no-ranking, 19.61 with flat payment and 16.75 with tournament (No-ranking vs. Flat Payment: Z = 0.371, p = 0.710; Flat Payment vs. Tournament: Z = 1.240, p = 0.251, Wilcoxon rank-sum tests using individual averages); the average count is 146.19 without ranking, 236.10 with flat payment and 168.04 with tournament (No-ranking vs. Flat Payment: $Z = -3.083$, $p = 0.002$; Flat Payment vs. Tournament: $Z = 2.529$, $p = 0.011$, Wilcoxon rank-sum tests using individual averages). In the math task, the average taxonomy score is 7.5 with no-ranking, 9.19 with flat payment and 7.60 with tournament (No-ranking vs. Flat Payment: $Z = -0.752$, $p = 0.157$; Flat Payment vs. Tournament: $Z = 0.496$, $p = 0.620$, Wilcoxon rank-sum tests using individual averages); the average count is 64.04 without ranking, 66.00 with flat payment, and 91.90 with tournament (No-ranking vs. Flat Payment: $Z = 0.813$, $p = 0.416$; Flat Payment vs. Tournament: $Z = -0.418$, $p = 0.675$, Wilcoxon rank-sum tests using individual averages. with the median test and the Kolmogorov-Smirnov tests ($p = 0.903$ and $p = 0.586$, respectively).

incentives ($Z = 0.060$ and $p = 0.952$, respectively). As per Dillon (1982), most artistic endeavors generally represent open problems so that perhaps a true artist cannot be incentivized; artistic talent may simply be lacking. But "thinking harder" with open tasks does not help and could conceivably hurt (e.g., the so-called creative blockage, see for instance Corgnet et al., 2016); furthermore, increasing output in open-ended creativity may be more limited by current creative talent and skill levels than doing so with closed creativity. As with closed tasks, peer ranking promotes creativity: the average creativity score drops from 5.079 to 4.068 with only pure ranking ($Z = -3.880$ and $p = 0.001$).

While peer ranking seems to be effective with open tasks, we observe no additional effectiveness for tournament incentives with such tasks. Our explanation is that financial incentives are likely to work better if the task objectives are defined more precisely and are consequently perceived as clearer (as happens with closed creativity) because the evaluation process is easier to forecast for the subject who will experience it. Being ranked also appears to work as a form of (non-monetary) incentive that induces one to exert effort so as to avoid the potential negative feelings deriving from a poor ranking. This mechanism is not affected by the specificities of the task; please see Section 5.2 for an intuition and some discussion at the end.

This result that monetary tournament incentives are ineffective for inducing creativity in open tasks is in line with contract-theory insights regarding the effect of imprecision in the definition of contracts on reducing agents' effort (e.g. Laffont and Martimort, 1997; Faure-Grimaud et al., 2000). In a related vein, one of the predictions derived from expectancy theory (Vroom, 1964) is that noisy performance measurement challenges the perceived relationship between effort and performance, and thus reduces effort levels.

Some further support is provided by the fact that the evaluations of the external judges show a slightly higher degree of consistency and stronger correlation in the closed condition than in the open condition (Cronbach's alpha = .646 and .617 in the respective conditions; Spearman correlation test with coefficient = .481, $p = 0.000$ in the closed condition and coefficient = .448, $p = 0.000$ in the open condition). Delfgaauw et al. (2015) conduct a field experiment in a retail chain to test the prediction of tournament theory and find that noise dilutes incentives to perform because it reduces the marginal effect of effort on the probability of winning.

### 4.5. Regression analysis

This section examines the role of tournament incentives, taking demographic controls and personal characteristics into account. Among these, we devote particular attention to subjects' attitudes towards risk and ambiguity. We characterize investment choices in the ambiguous lottery in terms of risk aversion and ambiguity aversion. Furthermore, we consider the interaction between the presence of monetary incentives and risk/ambiguity aversion.[23]

Peer ranking increases creativity scores significantly with both closed and open tasks. When considering closed creativity, tournament incentives succeed in fostering creativity. Table 2 shows that tournament incentives matter *per se* for closed creativity, but Table 3 shows this is not true for open creativity. While neither risk-aversion nor any interaction between risk, ambiguity and incentives is significant, ambiguity aversion seems to consistently have a slightly negative influence on creativity.

Column 2 shows that participants putting more effort in the task (using more words or operations) are the people who receive a significantly higher creativity score.[24] Column 3 indicates that results do not change if we also control for the taxonomy score; the same holds when distinguishing between the math

[23] 137 people showed no ambiguity-risk gap; of the rest, 102 people invested more with risk than with ambiguity, while 86 people invested less with risk than with ambiguity[23]. This is significantly different from random behavior ($Z = -2.139$, $p = 0.033$, binomial test). Overall, the average investment with risk was 55.82 and the average investment with ambiguity was 52.76 ($t = 1.843$, $p = 0.066$, one-sample $t$ test).

[24] Since the creativity score is correlated with effort and taxonomy score, it seems that, on some level, judges' creativity evaluations reflect these objective measurements. But of course the correlation is not perfect; we suspect that there is some residual that matters in a creativity evaluation and cannot be readily captured by objective measures. Perhaps it is this residual that is enhanced by monetary incentives.

and the verbal tasks (column 4), when controlling for the interaction between incentives and count (Column 5) and when considering demographic and personal features (Column 6).

[Table 2 about here.]

The regressions in Table 3 examine the role of tournament incentives, ranking, count, risk and ambiguity aversion, as well as interactions between these attitudes and tournament incentives in case of open creativity.

[Table 3 about here.]

As with closed tasks, peer ranking increases the creativity score significantly. On the contrary, tournament incentives are ineffective in shaping open creativity. Again, participants exerting more effort in the task (using a higher number of words) generally receive a significantly higher creativity score (column 2), and this effect holds also when controlling for the selected task (column 3) or for the interaction between count and incentives, as shown in column 4. Overall, there is no significant effect of risk or ambiguity attitudes or their interactions with incentives. However, while not statistically-significant, note that the coefficients of ambiguity attitude are all negative,[25] suggesting again some detrimental effect on creativity.

Consistent with the conceptual versus experimental classification (conceptual creative people have definite goals and methods, whereas experimentally creative people do not have clearly established methods or definite goals and use trial and error),[26] our data suggest that open creativity tasks correspond to experimental creativity: Column 5 shows that the more one has an experimental creative style, the higher the score for open creativity. We considered the fit of our measures of creativity with previous measures of

[25] The Z-statistics are over 1.17 in all cases,

[26] Based on Nielsen et al. (2008), which introduces operational definitions of Galenson (2004)'s creative methods.

creative style and personality, finding that open creativity overlaps some with Galenson (2004)'s definition of experimental creativity.[27]

We now focus on the role of demographic features and personal attitudes like creative style and sensation-seeking mind-set. Regarding closed creativity tasks, neither creative style nor preferences for sensation-seeking tasks plays a role. A marginally-significant gender effect emerges: males reach higher creative scores. Furthermore, people with larger past involvement in creative endeavors seem to be less creative. All in all, in closed tasks creativity appears to respond to financial incentives, but little else.

Turning to open creativity tasks, the more a subject's creative style is experimental rather than conceptual, the higher the creativity score in open tasks. Second, people with more elder siblings are more creative. This finding is consistent with the psychological research on the role of the characteristics of siblings (Baer et al., 2005) and birth-order (Eisenman, 1964), which emphasizes that later-born are generally less prone to conservatism and conformism.

**4.6. Quantile regression**

Do highly-creative subjects react more to financial incentives than less-creative ones? This sub-section investigates whether the effect of tournament incentives on subjects' creative output differs at different points of the distribution of subjects' creative output. We use quantile regressions to estimate the conditional distribution of the creative output of subjects at each decile [0,1] for both closed (Table 4) and open tasks (Table 5).

---

[27] An additional result pertains to the within-subject difference between investment in the ambiguity lottery and in the risky lottery. In the ranking and tournament treatments, participants with a non-negative difference have a significantly higher open-creativity score than those with a negative difference (5.30 versus 4.82, $Z = -1.980$, $p = 0.047$, Wilcoxon rank-sum test using individual averages). However, there is no effect of ambiguity when we add the data of the No-ranking treatment. More research is needed on this topic.

[Table 4 about here.]

Table 4 compares our baseline Tobit regression (already presented in Table 2 and now reported in column 1) with decile regressions (DR) from 1st to 9th. It shows that, for what concerns closed tasks, tournament incentives generally play a significant role for subjects from the 3rd decile of the distribution of creative score onward. In particular, tournaments incentives are not significant for subjects in the first and second decile. Some of the coefficients appear to differ across the deciles, and we use the Wald test on the hypothesis that the coefficients of tournament incentives are the same for the nine deciles: the null hypothesis of the coefficient equality cannot be rejected (Wald test: $F = 1.04$, $p = 0.413$). In general, the introduction of tournament incentives significantly modifies the distribution of creative scores in the closed task (Kolmogorov-Smirnov two-sample test: $D = 0.372$, $p < 0.001$).

[Table 5 about here.]

Table 5 reports decile regressions for the open task, comparing them with the basic Tobit regression in Table 3 (here reported in column 1). Tournament incentives play no significant role, with the exception of subjects in the first decile of the distribution of creative scores. The null hypothesis of coefficient equality cannot be rejected (Wald test: $F = 1.53$, $p = 0.152$).

Overall, the quantile analysis shows that subjects with the worst creative performances are mildly sensitive to tournament incentives in case of open tasks, and that subjects with the worst creative performance do not react to incentives in case of closed tasks. However, the difference with the rest of the population is not significant; the introduction of tournament incentives does not significantly modify the distribution of creative scores with open tasks (Kolmogorov-Smirnov two-sample test: $D = 0.153$, $p < 0.226$).

## 5. Structural model

How do financial incentives and ranking affect creative output? We would like to be able to interpret our estimates of the experimental data in the context of a clear theoretical model presenting different incentive schemes, so we build a simple structural model to capture directly the relationship between unobserved features and choice of exerting creative effort. We are interested in (i) estimating the magnitude of the return of tournament incentives to creative output, and (ii) predict behavioral responses to a richer set of incentive mechanisms (piece rate and target bonus) than those in the experiment.

### 5.1. Specifying the model

In this sub-section, we build a micro-economic model to consider the "supply" of individual creative output $q_i$. Specifically, in our model subjects balance the cost and benefit of engaging in creative tasks, and do so on the basis of two unobservables: creative skill ($s_i$) and attitude towards the specific task ($\varphi_i$). $\varphi_i$ is an idiosyncratic variable capturing how the individual's creative output depends upon the individual "reaction" to the openness of the task, where a task is defined as "open" when the goal is non-specific and loosely-delineated. Thus, individuals are heterogeneous with respect to the manner in which the openness of the task reduces the marginal benefit from producing creative output. We can refer to $e_i$ as creative effort and define creative output $q_i$ as

$$q_i = q(s_i, e_i, \varphi_i) = s_i + \varphi_i e_i \qquad (1)$$

**Assumption 1. Utility**

Every subject *i* has preferences of an identical form and maximizes the following expected utility:

$$E(Q) = W - V(e_i, s_i)$$

Thus, a given subject *i* receives utility from the payment $W$, that may or not depend on her creative output, according to the type of incentives provided. Additionally, she faces a disutility of effort $V(e_i, s_i)$ that is U-shaped: since intrinsically-creative people enjoy taking part in creative tasks, the disutility of effort decreases as long as subjects have to exert "low" levels of creativity (with respect of their skill). However, when they have to push beyond their "natural" creative level, disutility increases in effort.[28] We shall assume $V(e_i, s_i) = e_i^2/2 - s_i e_i$. The subject chooses the optimal level of effort $e_i^*$ by maximizing $E(Q)$. The key mechanism in the model is the following: subjects choose their effort and effort translates into creative output depending on subject's attitude towards the type of task. Creative output is also affected by skill and might shape payment (depending on the specific incentive scheme in place).

Subjects are randomly assigned to the task, thus there is no self-selection in the task (and no endogeneity issues). Each subject is endowed with some combination of skill and attitude towards the task, so that the combination of the two, together with the decision of how much effort to exert, will determine the subject's creative output and her earnings.

**Assumption 2. Incentive schemes**

- **No ranking.** Subjects all receive a fixed sum $Y$ in any case.
- **Tournament.** For the sake of simplicity, we consider the simplest case of a two-subject tournament in which the winner gets a fixed prize $W_H = Y + x$ and the loser gets a fixed prize $W_L = Y - x$, with $x > 0$ that can be interpreted as the risk or reward to be added to the safe payment $Y$. The winner of the tournament is determined by the largest creative output $q$. In contrast to the flat-payment condition, now the payment is no longer constant, but its expected value is exactly equal to $Y$ to keep the average payment constant across incentive schemes. The probability $P$ of winning depends on both contestants' effort and on the distributions of $s$ and $\varphi$, denoted with $g$ and $h$ respectively. The more the openness of the task "disturbs" the subject, the lower the subject's level of $\varphi_i$, the lower the effectiveness of effort in determining the subject's creative output, and the higher the subject's probability to win the contest.

[28] Contemporary work in psychology assumes that most individuals are capable of producing at least moderately creative work (Amabile, 1996).

- **Piece Rate.** The piece-rate pay scheme compensates subjects for each unit of creative output $q_i$ at a fixed rate: if the experimenter observes a creative output $q_i > 0$, the subject receives a payment $bq_i$. To keep the average payment constant across incentive schemes and equal to the flat pay $Y$, rate $b$ is set as the ratio between the flat pay $Y$ and the expected creative output of the average subject $\tilde{q}$: $b = Y/\tilde{q}$ (as in Bandiera et al., 2007).
- **Targeted Bonus.** The bonus pay scheme pays a sum equal to $B$ for each subject whose creative output exceeds an exogenously set threshold $\overline{q_z}$ (as in Healy, 1985). The threshold $\overline{q_z}$ is set as the creative output above an exogenously fixed percentile $z$. To keep the average payment constant across incentives schemes and equal to the flat pay $Y$, rate $B$ is set as the product between the flat pay $Y$ and the number of subjects whose creative output is above the $z^{th}$ percentile: $B = 100/(100 - z)\,Y$.

**Assumption 3. Peer ranking**

Since creativity must be "recognized in its utility by peers" (Mumford, 2003), the model accounts not only for monetary incentives, but also for the effect of peers' evaluation in the form of ranking: subjects' creative output might be evaluated, and in particular ranked, by other subjects engaged in the same creative task. As shown by Kreps (1997), peer evaluation represents an extrinsic incentive that is able to foster higher effort among employees. We assume an analogous effect in case of creative effort. Thus, peer ranking translates into a reduction of the disutility of effort (as in Kandel and Lazear, 1992), so that $V(e_i, s_i)$ is modified as follows:

$$V(e_i, s_i) = \frac{e_i^2}{2} - s_i e_i - re_i \quad (2)$$

where $r$ represents a common "environmental" variable capturing the effect of being ranked.

## 5.2. Solving the model

**No ranking.** In the case of no ranking, the optimal effort level is $e_i^* = s_i$.

**Tournament.** Consider contestant 1's problem and denote the opponent by 2. Subject 1's expected utility is:

$$P(e_1, e_2, \varphi_1, \varphi_{2,})[W_H - V(e_1, s_1)] +$$

$$+\left(1 - P(e_1, e_2, \varphi_1, \varphi_{2,})\right)[W_L - V(e_1, s_1)] \quad (3)$$

or

$$P(e_1, e_2, \varphi_1, \varphi_{2,})(W_H - W_L) - W_L + V(e_1, s_1)$$

One will supply effort until one's marginal disutility from exerting creative effort (attenuated by creative skill) is compensated by the increase in the chance of winning the prize. Subjects choose creativity effort in accordance with:

$$2x\frac{\partial P(e_1,e_2,\varphi_1,\varphi_{2,})}{\partial e_1} - \frac{\partial V(e_1,s_1)}{\partial e_1} = 0 \qquad (4)$$

where $2x = (W_H - W_L)$.

In order to compute $\frac{\partial P(e_1,e_2,\varphi_1,\varphi_{2,})}{\partial e_1}$, we recall that subject 1 wins the contest if $q_1 > q_2$, i.e. if $\varphi_1 e_1 + s_1 > \varphi_2 e_2 + s_2$. For a given $s_2$, the probability that $s_1 > \varphi_2 e_2 - \varphi_1 e_1 + s_2$ is equal to $1 - H[G(\varphi_2 e_2 - \varphi_1 e_1 + s_2)]$. The total probability of winning is obtained by integrating over all the possible values of $s_2$ and $\varphi_2$, weighted by the density of $s_2$, namely $g(s_2)$, and by the density of and $\varphi_2$, namely $h(\varphi_2)$. Hence, in the symmetric solution (when $e_1 = e_2, \varphi_1 = \varphi_2$ and $P = 1/2$), subject i's increased chance of winning by raising creative effort is

$$\frac{\partial P(e_1,e_2,\varphi_1,\varphi_{2,})}{\partial e_1} = \int\int g(s_2)g(s_2)ds_2 h(\varphi_2\ )h(\varphi_2\ )d\varphi_{2,} = \tilde{\varphi}\tilde{s} \quad (5)$$

where $\tilde{s} = E[g(s)]$ and $\tilde{\varphi} = E[h(\varphi)]$.

We obtain a level of optimal creative effort in tournaments that equals

$$e_i^{**} = 2\tilde{s}\tilde{\varphi}x + s_i$$

that can be generalized to

$$e_i^{**} = n\tilde{s}\tilde{\varphi}x + s_i \quad (6)$$

where *n* is a parameter depending on the number of subjects that compose each group.

The expression above illustrates that optimal effort increases in: (a) own creativity skill $s_i$ (as in case of flat payment), (b) the average creativity skill of subjects taking part into the tournament $\tilde{s}$, (c) the distance between the prize derived from winning the contest and the prize derived from losing it, which is $x$. On the other hand, optimal effort decreases in: (d) the average effect on creativity output of the openness of the task $\tilde{\varphi}$.

Two main implications derive from (6):

I) The monetary incentives provided by the tournament are effective in stimulating creativity effort: subjects increase their creativity effort with respect to flat payments ($e_i^{**} > e_i^{*}$) in the aim of raising the chances of winning the contest. More specifically, creativity effort rises in the monetary gap between the winner's and the loser's payment.

II) When $\tilde{\varphi} \leq 0$ (i.e., when the average marginal benefit on creative output of exerting effort is very small), $e_i^{**} \leq e_i^{*}$. Thus, the characteristics of the task determine the effectiveness of incentives in stimulating creative effort: when a task is loosely-defined, subjects perceive lower probability to win the task and become less sensitive to monetary incentives. In the extreme case of subjects experiencing the task as "fully open" and being unable to capture its goals (that is not the case we consider in the experiment), note tournament incentives will induce a lower amount of effort than under flat payment when $\tilde{\varphi} < 0$.

**Piece Rate.** With a piece-rate payment, the subject maximizes the expected utility *E(Q):*

$$E(Q) = bq_i - V(e_i, s_i)$$

The optimal effort level is:

$$e_i^{***} = b\varphi_i + s_i \qquad (7)$$

This expression shows that optimal effort increases in: (a) own creativity skill $s_i$ (as in previous cases) and (b) the rate *b*. The optimal effort decreases in: (c) the effect on creative output $q_i$ of the openness of the task.

As in the case of a tournament, the monetary incentives provided by the piece rate are effective in stimulating creativity effort: subjects increase their creativity effort with respect to flat payments ($e_i^{***} > e_i^{*}$) in the aim of increasing their creative output and thus their payment. However, this does not hold when $\varphi_i \leq 0$: in the presence of open tasks with very loosely-defined goals, subjects experience difficulties in translating effort in higher creativity output and become less sensitive to monetary incentives.

**Targeted Bonus.** With a targeted bonus payment, the subject receives a bonus only if her creativity output $q_i$ exceeds an exogenously fixed threshold ("target") of creative output $\overline{q_z}$ . The bonus is equal to $B = \frac{100}{100-z}\, Y$ and is paid only if $q_i \geq \overline{q_z}$ ; if the target threshold is not met, the subject gets zero (Healy, 1985).

Subjects therefore maximize the following expected utility *E(Q)* as above:

$$E(Q) = B - V(e_i, s_i) \text{ if } q_i \geq \overline{q_z}$$

$$= 0 \; - \; V(e_i, s_i) \text{ if } q_i < \overline{q_z}$$

Under the usual assumptions on $V(e_i, s_i)$, the first-order conditions lead to

$$e_i^{****} = B + \; s_i \;\; \text{if } q_i \geq \overline{q_z}$$

$$= s_i \qquad \text{if } q_i < \overline{q_z} \qquad (8)$$

This expression shows that optimal effort increases in: (a) own creativity skill $s_i$ (as in the previous cases) and (b) the rate *B*. Furthermore, it emerges that (c) the higher the target threshold $\overline{q_z}$, the lower is the chance to obtain the bonus. Since the rate bonus $B = 100/(100 - z)\, Y$ is set in order to reflect the number of subjects whose creative output is above the $z^{th}$ percentile, the subject's expected payment depends on her own probability to reach the target. What is important to note is that, differently from tournament and piece-rate incentive schemes, the optimal effort depends on $\varphi_i$ only *via* the probability to reach the target. Therefore, the monetary incentives for the target bonus are only indirectly affected by how the openness of the task affected one's productivity.

**Peer ranking.** If subjects are evaluated by peers in addition to receiving monetary incentives in the forms described above, the first-order conditions reported above are modified since we have:

$$V(e_i, s_i) = \frac{e_i^2}{2} - s_i e_i - r e_i \qquad (9)$$

so the model predicts higher optimal effort. For instance, with pure ranking, we get $e_i^{*\prime} = s_i + r$. The same applies to other incentive schemes.

It is worth noting that the incentivizing effect of peer ranking does not depend on the openness of the task. Whereas peer ranking reduces the cost of creative effort (see equation 2 above), financial incentives pay subjects according to subjects' creative output, which depends on the openness of the task: the less the goal is well-defined, the more difficult subjects are able to predict how their performance will be evaluated.

One may very well wonder why this is the case. What we observed in the experiments is that peer ranking is very powerful in stimulating people to put at least a minimum effort level in what they are doing, most likely because they feel some form of peer pressure. The model helps to get to a clearer interpretation on why this works no matter the task, differently from monetary incentives: equation (9) shows that peer ranking depends only on individuals' creative effort, whereas financial incentives pay on the basis of creative output; this is jointly determined by effort, skill and $\varphi_i$. Thus, obtaining improvement in creative output is a more complex issue. It is the creative output that depends on the openness of the task: the less the goal is well-defined, the more difficult it is for one to predict how one's performance will be evaluated.

### 5.3. The distribution of unobservables

Section 5.2 above reports the solution of the model: if we know (or posit) values of skill $s_i$ and $\varphi_i$ and the incentive schemes the subject receives, we know one's creative output and how much one will earn from the experiment. This relies on the assumption of people maximizing utility by choosing the optimal level of effort (contingent on the incentive scheme).

Before moving to the empirical analysis, we need to make distributional assumptions about the unobservables in the model ($s_i$ and $\varphi_i$). The simplest way to proceed is to assume that $s_i$ has a normal distribution; we do not need to make an assumption on $\varphi_i$ because it can be derived from equation (1), since we observe subjects' creative output $q_i$ and subjects' creative effort $e_i$ in the experiment.

The model makes predictions both about how incentives and $\varphi_i$ determine effort $e_i$, and how $s_i$, $e_i$ and $\varphi_i$ determine creative output $q_i$. We can encompass the four above equations in a single nested model that also accounts for the role of peer ranking:

$$e_i = \beta_1 s_i + \beta_2 n\tilde{s}\tilde{\varphi}x + \beta_3 b\varphi_i + \beta_4 B + \beta_5 r \qquad (10)$$

Each parameter $\beta_j$ (with *j=1, …, 5* indicating the generic incentive scheme) weights the determinant(s) of the optimal effort level in each incentives scheme *j*, as described above. Thus, the object of interest is a vector of parameters to which we can refer as Θ:

$$\Theta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \sigma^2 \end{pmatrix}$$

**5.4. Structural estimation**

We proceed to the structural estimation of the experimental data using maximum-likelihood estimation. The estimation is performed using data from the Tournament, the Pure-ranking and the No-ranking treatments and considers both open and closed tasks; the experimental variation in the presence of incentives and in the openness of the task permits separate identification of the structural parameters. These parameter estimates then allow the model to be used for an out-of-sample prediction exercise with the Piece Rate and Targeted Bonus incentive schemes.

In the empirical implementation we specify a parametric distribution of the unobserved heterogeneity in skill $s$. The distribution of $s$ varies with individual characteristics $X_i$, and the mean of the distribution is allowed to vary with demographic and personal attitudes through a linear index restriction $\mu_s = X'\beta_s$. The proxy that we use for effort is the number of words or math operations (called "count" in Section 4.3.) each subject provided in the response.

The maximum-likelihood estimates of the effect of tournament incentives can then be obtained from two independent Tobit regressions: the first uses data from the Pure-ranking treatment, regressing $e_i = s_i$ on the set of individual controls $X_i$ (including a constant); the second uses data from the Tournament treatment and regresses $e_i = n\tilde{s}\tilde{\varphi}x + s_i$ on the same set.

Table 6 presents the parameter estimates from our model.

[Table 6 about here.]

The first column summarizes the effect of tournament incentives with respect to flat payment, whereas the second column reports the effect of peer ranking with respect to no-peer ranking when tournament incentives are not in effect. As shown, skill, tournament incentives, and ranking are significant at the 1% level.

The Likelihood Ratio test compares the goodness of fit of the two optimal effort levels ($e_i^{FLAT} = s_i$ and $e_i^{TOURN} = n\tilde{s}\tilde{\varphi}x + s_i$) using data from the Pure-ranking and Tournament treatments, and shows there is a significant effect of tournament incentives on the creative score (LR test: $\chi^2 = 61.43$, $p = 0.000$).

Focusing on the role of peer ranking, the Likelihood Ratio test compares the goodness of fit of the two optimal effort levels ($e_i^{NO-RANK} = s_i$ and $e_i^{PURE-RANK} = s_i + r$) using data from the No-ranking and the Pure-ranking treatments, showing that there is a significant effect of peer ranking on the creative score (LR test: $\chi^2$ = 16.90, $p$ = 0.000).

**5.5. Counterfactual incentive schemes**

We now use our structural model to explore the effectiveness of alternative incentive schemes, as was done in Huck et al. (2015) who utilized the model to predict giving behavior for a series of counterfactual schemes in the context of fundraising. We conduct counterfactual experiments using the estimated parameters from the structural model: all moments are computed using the empirical distributions observed in the experiment. We simulate data according to our model, allowing for 1000 observations for each treatment. The reduced-form evidence presented above suggests that, on average, financial incentives are effective, but not with open tasks. The structural evidence allows for subjects' heterogeneity in the effect of the openness of the task on the marginal benefit of effort: only subjects who are not "disturbed" by the openness of the task enhanced their creative output under tournament incentives. The counterfactual analysis explores the effectiveness of Piece Rate and Target Bonus schemes in order to test whether there are other incentive schemes that out-perform the Tournament scheme.

We assume that one "piece" of creativity is represented by one score point (in a scale from 1 to 10) that a judge assigns to the subject's creative output. The same holds for the Target Bonus scheme, which is implemented as a special case of Piece Rate where subjects receive a bonus when they produce a creativity output reaching a given threshold ("target").

Table 7 presents the maximum-likelihood estimates for the two counterfactual incentives schemes taken into consideration. Columns 1 and 2 present the empirical and the predicted outcomes for the Tournament treatment; columns 3 and 4 show the predicted outcomes for the Piece Rate and the Target Bonus treatments. The Likelihood Ratio test compares the goodness of fit of each incentive scheme predicted optimal effort ($e_i^{PIECE} = b\varphi_i + s_i + r$ and $e_i^{BONUS} = B + s_i + r$) with $e_i^{FLAT} = s_i + r$ using simulated data, and shows that incentives are always effective in stimulating higher creative output.

[Table 7 about here.]

The Table reports the average creative output of each scheme at the bottom of the corresponding column. What emerges is that the predicted average creative output is significantly different across schemes, with the Target Bonus (with a predicted average creative output equal to 6.737) being the more effective scheme in fostering creativity, and Piece Rate (with a predicted average creative output equal to 6.404) being more effective than Tournament (with a predicted average creative output equal to 6.077) (Wilcoxon rank-sum test; Tournament vs. Piece Rate: $Z = 13.603$, $p = 0.000$; Tournament vs. Bonus: $Z = 24.311$, $p = 0.000$; Piece Rate vs. Bonus: $Z = 13.604$, $p = 0.000$). The better performance of the Target Bonus is consistent with the fact that subjects' attitude toward open tasks (captured by $\varphi_i$) enter the optimality condition only indirectly via the probability of reaching the target threshold.

## 6. Discussion

Our results seem clear. Creativity is markedly higher when there is a closed task and extrinsic incentives are provided. Regarding our open tasks, financial incentives are effective for subjects for which the openness of the task is not too harmful, as shown by our structural estimation. Knowing that peers will rank one's creative work (anonymously!) is also a powerful motivation.

Contrary to the predictions of some relevant literature in psychology, we see no evidence that providing financial incentives has a crowding-out effect on creativity. With tasks with reasonable clarity about goals and constraints, monetary incentives do work, no matter the specific form (tournament, piece rate, or bonus) they assume. This is good news in that, if true, providing financial rewards for creative performance will be at most costly only to the extent of the cost for the rewards. One might argue that there is little or no intrinsic motivation in the first place, but this belies the mental effort most people put into the task when there was a flat payment and the work *per se* clearly did not benefit the researchers.[29,30]

According to Baer et al. (2003), the inconsistent relationship between rewards and creativity could result from the interaction between intrinsic and extrinsic motivations: creativity is enhanced by intrinsic motivations that are boosted by the presence of extrinsic rewards. A relevant dimension is that of cognitive style, either innovative or adaptive: highly-motivated people reach greater achievements, are more open to new experiences, and exhibit higher productivity in a variety of aspects of life (Heckman, 2007), intrinsic motivational qualities are likely to be stronger for those with an innovative style than for those with an adaptive style; the latter tend to perceive their jobs as being instrumental for obtaining extrinsic rewards.

This fits well with our results on creativity, as providing financial incentives has no beneficial effect in the more innovative open task, but does have an effect in the more adaptive closed task. Perhaps when employers wish to stimulate employees' creativity in organizations, monetary incentives should be offered according to the type of job, and employees should be trained to deal with loosely-defined problems in order to develop the capability to reason in "open" contexts where there is no clear path to follow. Another possible explanation is based on the "short-term" structure of our incentive mechanism: Ederer and Manso (2013) find that long-term (vs. short-term) reward is able to motivate what they call "exploration", which presents similarities with the attitude required in our open creativity task (whereas "exploitation" resembles a closed creativity task).

We use subjective measurement (participants' and external judges' evaluation) together with objective measurement. When judges evaluate creativity without indication of which criteria to follow, they

---

[29] In Appendix D, we present some examples of the creative responses made by the participants.

[30] Of course, it may also be possible to "crowd-in" intrinsic motivation. For example, Charness and Gneezy (2009) found strong effects from paying students to go to the gym multiple times and exercise. The main driver of this result was that people who had not previously been regular gym attendees continued to go to the gym after the payment period had ended.

appear to effectively share certain objective principles, but also focus on something that is idiosyncratic and therefore difficult to capture. Interestingly, our findings suggest it is the latter component that is more responsive to financial incentives.

**Implications for innovation**

A natural consideration for economists is the implication of our findings for innovation.[31] Patents have been used to prize innovators through the creation of (temporary) market power, yet there is a debate focused on the tradeoff between the gains generated by innovation - with the consequent need to provide incentives for stimulating R&D investment - and the costs of patent monopoly power (Gilbert and Shapiro, 1990). Following Kremer and Glennerster (2004)'s taxonomy of government interventions, the patent system is the more familiar "pull" program; these pay off only if an innovation is developed, whereas "push" programs subsidize the search for a socially-desirable innovation - such as a vaccine - whether or not the search is successful. Kremer and Glennerster advocate a monetary prize large enough to get the attention of the pharmaceutical companies and have them invest in the discovery of a new effective vaccine.

We find that providing incentives can indeed have favorable effects for innovation when the goal is already delineated, as is often the case with incremental innovations. On the other hand, drastic innovations are typically less defined *ex ante*. In a similar vein, Hellmann and Thiele (2011) provide a theoretical model that shows that incentive contracts are feasible for those tasks that are well understood and measurable *ex ante*. To the extent that drastic innovations map onto what we have termed an open task, direct incentives seem unnecessary. Instead, firms and governments may wish to simply support or subsidize basic research, which is executed without any specific applications or products in mind.[32] Since innovation involves the exploration of untested approaches that are likely to fail, incentive schemes that

[31] Our results also offer implications for incentivizing artistic and musical creativity, topics of interest to economists.

[32] Basic research lays the foundation for advancements in knowledge that lead to applied gains later on, occasionally as a result of unexpected discoveries.

punish failures with low rewards and termination may have adverse effects on innovation: as shown by Manso (2011), the optimal incentive scheme that motivates innovation should exhibit substantial tolerance for early failure and reward for long-term success.

## 7. Conclusion

Creativity is a main driver of the world's economy. Without creativity in areas such as science, technology, and the arts, our lives would be considerably poorer economically and aesthetically. From an economist's standpoint, one critical question is whether it is possible to incentivize creativity. We investigate whether incentives for performance can lead to higher levels of creativity at the individual level. To the best of our knowledge, we are the first to explore theoretically and experimentally how the effect of financial incentives on creativity can vary across types of creative tasks and across incentive schemes.

We consider individual creativity in a laboratory environment. First, our results show that peer ranking is a powerful form of non-financial incentive that can stimulate creative effort no matter the type of the task. Second, when a task has specific *ex-ante* goals, tournament financial incentives successfully induce a higher degree of creativity. However, we find no evidence that these induce creativity that is relatively unconstrained and non-goal-oriented. In this case, tournament incentives appear to be ineffective. A structural estimation of both experimental and simulated data allows us to account for individual heterogeneity in how the openness of the task alters the marginal benefit of exerting creative effort, and sheds further light on the effectiveness of alternative incentives schemes: piece rate and target bonus.

Our results are applicable to a wide range of economic environments, particularly when a clear need has been identified. Perhaps the best that can be done to achieve creativity in these realms is to create a research environment where funds are available as needed for talented researchers. This seems preferable to having competitions for research grants, as this latter approach seems much more conducive to incremental advances.

We have scratched the surface on the relationship between incentives and creativity, and there is much more work to be done. For example, how do people select into creative versus non-creative activities?

This cross-person variation is likely to be an important part of the creativity production function.[33] Nevertheless, we at least offer some novel and insightful results. We provide clean theoretical predictions and experimental evidence concerning the impact of financial rewards on two forms of creativity problems, which certainly points to the need for further research on this important issue.

## References


Aghion, Philippe and Jean Tirole (1994). "The Management of Innovation." *Quarterly Journal of Economics*, 109, 1185-1209.

Alekseev, Aleksandr, Gary Charness, and Ury Gneezy (2017). "Experimental Methods: When and Why Contextual Instructions Are Important." *Journal of Economic Behavior & Organization*, 134, 48-59.

Amabile, Teresa M. (1989). *Growing Up Creative: Nurturing a Lifetime of Creativity*. Norwalk, CT, US: Crown House Publishing Limited.

Amabile, Teresa M. (1996). *Creativity In Context: Update To The Social Psychology Of Creativity*. Boulder: Westview Press.

Ariely, Dan, Uri Gneezy, George Loewenstein and Nina Mazar (2009). "Large Stakes and Big Mistakes." *Review of Economic Studies*, 76, 451-469.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. (2011). "Incentives and Creativity: Evidence From the Academic Life Sciences." *RAND Journal of Economics*, 42(3), 527-554.

Baer, Markus, Greg Oldham, and Anna Cumming (2003). "Rewarding Creativity: When Does It Really Matter?" *Leadership Quarterly*, 14(4), 569-586.


[33] We thank John List for this comment.

Baer, Markus, Greg R. Oldham, Andrea B. Hollingshead, and Gwendolyn Costa Jacobsohn (2005). "Revisiting the Birth Order-Creativity Connection: The Role of Sibling Constellation." *Creativity Research Journal*, 17(1), 67-77.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2007). "Incentives for Managers and Inequality Among Workers: Evidence from a Firm-level Experiment." *Quarterly Journal of Economics*, 122(2), 729-773.

Boudreau, Kevin J., Nicola Lacetera and Karim R. Lakhani (2011). "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis." *Management Science*, 57(5), 843-863.

Bowden, Edward and Mark Jung-Beeman (2003). "Aha! Insight Experience Correlates with Solution Activation in the Right Hemisphere." *Psychonomic Bulletin & Review*, 10, 730-737.

Bradler, Christiane, Susanne Neckermann, and Warnke, Arne J. (2013). "Incentivizing Creativity: A Large-Scale Experiment with Tournaments and Gifts." ZEW Discussion Papers 16-040, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Charness, Gary and Ury Gneezy (2009). "Incentives to Exercise." *Econometrica*, 77, 909-931.

Charness, Gary, David Masclet, and Marie C. Villeval (2014). "The Dark Side of Status." *Management Science*, 60, 38-55.

Chen, Clara X., Michael G. Williamson, and Flora H. Zhou (2012). "Reward System Design and Group Creativity: An Experimental Investigation." *The Accounting Review*, 87(6), 1885-1911.

Collins, Mary A. and Teresa M. Amabile (1999). "Motivation and Creativity". In Robert J. Sternberg (Ed.), *Handbook of Creativity* (pp. 297-312). Cambridge, England: Cambridge University Press.

Corgnet, Brice, Antonio M. Espín, and Roberto Hernán-González (2016). "Creativity and Cognitive Skills among Millennials: Thinking Too Much and Creating Too Little." *Frontiers in Psychology*, 7.

Deci, Edward and Richard Ryan (1985). *Intrinsic Motivation and Self-determination in Human Behavior*. New York: Plenum.

Delfgaauw, Josse, Robert Dur, Arjan Non, and Willem Verbeke (2015). "The Effects of Prize Spread and Noise in Elimination Tournaments: A Natural Field Experiment." *Journal of Labor Economics*, 33 (3), 521-569.

Diamond, Jared (1999). *Guns, Germs, and Steel: The Fates of Human Societies*. WW Norton & Company.

Dillon, Jessica (1982). "Problem Finding and Solving." *Journal of Creative Behavior*, 16, 97-111.

Dow, Gayle T. and Richard E. Mayer (2004). "Teaching Students to Solve Insight Problems: Evidence for Domain Specificity in Creativity Training." *Creativity Research Journal*, 16, 389-402.

Eckartz, Katarina, Oliver Kirchkamp, and Daniel Schunk (2012). "How Do Incentives Affect Creativity?" *CESifo Working Paper Series* 4049, CESifo Group Munich.

Ederer, Florian and Gustavo Manso (2013). "Is Pay-for-Performance Detrimental to Innovation?" *Management Science*, 59(7), 1496-1513.

Eisenberger, Robert, Stephen Armeli, and Jean Pretz (1998). "Can the Promise of Reward Increase Creativity?" *Journal of Personality and Social Psychology*, 74, 704–714.

Eisenberger, Robert and Linda Rhoades (2001). "Incremental Effects of Reward on Creativity." *Journal of Personality and Social Psychology*, 81, 728–741.

Eisenman, Russel (1964). "Birth Order and Artistic Creativity." *Journal of individual Psychology,* 183-185.

Ellsberg, Daniel (1961). "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics*, 75(4), 643-669.

Erat, Sanjiv and Ury Gneezy (2015). "Incentives for Creativity." *Experimental Economics*, 19(2), 269-280.

Erat, Sanjiv and Vish Krishnan (2012). "Managing Delegated Search over Design Spaces." *Management Science*, 58(3), 606-623.

Eysenck, Michael H. (2003). "Creativity, Personality, and the Convergent-Divergent Continuum." In Mark A. Runco (Ed.), *Critical Creative Processes* (pp. 95–114). Cresskill, NJ: Hampton.

Faure-Grimaud, Antoine, Jean J. Laffont and David Martimort (2000). "A theory of supervision with endogenous transaction costs." *Annals of Economics and Finance*,1, 231-263.

Feinstein, Jonathan (2006). *The Nature of Creative Development*. Stanford: Stanford University Press.

Feinstein, Jonathan (2009). "Creativity & The Economic System." In Kerr, B. (Ed.). *Encyclopedia of Giftedness, Creativity, and Talent*. New York: Sage Publications.

Fiore, Stephan and Jonathan W. Schooler (2001). “Convergent or Divergent Problem Space Search. The Effect of Problem Structure on Group versus Individual Problem Solving.” *Proceedings of the Human Factor Ergonomic Society 45th ANNUAL MEETING*, 483-487.

Galenson, David (2004). “A Portrait of the Artist as a Very Young or Very Old Innovator: Creativity at the Extremes of the Life Cycle.” *NBER*, Working Paper No. 10213.

Getzels, J. W. and Mihaliy Csikszentmihalyi (1967). “Scientific Creativity.” *Science Journal*, 3, 80-84.

Gilbert, Richard and Carl Shapiro (1990). “Optimal Patent Breadth and Length.” *The Rand Journal of Economics*, 21(1), 106-112.

Gino, Francesca and Dan Ariely (2012). “The Dark Side of Creativity: Original Thinkers Can Be More Dishonest.” *Journal of Personality and Social Psychology*, 102(3), 445-459.

Gneezy, Ury and Jan Potters (1997). “An Experiment on Risk Taking and Evaluation Periods.” *Quarterly Journal of Economics*, 112, 631-645.

Gneezy, Ury and Aldo Rustichini (2000). “Pay Enough or Don’t Pay at All.” *Quarterly Journal of Economics*, 115(3), 791-810.

Goldenberg, Jacob, David Mazursky, and Sorin Solomon (1999). “Creative Sparks.” *Science*, 285 1495-1496.

Gough, Harrison G. (1979). “A Creative Personality Scale for the Adjective Check List.” *Journal of Personality and Social Psychology*, 37, 1398-1405.

Guilford, Joy P. (1950). “Creativity.” *American Psychologist*, 5, 444-454.

Guilford, Joy P.  (1962). *A Source Book for Creative Thinking*. New York: Charles Scribner's Sons.

Healy, Paul M. (1985). “The Effect of Bonus Schemes on Accounting Decisions.” *Journal of Accounting and Economics*, 7(1-3), 85-107.

Heckman, James (2007). “The Economics, Technology and Neuroscience of Human Capability Formation.” *PNAS*, 104(33), 13250–13255.

Heckman, James and Yona Rubinstein (2001). “The Importance of Noncognitive Skills: Lessons from the GED Testing Program.” *American Economic Review*, 91(2), 145-149.

Hellmann, Thomas and Veikko Thiele (2011). “Incentives and Innovation: A Multitasking Approach.” *American Economic Journal: Microeconomics*, 3(1), 78-128.

Hennessey, Beth and Teresa M. Amabile (1998). “Reward, Intrinsic Motivation, and Creativity.” *American Psychologist*, 53, 674-675.

Hocevar, Dennis (1980). “Intelligence, Divergent Thinking, and Creativity.” *Intelligence*, 4, 25-40.

Huck, Steffen, Imran Rasul and Andrew Shephard (2015). “Comparing Charitable Fundraising Schemes: Evidence from a Natural Field Experiment and a Structural Model.” *American Economic Journal: Economic Policy*, 7(2), 326-69.

Hudson Liam (1966). *Contrary Imaginations: A Psychological Study of the English Schoolboy*. London: Methuen.

Kahneman, Daniel (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kandel, Eugene and Edward P. Lazear (1992). “Peer Pressure and Partnerships.” *Journal of Political Economy*, 100(4), 801-817.

Kirton, Michael (1976). “Adaptors and Innovators: A Description and Measure.” *Journal of Applied Psychology*, 61(5), 622-629.

Kremer, Michael (1993). “Population Growth and Technological Change: One Million BC to 1990.” *The Quarterly Journal of Economics*, 681-716.

Kremer, Michael and Rachel Glennerster (2004). *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases.* Princeton, NJ: Princeton University Press.

Kreps, David M. (1997). “Intrinsic Motivation and Extrinsic Incentives.” *American Economic Review*, 87(2), 359-364.

Kohn, Alfie (1993). *Punished by Rewards*. Boston: Houghton Mifflin.

Laffont, Jean J. and DavidMartimort, D. (1997). “The Firm as a Multicontract Organization.” *Journal of Economics & Management Strategy*, 6(2), 201-234.

Laske, Katharina and Marina Schroeder (2016). “Quantity, Quality, and Originality: The Effects of Incentives on Creativity.” *Cologne Graduate School Working Paper Series*, *07-01*.

Lerner, Josh and Julie Wulf (2007). “Incentives and Innovation: A Multitasking Approach.” *The Review of Economics and Statistics*, 89(4), 634-644.

Mayer, Richard, Gayle Dow and Sarah Mayer (2003). “Multimedia Learning in Interactive Self-Explaining Environment: What Works in the Design of Agent-Based Microworlds?” *Journal of Educational Psychology*, 95(4), 806-813.

Manso, Gustavo (2011). “Motivating Innovation.” *The Journal of Finance*, LXVI(5), 1823-1860.

McCullers, John C. (1978). “Issues in Learning and Motivation.” In Mark R. Lepper, David Greene (Eds.). *The Hidden Costs of Reward*. Erlbaum, Hillsdale, NJ, 5-18.

McGraw, Kenneth O. (1978). “The Detrimental Effects of Reward on Performance: A Literature Review and a Prediction Model.” In Mark R. Lepper, David Greene (Eds.). *The Hidden Costs of Reward. Erlbaum*. Hillsdale, NJ, 33-60.

Mumford, Michael D. (2003). “Where Have We Been, Where Are We Going? Taking Stock in Creativity Research.” *Creativity Research Journal*, 15, 107-120.

Nickerson, Raymond S. (1999). “Enhancing Creativity.” In Robert J. Sternberg (Ed.). *Handbook of Creativity* (pp. 392-430). Cambridge, England: Cambridge University Press.

Nielsen, Bayard D., Cynthia L. Pickett, and Dean K. Simonton (2008). “Conceptual versus Experimental Creativity: Which Works Best on Convergent and Divergent Thinking Tasks?” *Psychology of Aesthetics, Creativity, and the Arts*, 2(3), 131-138.

Ohly, Sandra, Sabine Sonnentag, and Franziska Pluntkent (2006). “Routinization, Work Characteristics and their Relationships with Creative and Proactive Behaviors.” *Journal of Organizational Behavior*, 27(3), 257-79.

Pfeffer, Jeffrey and Nancy Langton (1993). “The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty.” *Administrative Science Quarterly*, 382-407.

Runco, Mark A. (1994). *Problem Finding, Problem Solving, and Creativity*. Greenwood Publishing Group.

Runco, Mark A. (2006). *Creativity: Theories and Themes: Research, Development, and Practice*. San Diego: Academic Press.

Runco, Mark A. (2007). “Correcting the Research on Creativity.” *Creativity Research Journal*, 17, 321-327.

Schulze, Anja and Martin Hoegl (2008). “Organizational Knowledge Creation and the Generation of New Product Ideas: A Behavioral Approach.” *Research Policy*, 37(10), 1742-1750.

Schweitzer, Maurice, Lisa Ordóñez and Bambi Douma (2004). “Goal Setting as a Motivator of Unethical Behavior.” *Academy of Management Journal*, 47 (3), 422-432.

Simonton, Dean K. (2000). “Creativity: Cognitive, Personal, Developmental, and Social Aspects.” *American Psychologist*, 55, 151-158.

Stanovich, Keith and Richard West (2000). “Individual Difference in Reasoning: Implications for the Rationality Debate?” *Behavioural and Brain Sciences*, 23, 645-726.

Stein, Morris I. (1974). *Stimulating Creativity*. New York: Academic Press.

Sternberg, Robert J. (1985). *Beyond IQ: A Triarchic Theory of Intelligence*. Cambridge: Cambridge University Press.

Sternberg, Robert J. (1988). “Mental Self-Government: A Theory of Intellectual Styles and their Development.” *Human Development*, 31,197-224.

Sternberg, Robert J. (Ed.) (1999). *Handbook of Creativity*. New York: Cambridge University Press.

Sternberg, Robert J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.

Surowiecki, James (2004). *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, and Nations*. New York: Random House.

Torrance, Paul E. (1962). *Guiding Creative Talent.* Englewood Cliffs, NJ: Prentice Hall.

Torrance, Paul E. (1974). *The Torrance Tests of Creative Thinking. Norms-Technical Manual Research Edition. Verbal Tests, Forms A and B- Figural Tests, Forms A and B*. Princeton, NJ: Personnel Press.

Torrance, Paul E. (1989). “The Nature of Creativity as Manifest in its Testing.” In Robert J. Sternberg (Ed.), *The Nature of Creativity: Contemporary Perspectives*. New York: Cambridge University Press.

Toubia, Olivier (2006). “Idea Generation, Creativity, and Incentives.” *Marketing Science*, 25(5), 411-425.

Unsworth, Kerrie (2001). “Unpacking Creativity.” *Academy of Management Review*, 26(2), 289-297.

Vroom, Victor (1964). *Work and Motivation*. New York: Wiley.

Walls, Theodore A. and Todd Little (2005). “Relations Among Personal Agency, Motivation, and School Adjustment in Early Adolescence.” *Journal of Educational Psychology*, 97(1), 23-31.

Wakefield, John F. (1991). “The Outlook for Creativity Tests.” *The Journal of Creative Behavior*, 25(3), 184-193.

Weitzman, Martin (1998). “Recombinant Growth.” *Quarterly Journal of Economics*, 113(2), 331-360.

Winston, Andres S. and Joanne E. Baker (1985). “Behavior Analytic Studies of Creativity: A Critical Review.” *Behavior Analyst*, 8, 191-205.

Woodman, Richard W., John E. Sawyer and Ricky W. Griffin (1993). “Toward a Theory of Organizational Creativity.” *Academy of Management Review*, 18, 293-321.

Zuckerman, Marvin, Elizabeth A. Kolin, Leah Price and Ina Zoob (1964). “Development of a Sensation-Seeking Scale.” *Journal of Consulting Psychology*, 28(6), 477-482.

# Figures and tables

**Figure 1. Creativity scores in open versus closed tasks in the three treatments.**

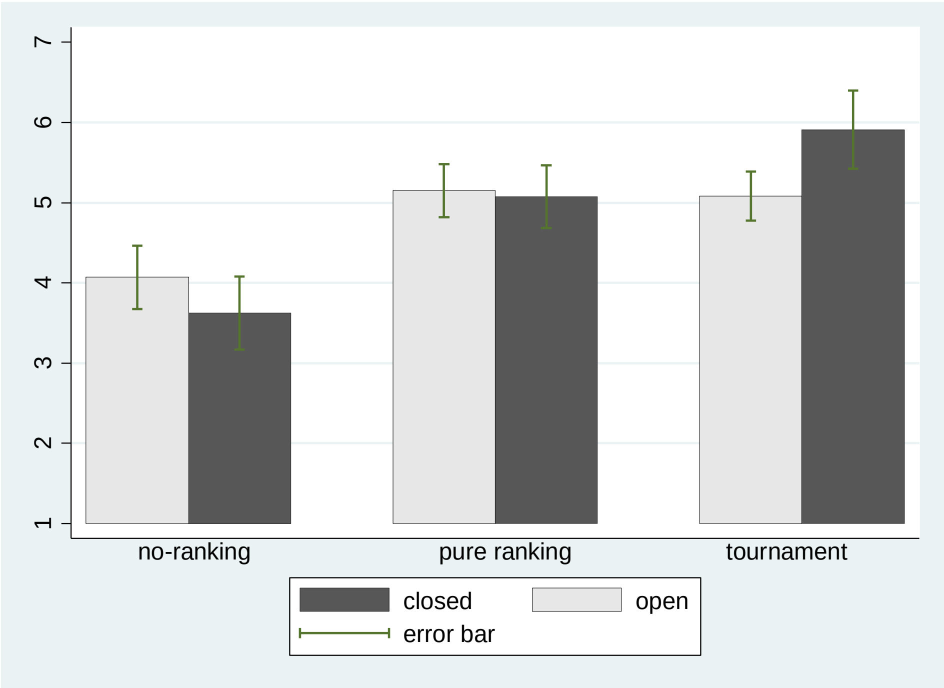


**Figure 2a: Distributions of creativity scores: effect of ranking**

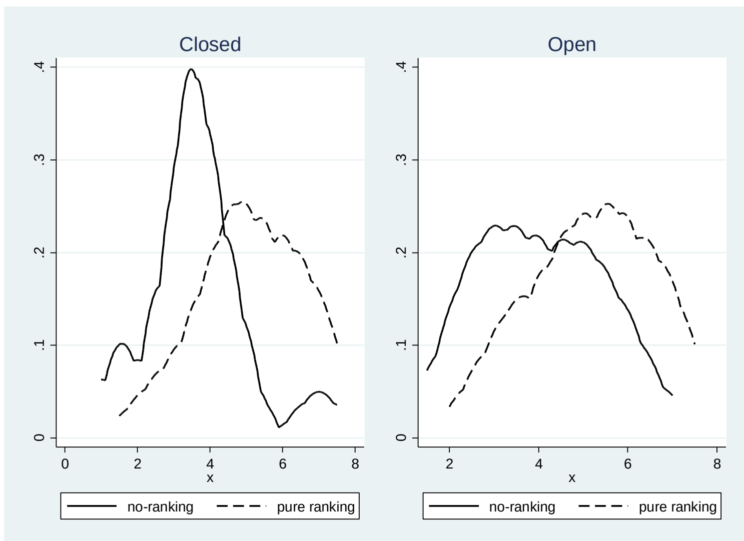



**Figure 2b: Distributions of creativity scores: effect of tournament incentives**

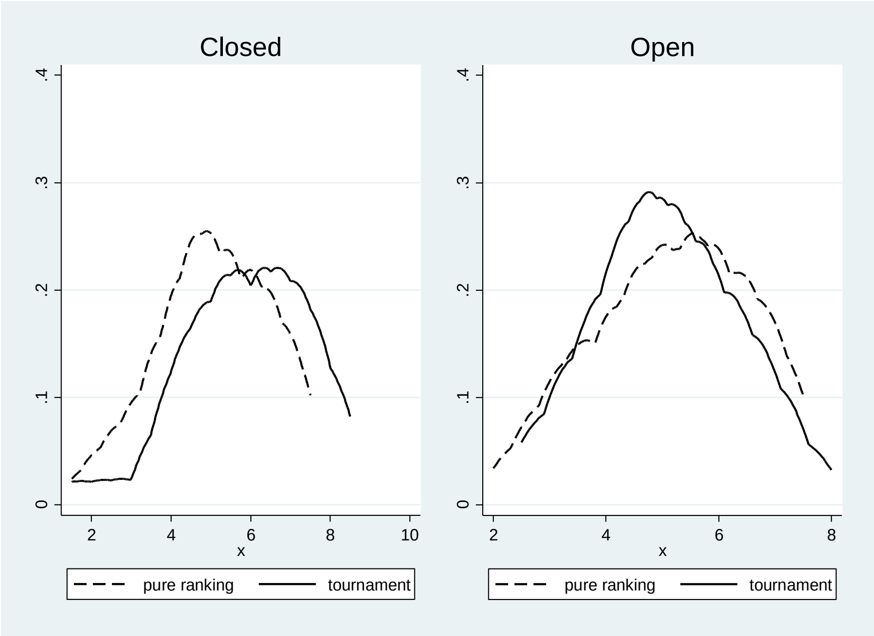



**Table 1. Creativity score by treatment: summary statistics**

| Treatments | Closed, no ranking | Open, no ranking | Closed , pure ranking | Open, pure ranking | Closed, tournament | Open, tournament |
|---|---|---|---|---|---|---|
| Average | 3.625 | 4.068 | 5.075 | 5.150 | 5.909 | 5.079 |
| Standard error | 0.224 | 0.197 | 0.193 | 0.165 | 0.240 | 0.152 |
| Min | 1 | 1.5 | 1.5 | 2 | 1.5 | 2.5 |
| Max | 7.5 | 7 | 7.5 | 7.5 | 8.5 | 8.0 |
| Obs. | 40 | 51 | 53 | 70 | 44 | 69 |

## Table 2. Closed creativity: Determinants of creativity score

| **Creativity score** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Tournament incentives | 0.834*** | 1.589** | 1.428** | 1.411** | 2.046** | 1.378** |
| | [0.303] | [0.682] | [0.700] | [0.683] | [0.885] | [0.669] |
| Ranking | 1.496*** | 1.198*** | 1.372*** | 1.480*** | 1.160*** | 1.126*** |
| | [0.312] | [0.375] | [0.410] | [0.403] | [0.377] | [0.411] |
| Count | | 0.005*** | 0.005*** | 0.003* | 0.006*** | 0.006*** |
| | | [0.001] | [0.001] | [0.002] | [0.001] | [0.001] |
| Risk aversion | | 0.002 | 0.003 | 0.003 | 0.002 | -0.002 |
| | | [0.007] | [0.007] | [0.007] | [0.007] | [0.007] |
| Ambiguity aversion | | -0.011* | -0.011 | -0.012* | -0.011* | -0.010 |
| | | [0.007] | [0.007] | [0.007] | [0.007] | [0.007] |
| Incentives*risk | | 0.001 | -0.001 | 0.000 | 0.002 | -0.003 |
| | | [0.011] | [0.011] | [0.011] | [0.011] | [0.012] |
| Incentives*ambiguity | | 0.013 | 0.011 | 0.013 | 0.013 | 0.014 |
| | | [0.010] | [0.010] | [0.010] | [0.010] | [0.011] |
| Experimental creative style | | | | | | 0.230 |
| | | | | | | [0.267] |
| Sensation seeking | | | | | | 0.111 |
| | | | | | | [0.167] |
| Male | | | | | | 0.461* |
| | | | | | | [0.260] |
| Past involvement in artistic tasks | | | | | | -0.180* |
| | | | | | | [0.102] |
| Major: stem vs. social/humanities | | | | | | 0.170 |
| | | | | | | [0.259] |
| Right-handed | | | | | | -0.034 |
| | | | | | | [0.527] |
| Siblings | | | | | | -0.189* |
| | | | | | | [0.112] |
| Birth-order | | | | | | 0.075 |
| | | | | | | [0.149] |
| Taxonomy | | | 0.002 | 0.005 | | |
| | | | [0.003] | [0.003] | | |
| Math task | | | | -0.763** | | |
| | | | | [0.319] | | |
| Incentives*count | | | | | -0.002 | |
| | | | | | [0.003] | |
| Constant | 3.580*** | 2.484*** | 2.403*** | 2.941*** | 2.427*** | 2.727*** |
| | [0.236] | [0.394] | [0.458] | [0.500] | [0.400] | [0.846] |
| Observations | 137 | 124 | 122 | 122 | 124 | 120 |

Tobit regression
Standard errors in brackets
*** p < 0.01, ** p < 0.05, * p < 0.1

**Table 3. Open creativity: Determinants of creativity score**

| **Creativity score** | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Tournament incentives | -0.070 | 0.362 | 0.369 | 0.385 | 0.232 |
| | [0.227] | [0.536] | [0.536] | [0.794] | [0.536] |
| Ranking | 1.081*** | 1.237*** | 1.258*** | 1.238*** | 1.533*** |
| | [0.247] | [0.323] | [0.326] | [0.324] | [0.348] |
| Count | | 0.003*** | 0.003*** | 0.004** | 0.004*** |
| | | [0.001] | [0.001] | [0.002] | [0.001] |
| Risk aversion | | -0.001 | -0.001 | -0.001 | 0.003 |
| | | [0.006] | [0.006] | [0.006] | [0.006] |
| Ambiguity aversion | | -0.007 | -0.008 | -0.007 | -0.008 |
| | | [0.006] | [0.006] | [0.006] | [0.006] |
| Incentives*risk aversion | | 0.001 | 0.001 | 0.001 | -0.004 |
| | | [0.009] | [0.009] | [0.009] | [0.009] |
| Incentives*ambiguity aversion | | 0.007 | 0.008 | 0.007 | 0.011 |
| | | [0.008] | [0.008] | [0.008] | [0.008] |
| Experimental creative style | | | | | 0.521** |
| | | | | | [0.219] |
| Sensation seeking | | | | | -0.068 |
| | | | | | [0.163] |
| Male | | | | | 0.321 |
| | | | | | [0.218] |
| Past involvement in artistic tasks | | | | | 0.110 |
| | | | | | [0.086] |
| Major: stem vs. social/humanities | | | | | 0.287 |
| | | | | | [0.212] |
| Right-handed | | | | | -0.060 |
| | | | | | [0.355] |
| Siblings | | | | | -0.107 |
| | | | | | [0.076] |
| Birth-order | | | | | 0.285** |
| | | | | | [0.131] |
| Story1 | | | -0.102 | | |
| | | | [0.226] | | |
| Incentives*count | | | | -0.000 | |
| | | | | [0.003] | |
| Constant | 4.069*** | 2.880*** | 2.951*** | 2.871*** | 1.650** |
| | [0.188] | [0.412] | [0.441] | [0.469] | [0.655] |
| Observations | 190 | 162 | 162 | 162 | 152 |

Tobit regression.
Standard errors in brackets
*** p < 0.01, ** p < 0.05, * p < 0.1

## Table 4. Closed creativity. Determinants of creative score: Decile regressions

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | Tobit | DR_10 | DR_20 | DR_30 | DR_40 | DR_50 | DR_60 | DR_70 | DR_80 | DR_90 |
| Tournament incentives | 1.589** | 1.079 | 0.632 | 1.399* | 1.468* | 1.567* | 2.697*** | 1.817* | 2.528** | 2.741*** |
| | [0.682] | [1.178] | [0.869] | [0.821] | [0.811] | [0.858] | [0.931] | [1.045] | [1.214] | [0.959] |
| Ranking | 1.198*** | 1.451** | 1.468*** | 1.103** | 0.931** | 0.680 | 0.829 | 1.505*** | 1.515** | 1.350** |
| | [0.375] | [0.645] | [0.476] | [0.449] | [0.444] | [0.469] | [0.510] | [0.572] | [0.664] | [0.525] |
| Count | 0.005*** | 0.009*** | 0.007*** | 0.005*** | 0.005*** | 0.005*** | 0.004** | 0.002 | 0.001 | 0.002 |
| | [0.001] | [0.002] | [0.002] | [0.001] | [0.001] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] |
| Risk aversion | 0.002 | 0.012 | 0.018** | 0.000 | -0.002 | -0.002 | -0.005 | -0.007 | -0.011 | -0.003 |
| | [0.007] | [0.012] | [0.009] | [0.008] | [0.008] | [0.009] | [0.009] | [0.010] | [0.012] | [0.010] |
| Ambiguity aversion | -0.011* | -0.012 | -0.009 | 0.000 | -0.002 | -0.002 | -0.004 | -0.009 | -0.010 | -0.033*** |
| | [0.007] | [0.012] | [0.008] | [0.008] | [0.008] | [0.008] | [0.009] | [0.010] | [0.012] | [0.009] |
| Incentives*risk aversion | 0.001 | 0.006 | -0.005 | 0.015 | 0.011 | 0.004 | 0.011 | 0.004 | 0.020 | 0.002 |
| | [0.011] | [0.019] | [0.014] | [0.013] | [0.013] | [0.014] | [0.015] | [0.017] | [0.019] | [0.015] |
| Incentives*ambiguity aversion | 0.013 | -0.001 | 0.001 | -0.004 | 0.004 | 0.010 | 0.013 | 0.015 | 0.001 | 0.031** |
| | [0.010] | [0.018] | [0.013] | [0.012] | [0.012] | [0.013] | [0.014] | [0.016] | [0.018] | [0.014] |
| Constant | 2.484*** | 0.991 | 1.955*** | 2.640*** | 2.711*** | 2.965*** | 3.068*** | 3.192*** | 3.352*** | 3.105*** |
| | [0.394] | [0.680] | [0.502] | [0.474] | [0.468] | [0.495] | [0.538] | [0.603] | [0.701] | [0.553] |
| Observations | 124 | 124 | 124 | 124 | 124 | 124 | 124 | 124 | 124 | 124 |

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

## Table 5. Open creativity. Determinants of creative score: Decile regressions

| | (1) | -2 | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | Tobit | DR_10 | DR_20 | DR_35 | DR_40 | DR_50 | DR_60 | DR_70 | DR_80 | DR_90 |
| Tournament incentives | 0.362 | 1.353** | 0.481 | 0.303 | 0.286 | -0.485 | -0.325 | -0.610 | 0.151 | 0.333 |
| | [0.536] | [0.685] | [0.712] | [0.686] | [0.695] | [0.801] | [0.949] | [0.933] | [0.831] | [0.926] |
| Ranking | 1.237*** | 0.837** | 1.333*** | 1.323*** | 1.302*** | 1.870*** | 1.624*** | 1.486*** | 1.476*** | 1.000* |
| | [0.323] | [0.413] | [0.429] | [0.413] | [0.419] | [0.482] | [0.572] | [0.562] | [0.501] | [0.558] |
| Count | 0.003*** | 0.006*** | 0.006*** | 0.005*** | 0.006*** | 0.005*** | 0.004* | 0.002 | 0.001 | 0.000 |
| | [0.001] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] | [0.002] |
| Risk aversion | -0.001 | -0.006 | -0.004 | -0.003 | -0.003 | 0.010 | 0.013 | 0.014 | -0.000 | 0.000 |
| | [0.006] | [0.008] | [0.008] | [0.008] | [0.008] | [0.009] | [0.010] | [0.010] | [0.009] | [0.010] |
| Ambiguity aversion | -0.007 | -0.005 | -0.007 | -0.010 | -0.011 | -0.014 | -0.017 | -0.016 | -0.006 | -0.000 |
| | [0.006] | [0.007] | [0.008] | [0.007] | [0.008] | [0.009] | [0.010] | [0.010] | [0.009] | [0.010] |
| Incentives*risk aversion | 0.001 | 0.019* | 0.005 | -0.005 | -0.003 | -0.014 | -0.018 | -0.021 | 0.003 | 0.011 |
| | [0.009] | [0.011] | [0.012] | [0.011] | [0.011] | [0.013] | [0.015] | [0.015] | [0.013] | [0.015] |
| Incentives*ambiguity aversion | 0.007 | 0.006 | 0.006 | 0.013 | 0.013 | 0.015 | 0.021 | 0.015 | 0.005 | -0.006 |
| | [0.008] | [0.011] | [0.011] | [0.011] | [0.011] | [0.013] | [0.015] | [0.015] | [0.013] | [0.014] |
| Constant | 2.880*** | 0.850 | 0.945* | 1.363** | 1.336** | 2.304*** | 3.096*** | 4.140*** | 4.813*** | 6.000*** |
| | [0.412] | [0.527] | [0.548] | [0.528] | [0.535] | [0.616] | [0.730] | [0.718] | [0.640] | [0.712] |
| Observations | 162 | 162 | 162 | 162 | 162 | 162 | 162 | 162 | 162 | 162 |

Standard errors in brackets
*** p < 0.01, ** p < 0.05, * p < 0.1

**Table 6. Structural Parameter Estimates on experimental data**

| **Parameter** | Tournament | Ranking |
|---|---|---|
| $\beta_1$ | 0.910*** | 0.647*** |
| | (0.244) | (0.190) |
| $\beta_2$ | 0.190*** | |
| | (0.023) | |
| $\beta_5$ | - | .898*** |
| | | (0.216) |
| Constant | 0.472 | 1.120 |
| | (1.292) | (0.830) |
| $\sigma_s$ | 1.207 | 1.363 |
| | (0.057) | (0.068) |
| Observations | 224 | 203 |
| Log Likelihood | -360.07 | -350.96 |
| Wald Chi2(2) | 87.61 | 51.42 |

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

## Table 7. Counterfactual incentive schemes

| Parameter | (1) Tournament (empirical) | (2) Tournament (predicted) | (3) Piece Rate (predicted) | (4) Target Bonus (predicted) |
|---|---|---|---|---|
| $\beta_1$ | .910*** | .901*** | .969*** | .698*** |
| | (0.244) | (0.129) | (0.017) | (0.098) |
| $\beta_2$ | .190*** | 0.190*** | n.a. | n.a. |
| | (0.023) | (0.008) | | |
| $\beta_3$ | n.a. | n.a. | .101*** | n.a. |
| | | | (0.001) | |
| $\beta_4$ | n.a. | n.a. | n.a. | .020*** |
| | | | | (0.001) |
| Constant | .472 | .526*** | .190* | 2.233*** |
| | (1.292) | (0.151) | (0.099) | (0.603) |
| $\sigma_s$ | 1.207** | .446*** | .051*** | .367*** |
| | (0.057) | (0.010) | (0.001) | (0.008) |
| Average creative output | 5.909 | 6.077 | 6.405 | 6.737 |
| Average payment | 9 | 9 | 9 | 9 |
| N. obs | 224 | 1000 | 1000 | 1000 |
| Log Likelihood | -360.07 | -611.73 | 1548.56 | -416.38 |
| Wald $\chi^2$ | 87.61 | 312.05 | 100814.27 | 1127.04 |

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

# Appendix A

## *A1: Instructions for the closed task (Flat payment Treatment and No-ranking Treatment)*

Thank you for coming to our experiment. You will receive a payment of $5 for showing up, plus additional money as explained below.

You will be in groups of 5 people. You are asked to answer to one of these questions:

1. **Create an interesting story using all the words you like: the only constraint you have is that you must use all of the following words, in addition to whatever other words you wish.**

   *house, zero, forgive, curve, relevance, cow, tree, planet, ring, send.*

2. **Choose a combination of numerical operations to start with the number 27 and reach the number 6. Here are some examples of such operations, but you are not limited to these.**

*addition, subtraction, multiplication, division, exponential, factorial, logarithm.*

For example: (**27**:3) **–** 3 = **6**.

Please choose one of these questions and respond on the sheets of paper that we have supplied.

*Each of the five people in another group will rank the level of creativity for each person in your 5-person group. You will do the same for the people in another 5-person group. The evaluator will not know which group he or she is evaluating. [deleted in case of the No-ranking Treatment]*

You will receive an addition an additional payment of $9 for answering the question as creatively as possible.

You have 20 minutes to complete your answer. Afterwards, we will collect all the answers, ask people to evaluate the response, and ask people to complete a final questionnaire.

Are there any questions?

## *A2: Instructions for the closed task (Tournament incentives Treatment)*

Thank you for coming to our experiment. You will receive a payment of $5 for showing up, plus additional money as explained below.

You will be in groups of 5 people. You are asked to answer to one of these questions:

1. **Create an interesting story using all the words you like: the only constraint you have is that you must use all of the following words, in addition to whatever other words you wish.**

*house, zero, forgive, curve, relevance, cow, tree, planet, ring, send.*

2. **Choose a combination of numerical operations to start with the number 27 and reach the number 6. Here are some examples of such operations, but you are not limited to these.**

*addition, subtraction, multiplication, division, exponential, factorial, logarithm.*

For example: (**27**:3) – 3 = **6**.

Please choose one of these questions and respond on the sheets of paper that we have supplied.

Each of the five people in another group will rank the level of creativity for each person in your 5-person group. You will do the same for the people in another 5-person group. The evaluator will not know which group he or she is evaluating.

We will pay people on the basis of the assessments made by the 5 judges. The person with the best ranking will receive an additional $15, the person with the second-best ranking will receive an additional $12, the person with the third-best ranking will receive an additional $9, the person with the fourth-best ranking will receive an additional $6, and the person with the fifth-best ranking will receive an additional $3.

You have 20 minutes to complete your answer. Afterwards, we will collect all the answers, ask people to evaluate the response, and ask people to complete a final questionnaire.

Are there any questions?

*A3: Instructions for the open task (Flat payment Treatment and No-ranking Treatment)*

Thank you for coming to our experiment. You will receive a payment of $5 for showing up, plus additional money as explained below.

You will be in groups of 5 people. You are asked to answer to one of these questions:

**1)"If you had the talent to invent things just by thinking of them, what would you create and why?"**

**2) "Imagine and describe a town, city, or society in the future."**

*Each of the five people in another group will rank the level of creativity for each person in your 5-person group. You will do the same for the people in another 5-person group. The evaluator will not know which group he or she is evaluating. [deleted in case of the No-ranking Treatment]*

You will receive an addition an additional payment of $9 for answering the question as creatively as possible.

You have 20 minutes to complete your answer. Afterwards, we will collect all the answers, ask people to evaluate the response, and ask people to complete a final questionnaire.

Are there any questions?

*A4: Instructions for the open task (Tournament incentives Treatment)*

Thank you for coming to our experiment. You will receive a payment of $5 for showing up, plus additional money as explained below.

You will be in groups of 5 people. You are asked to answer to one of these questions:

**1)"If you had the talent to invent things just by thinking of them, what would you create and why?"**

**2) "Imagine and describe a town, city, or society in the future."**

Each of the five people in another group will rank the level of creativity for each person in your 5-person group. You will do the same for the people in another 5-person group. The evaluator will not know which group he or she is evaluating.

We will pay people on the basis of the assessments made by the 5 judges. The person with the best ranking will receive an additional $15, the person with the second-best ranking will receive an additional $12, the person with the third-best ranking will receive an additional $9, the person with the fourth-best ranking will receive an additional $6, and the person with the fifth-best ranking will receive an additional $3.

You have 20 minutes to complete your answer. Afterwards, we will collect all the answers, ask people to evaluate the response, and ask people to complete a final questionnaire.

Are there any questions?

# Appendix B

## *Final Questionnaire*

1) You are endowed with 100 units and can invest any portion of this amount in an asset that has a 50% chance of success and pays 2.5 times the amount invested if successful; you can retain the units not invested. Please note that a regular coin will be tossed at the end of the experiment, and one of you will be selected randomly and paid according the amount stated in this question. How much do you want to invest in this asset? ……….

2) You are endowed with 100 units and can invest any portion of this amount in an asset with unknown chance of success are and pays 2.5 times the amount invested if successful; you can retain the units not invested. Please note that a regular coin will be tossed at the end of the experiment, and one of you will be selected randomly and paid according the amount stated in this question. How much do you want to invest in this asset? ……….

3) You are presented with 10 couples of sentences: in each couple, please pick the one that better describes your personality.

   1A. Planning is essential for me to be creative. I often have detailed sketches for what I am going to do before I do anything.
   1B. Planning is not important for me to be creative. I rarely have detailed sketches for what I am going to do before I do anything.

   2A. I view working creatively as the systematic execution of a plan; I work easily and swiftly.
   2B. I view working creatively as mainly trial and error; I make choices, change them, and react to my changes.

   3A. I have a discontinuous creative career. Once I master one idea or topic, I move on to the next.
   3B. I am a perfectionist who is constantly searching. I am frustrated by my inability to achieve my goals.

   4A. I am finished working creatively when I complete my preconceived plan.
   4B. I am finished working only after inspecting and judging my work.

   5A. When working creatively, I precisely state my goals before beginning, either as an image or an exact procedure.
   5B. When I am working creatively, my goals are imprecise. Having imprecise goals leads me to use a tentative procedure.

   6A. I work creatively to produce something that achieves a purpose.
   6B. I work creatively to search for and discover the meaning of my work.

   7A. My innovation appears suddenly. My new ideas are very different from my old ideas.
   7B. My innovation appears through pursuing one image at a time. My new ideas tend to be different versions of the same thing.

   8A. I get bored seeing the same old faces.
   8B. I like the comfortable familiarity of everyday life.

   9A. I like to explore a strange city or section of town by myself, even if it means getting lost.
   9B. I prefer a guide when I am in a place I don't know well.

   10A. If I were a salesman I would prefer a straight salary, rather than the risk of making little or nothing on a commission basis
   10B. If I were a salesman I would prefer working on a commission if I had a chance to make more money than I could on a salary.

4) Please indicate your

- Gender F M
- Age …………
- Major ……………………………………………………
- Number of siblings ……….
- Birth-order 1st 2nd 3rd 4th 5th more than 5th
- Right/Left-handed R L
- Married/Unmarried/Divorced parents M U D

5) Have you ever

| | | |
|---|---|---|
| (a) painted an original picture | YES | NO |
| (b) wrote an original computer program | YES | NO |
| (c) composed a poem or a novel/tale | YES | NO |
| (d) composed a song | YES | NO |
| (e) made a craft item | YES | NO |
| (f) directed or acted in a play | YES | NO |

## Appendix C

*Taxonomy – verbal closed task*

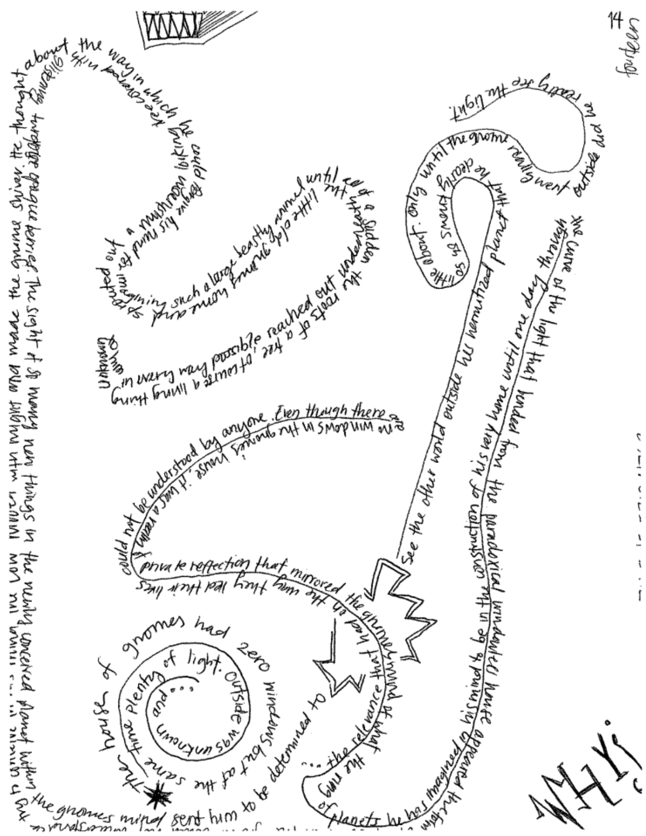
| Word / Meaning | Score | Word / Meaning | Score |
|---|---|---|---|
| **House** | | **cow** | |
| 1. residence building | 1 | 1. female bovine | 1 |
| 2. verb: keep in a dwelling | 2 | 2. figurative: unpleasant/fat woman | 2 |
| 3. verb: provide a storage place | 3 | 3. figurative: person who eats a lot | 3 |
| 4. family/household | 4 | 4. others | 10 |
| 5. shelter | 5 | | |
| 6. legislative body | 6 | **tree** | |
| 7. members of a college | 7 | **meaning** | |
| 8. convent, abbey, church | 8 | 1. plant | 1 |
| 9. others | 10 | 2. diagram | 2 |
| | | 3. tree-like shrub | 3 |
| **Zero** | | 4. tree-like stand | 4 |
| 1. number | 1 | 5. others | 10 |
| 2. figurative: starting point, absence | 2 | | |
| 3. figurative: unimportant person | 3 | **planet** | |
| 4. verb: change to zero | 4 | **meaning** | |
| 5. others | 10 | 1. Mars, Venus… | 1 |
| | | 2. others | 10 |
| **Forgive** | | **ring** | |
| 1. pardon/stop resenting | 1 | 1. jewelry worn on finger | 1 |
| 2. cancel a debt | 2 | 2. circular band | 2 |
| 3. others | 10 | 3. sund of a bell | 3 |
| | | 4. circular shape | 4 |
| **Curve** | | 5. verb: sound of a bell/telephone | 5 |
| 1. line or form that bends | 1 | 6. verb: draw a circle around | 6 |
| 2. bend in a road | 2 | 7. circle of people/objects | 7 |
| 3. verb: bend, not be straight | 3 | 8. arena for circus/boxing | 8 |
| 4. others | 10 | 9. cooking hob | 9 |
| | | 10. others | 10 |
| **Relevance** | | **send** | |
| 1. effect, connection | 1 | 1. cause to go/deliver | 1 |
| 2. others | 10 | 2. emit | 2 |
| | | 3. informal: delight | 3 |
| | | 4. others | 10 |

*Taxonomy – math closed task*

| **Operations** | **School level** | **Score** |
|---|---|---|
| + , - , * , : , fractions | Elementary | 1 |
| exp, log, roots, equations, inequalities | Secondary (Middle and High) | 3 |
| integral, factorial, matrixes, trigonometrics, limits, derivatives | University | 6 |
| Others | | 10 |

# Appendix D

*D1: Example of an answer to the closed task (verbal)*



*D2: Example of an answer to the closed task (math)*

[7]

$$\Big[27 - \sum_{i=2}^{5} \int_0^{\pi} i\sin(ix)\,dx \div \sqrt[3]{27} - \prod_{i=2}(3i) + \frac{25}{5^2} \cdot \frac{7}{2}$$

$$+ \int_0^{\pi}\int_0^{i}\int_0^{3} x^2\,dx \cdot \frac{1}{i\pi} - (8^2)^{1/3} + \cos(\sin(\cos(\tfrac{\pi}{2}))) -$$

$$i^2 \cdot \frac{1}{i^2} - \log\left[\left(\lim_{x\to\infty} \frac{x^3-x^2-1}{x^3+x^2}\right) \Big/ \left(-\frac{2}{7} \cdot \frac{14}{3} \cdot \frac{6}{4}\right)\right] +$$

$$\frac{\partial}{\partial x} f(x^7/7 - x^6/6)\Big|_0^1 \cdot \sum_{i=1}^{\infty} \int_0^{\infty}\int_0^{\theta}\int_1^{\frac{\phi-x^2}{2}} \rho^2 \sin\phi\, e^{\phi-1}\, d\rho\, d\phi\, d\theta -$$

$$\lim_{t\to\infty}\int_0^{t} e^{\left(x^2 - \ln x^e/e^{1\cdot 2} + \frac{\partial}{\partial\phi} f(x\phi)\right)}\, dt \cdot \frac{1}{t} + 10\Big]^{-1/2} \cdot 2^3 +$$

$$\Big[\int_0^{\pi/2} \sin\theta\, d\theta + 5/2 + \frac{17}{6} \cdot \frac{18}{51} + 0.5\Big]^2 - e^{\ln 4}$$

$$- \sqrt{\sqrt{\sqrt{i^6}}} \cdot \frac{6^{-2}}{6^2} + e^{\infty^{1/2}} = 6 \checkmark$$

*D3: Example of an answer to the open task*

Question 1

(1) If I had the talent to invent things by just thinking of them, I would invent a certain kind of peanut butter jar that would allow you to get more peanut butter out of the jar and not waisting any. The jar would have a twisty bottom (if you think of the bottom of a deodorant, it's basically the same concept). When you twist the bottom of the jar, the peanut butter gets pushed up to the top, making it easier to reach with a knife or spoon. I think we can all agree that when we throw away a peanut butter jar, there's always some left in it, but with the twisty bottom, you get more out of the jar than you did before and waste less. Plus, the last bit of peanut butter in the jar is always a struggle to get out — but wouldn't be with the twisty bottom

The same concept can be used for mayonaise :)