Peers and productivity: Evidence from an experimental factory

Reyniers, Diane

London School of Economics

November 2018

Peers and productivity: Evidence from an experimental factory

Diane Reyniers*

London School of Economics

This version: November 2018

ABSTRACT

Workplace peer effects are well documented, but why they arise remains a puzzle. This paper investigates the issue experimentally. Subjects are brought together to perform a real-effort task in a simulated factory environment. Varying the returns to effort by altering free-riding incentives or piece rates does not affect productivity but psychological factors do matter. Even though there are no technological complementarities, co-workers' productivity levels are highly correlated. Three psychological mechanisms which can generate these correlations are examined: (a) workers' desire to conform to a work norm, (b) inequity aversion and (c) concern about relative performance. Subjects' enjoyment of the task depends on their relative performance and not on how close their productivity is to the norm or on the inequity of outcomes. This finding suggests that peer effects arise because of intrinsic competitiveness. Subjects hate to do worse than their co-workers and love to do better.

Keywords: Peer Effects, Job satisfaction, Relative Performance Concerns, Social Norms

JEL Classification: C91, C92, D23, J01, J24

* Corresponding author. D.j.reyniers@lse.ac.uk, LSE, Houghton Street, London WC2A2AE.

442079557921.

1. INTRODUCTION

There is huge heterogeneity in group performance. For example, Syverson (2004) finds that plants at the 90th percentile of the productivity distribution produce double the output with the same measured inputs as those at the 10th percentile. A possible explanation is peer effects. People are influenced by the behaviour of those around them. Multiple equilibria are then on the cards. A good equilibrium in which all work assiduously is possible, but so too is a bad equilibrium in which a culture of lethargy prevails. Several experimental and field studies, discussed below, document the tendency to do what others do. Less studied is why a correlation in productivity of co-workers. This paper focusses on the psychology of work norms and propose a new psychological explanation as to why peer effects arise. The setting is an experimental candle factory. Subjects care about their relative performance and dislike falling behind their peers. This implies that they may slow down when their co-workers are slow and speed up if they are in a fast group. It is as if they are in a race with their fellow workers, even though there is no tournament element in the pay schedule. As far as we are aware, this relative performance motive has not thus far been documented in the peer effects literature. It stands in contrast to other explanations such as conformism (the desire to behave like one's peers) or inequity aversion which can be experimentally distinguished.

In the study, subjects' enjoyment of the experimental task is significantly positively correlated with the ratio of their output to the mean output of others in their small group. Subjects who produced the largest (smallest) quantity in their group (controlling for their absolute output level) enjoyed the task most (least). Furthermore, job satisfaction is not decreasing in deviation from the norm or the level of inequity.

These findings suggest that competitiveness may be rather more important than conformism or inequity aversion in the generation of work norms.

The experiment is set in a simulated (candle making) factory. Subjects are randomly allocated to sessions (timeslots) and each session accommodates up to four subjects. Sessions are randomly allocated to one of four financial incentives treatments. In two of these treatments, individuals are paid a piece rate (of £1 and £0.50) for their own output only. In the other two treatments, subjects are paired and paid the same piece rates for *the average output of the two pair members*. Productivity is not significantly affected by these economic factors. Output does not increase when a higher piece rate is on offer and in contrast with standard economic theory predictions we also find no evidence for free-riding in the pair treatments. Psychological factors do matter. Agreeableness is significantly negatively correlated with individual productivity and conscientiousness is significantly positively correlated. In all treatments, subjects' productivity is highly correlated within sessions. Although all subjects performed the same task (candle making) under identical conditions, in some sessions average output is as low as 6 candles whereas in other sessions it is over 16 candles and within-session variation is very low relative to between-session variation. This pattern reflects the field data. Were productivity differences down to individual differences, within-session dispersion would exceed between-session.

There are at least two broad classes of possible explanations for the finding that within-session variance of productivity is low. Firstly, an information effect may be involved. This could be either in the form of learning (from others in the room) how to work faster (knowledge spillovers) and/or in the form of deducing what the cost of

effort is, however the scope for learning is limited and the results (see Section 4.6) suggest that the information effect is not responsible for generating work norms. Secondly, peer effects can be caused by a number of psychological mechanisms. There may be a norm effect because subjects are driven by a desire to conform to the group productivity norm. As the famous Asch (1951) experiments (and many subsequent studies) show, individuals have a desire to belong or not to be seen to be different. They are likely to conform to the judgment of others, especially in face - to - face situations. Subjects may be concerned with how they are perceived by other subjects and suffer a loss in utility from not conforming (see e.g. Bernheim, 1994). Alternatively, as suggested by Mohnen et al. (2008), subjects may be inequity averse. Yet another possible psychological explanation involves subjects trying to do at least as well as others in the room (the competitiveness or relative performance hypothesis). [1]Subjects may wish to avoid shame (and guilt in the pair treatment) associated with underperforming relative to their peers. Hence, in sessions where most subjects are slow, nobody has a strong incentive to produce much output whereas if there are some fast workers in the group this tends to elicit more effort from everyone.

The results favour this last hypothesis: although all the subjects are rewarded by piece rates (and hence others' output should not affect own effort as it might in tournaments), it appears that they are motivated by a desire to perform at least as well as their co-workers. Controlling for absolute output level, subjects' reported enjoyment of the task increases in relative performance. [2] This effect is larger (in the

[1] Recent brain imaging research has found that outperforming other workers activates brain reward centres (Dohmen et al., 2011).

[2] Manski (2000) criticises economists' approach to the study of social interactions which is based on observing outcomes since outcomes can usually be generated by a variety of interaction patterns which leaves the findings open to a range of possible interpretations. In this context he discusses economists'

pair treatment) if subjects know each other. This suggests that in real workplaces, where co-workers do usually know each other, relative performance concerns may be even more important.

A number of studies investigate workplace peer effects. Charnes et al.'s (2010) find subjects perform better in an experimental setting when their relative performance is revealed, even though their monetary payoff does not depend on performance.[3] In a fascinating natural experiment, Blanes i Vidal and Nossol (2011) find large increases in productivity (6.8% on average) after the introduction of relative performance evaluation. In addition to a fixed wage, workers received piece rate payments. Financial incentives are kept constant and the only change is the private revelation of worker's rank in the productivity distribution. On introduction of this scheme, workers at all productivity levels increase effort. Kuhnen and Tymula (2012) also find that providing feedback on relative performance increases productivity. Their setting is experimental and subjects receive a flat wage.

This paper provides exceptional thorough evidence of the presence of peer effects in a controlled but realistic setting but its main contribution is to establish the underlying behavioural cause of peer effects. A real-effort experiment is used to simulate a workplace environment. Subjects can observe each other constantly and, importantly, are aware of their relative productivity. There is strong evidence of peer effects and the results suggest that these effects are driven by competitiveness rather than conformism or inequity aversion.

hostility to subjective data and suggests it is misplaced and that elicitation of subjects' perceptions of the interactions in which they participate would be useful.

[3] About 12% of subjects were willing to sacrifice income to improve their rank.

The rest of the paper is organised as follows. Section 2 contains a literature review. Section 3 describes the experimental design. Section 4 presents the peer effect results. In Section 5 the job satisfaction (task enjoyment) results are discussed and section 6 concludes.

2. LITERATURE REVIEW

The results relate to the vast literature on happiness as a relative, reference-dependent phenomenon, originating from Easterlin's (1974) observation that substantial economic growth in Western countries has not resulted in a corresponding growth in happiness levels.[4] Luttmer (2005), for example, finds that, controlling for own income, higher earnings of neighbours are associated with lower levels of self-reported happiness.[5] Several studies document that comparison income is negatively correlated with job satisfaction.[6] In Brown et al. (2008), the wage rank of an individual in the firm wage distribution, controlling for own wage, correlates positively with satisfaction.[7] Whilst this literature emphasises the satisfaction people get from a favourable comparison of their financial outcome, the (pair treatment) results suggest that individuals, possibly as a result of the desire to maintain a positive self-image, may be driven to outperform their peers even in the absence of financial consequences.

[4] See e.g. McBride (2001), Graham and Felton (2006), Helliwell and Huang (2009).
[5] Similarly, Knight and Song (2009), using Chinese data, report that households whose income is above their village average are far happier than those whose income is below the average.
[6] See e.g. Clark and Oswald (1996), Levy-Garboua and Montmarquette (2004), Sloane and Williams (2000).
[7] See also Clark (1996) who shows that job satisfaction is lower, the higher the wages of other workers in the household.

There is a well-established literature in psychology on peer effects. The economics and psychology literatures in this area appear remarkably ignorant of each other. In psychology, the concept of motivation gains in social contexts compared to isolated working conditions has received a great deal of attention since the pioneering studies by Otto Kohler (1926, 1927). [8] The 'Kohler effect', whereby inferior group members increase their effort in the presence of more capable co-workers, is attributed to two mechanisms, 'social comparison' and 'indispensibility of effort'. The social comparison explanation involves individuals revising their performance goals upward when paired with a more competent subject. [9] The indispensibility of effort explanation involves individuals' concern about 'holding back' the group when a team outcome is strongly dependent on their effort. Whose effort is most indispensible depends on the nature of the task and on the payoff structure. For example, the most capable member's effort is most indispensible if the best performance determines the group outcome (Kerr and Hertel, 2011).

Parallel to the psychology literature, there is a recent and rapidly growing economics literature on peer effects in field and laboratory experiments. The Ichino and Maggi (2000) study of a large Italian bank hints at performance differentials generated by work norms. They report substantial variation in absenteeism and misconduct episodes between bank branches in spite of identical financial incentives. Although individual worker characteristics play an important role in explaining these differences, there is evidence of a strong work environment effect: a positive relationship exists between a worker's shirking level and the average shirking level of his co-workers.

[8] For a meta-analysis of the Kohler effect see Weber and Hertel (2007).
[9] For evidence of increased effort when learning about being outperformed, see e.g. Hertel et al. (2000), Kerr et al. (2007) and Hertel et al. (2008).

The Falk and Ichino (2001) experiment on peer effects shows that the mere presence of another subject can induce subjects to work harder. They report that a 10% increase in a peer's output results in a 1.4% increase in subjects' effort. Subjects were paid a fixed rate and in the base treatment worked alone. In the 'pair' treatment, pairs of subjects worked in the same room, in sight of each other but were not allowed to combine efforts. The paired subjects significantly outperformed the lone subjects. Output levels within pairs were very similar whereas there was substantial variation between pairs. In the Falk-Ichino setup, there is a possibility that the higher output in pairs is caused by the major difference in working conditions (solitary versus non-solitary); subjects may have been more motivated to work hard if they considered their work environment more pleasant. [10] Zajonc (1965) reviews the psychology literature on precisely this phenomenon of 'social facilitation', the consequences for behaviour of the sheer presence of others. The results suggest that social facilitation is not the only possible explanation for the Falk-Ichino results however. [11] In this setting, no subject worked on their own so this aspect of the work environment is held constant. Yet, we also find evidence of 'pure' peer effects.

The field study by Bandiera et al. (2010) shows that, in a piece rate regime, workers who are friends conform to a common productivity norm. Workers are asymmetrically influenced by the norm in that the presence of friends increases productivity of workers who are less able than their friends and decreases productivity of workers who are more able than their friends. Whereas Bandiera et al. (2007) find

[10] Note that Eriksson et al. (2008) don't find peer effects for their paired subjects (under piece rates) but in their study the work environment is held constant in that the presence of a co-worker is not salient; the only evidence of a co-worker is information about the latter's performance displayed on a computer screen.

[11] Although the findings of Falk and Ichino (2001) are consistent with the hypothesis that workers experience disutility if they fall behind their peers, this hypothesis is not explicitly explored.

norm effects for workers who are friends, here similar results apply to all subjects independent of social ties. Bandiera et al (2007) interpret their findings as representing workers' desire for conformity or possibly wage equality but their results are not inconsistent with workers caring about their performance relative to their friends. If workers have a dislike of being outperformed (too much) they will speed up in the presence of more able friends and slow down (because they can without too much loss in utility) in the presence of less able friends.

Mas and Moretti (2009), in a study of cashiers in a national supermarket chain, find positive productivity spillovers when highly productive co-workers are introduced. According to their estimate, a 10% increase in co-worker productivity is associated with a 1.5% rise in worker's effort. Mutual monitoring appears important in their setting as there are negative externalities: a slow worker may increase the workload of co-workers. In the experimental setting there is no moral hazard problem of this type. Guryan et al. (2009) find that the scores of professional golfers are *not* influenced by the golfers they are drawn to play with. They suggest that social incentives may be less powerful when financial incentives are strong or that with experience workers learn not to be affected by social forces. Also, professional golfers may be selected to be immune to peer effects. It is likely that golf players (and other professional sports people) are not very 'agreeable'[12] and hence less easily influenced by performance norms. [13] In addition, experience and ability matter more in golf than in the tasks considered in other studies on peer effects. If some workers are unable or too inexperienced to keep up with their co-workers, peer effects are less likely to materialise. There is of course a major difference between a golf tournament setting

[12] We show in section 4.7 that 'agreeable' subjects produce output levels close to the norm.

[13] As Guryan et al. (2007) point out, there are undoubtedly selection effects in the sense that success in this occupation can only be achieved by individuals who are immune to peer pressure.

and the (and Falk and Ichino's) experimental setup: a golf player will naturally compare his performance to the performance of all participants in the tournament rather than to that of the player next to him. In the experiment, subjects do not find out how their performance compares to that of subjects in other sessions and their obvious reference group consists of other subjects in the same session.

In contrast to Blanes i Vidal and Nossol's (2011) results, Eriksson et al. (2009) find that feedback about relative performance does not change effort levels under piece rates. However, the scope for social incentives is much more limited in their experimental setup. Subjects don't know their 'co-worker', cannot observe him and the process is entirely anonymous with no present or possible future interaction between pair members. In other words, this setup does not model a workplace. The conjecture that peer effects are not likely to occur under piece rates because of subjects' focus on the monetary aspect of the work environment must be seen in this context and is not likely to hold in settings where social comparisons and social interactions between workers occur i.e. in most real life workplaces.

Similarly, Bellemare et al. (2010) fail to find peer effects in a setting where peer pressure is simulated by a signal representing the productivity level of an earlier subject. Clearly under these types of lab conditions there is no visceral stimulus to competition and the scope for emotional responses such as shame is very limited. Psychologists have identified significantly higher motivation gains for subjects who are in the physical (rather than virtual) presence of a co-worker (Lount et al., 2008). [14] Mohnen et al. (2009) seems to be the only other paper which explicitly investigates a possible cause of work norms. They compare a transparent setting, where subjects find out their partner's output after each round, with a non-transparent setting where

[14] Hertel et al. (2008) also find greater motivation gains in face-to-face than in anonymous settings.

no information about partner's productivity is revealed. In line with Falk and Ichino (2001), they find increased effort levels in the transparent setting which they attribute to inequity aversion. Only the pair treatments can be compared to Mohnen et al.'s setting, but even in the pair treatments, subjects report higher enjoyment of the task if they outperform their partner, a finding which is hard to reconcile with inequity aversion. This paper presents evidence that, at least in some circumstances, competitiveness or concern about relative performance rather than inequity aversion, may be an important factor in driving peer effects.

3. EXPERIMENTAL DESIGN

Subjects were recruited from the London School of Economics (Management) summer school student population. Potential subjects were told that they would receive a £5 participation fee in addition to performance related pay. 114 students were recruited and 38 sessions were run (with 2, 3 or 4 subjects at a time[15]). The age range of subjects was 19 to 43 with an average of 23 and median age 22, 67.5% were male and their nationalities were 48% European, 19% North American, 22% Asian, 11% other.

The experimental task consists of making candles. The aim was to select a real work task which is simple yet economically valuable, requires no prior knowledge and for which output could be easily measured. Subjects were seated at a desk and were provided with all the materials needed: beeswax, a hair dryer (to soften the wax), candle wick, a cutting board, scissors, a penknife, a pen (see Figure 1).

[15] Students signed up for specific time slots. The only constraints were that at least 2 subjects per slot were required and more than 4 subjects at a time could not be accomadated.

The candle making process involves cutting the wax and wick, heating up the wax, inserting the wick, rolling up the wax and sealing the wax. Subjects were told they would only get paid for candles which conformed to certain quality standards e.g. the candles should not be dented and should be able to stand up unsupported. All candles produced passed the quality control.

The room contained four work areas and subjects in the pair treatment were seated opposite each other. It is worth noting that all subjects in the same session could easily monitor each other's output. [16] Co-workers' output and effort levels were very salient.

Subjects were asked to fill out a personality questionnaire. We used the 10 item version of the Big-Five Inventory (Rammstedt and John, 2007) to obtain subjects' personality measurements. In this version, the five personality dimensions (extraversion, agreeableness, conscientiousness, neuroticism, openness) are assessed using two items for each. The five personality traits are presumed to be invariant to time and situation.

Each of the five personality traits – extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N), openness (O) - is measured as the sum of scores on two questions as indicated in Table 1 below.

After the administration of the personality questionnaire, subjects were shown a short candle making demonstration DVD and made a practice candle. They were then given 20 minutes to make as many candles as they could. After completion of the task, subjects filled out a questionnaire relating to the task.

[16] Mas and Moretti (2009) find that a worker's effort is positively related to the presence and speed of workers who can observe him but not to the presence and speed of workers whom he faces (and do not face him).

Four monetary incentive treatments were used varying along two dimensions: the magnitude of the piece rate (£1 or £0.50) and whether subject's payment depends on individual or team performance. In the two individual treatments subjects were paid according to their own output only. In the two pair treatments subjects were assigned a partner and were paid according to the total output of the pair with equal division of payoff. These payments were added to the £5 participation fee. When subjects arrived they were told which of the treatments they were allocated to i.e. individual versus pair and piece rate £1 versus £0.50. All subjects in a given session were allocated to the same treatment. Pair treatments were implemented in some of the sessions containing 2 or 4 subjects.

In both the individual and pair treatments, subjects had to work on their own. No technical assistance from other subjects was allowed but communication was. Subjects were monitored throughout the task. Subjects were seated around a square table in direct view of each other. For the pair treatment subjects were seated facing each other but were not allowed to combine efforts (no division of labor) so that the task was identical for all subjects.

4. PEER EFFECT RESULTS

To establish whether peer effects exist in the experimental factory, a number of tests were performed. Section 4.1 presents overall productivity results and some preliminary evidence of correlation in within-session output levels. Section 4.2. contains OLS results for individual output. In section 4.3, it is shown that peer effects exist for the individual treatments as well as the pair treatments. The presence of a high (low) productivity co-worker significantly increases (decreases) a subject's

productivity in section 4.4. In section 4.5 examine the productivity ranges for sessions is examined. Evidence that 'learning' is unlikely to be the cause of within-session correlation of outputs is in section 4.6. Finally, the role of personality factors in subjects' susceptibility to norms is examined in section 4.7.

4.1. Total sample output distribution and correlations in session and pair outputs

The output distribution for the entire sample is given in Table 2. Considering that the experimental task took only 20 minutes and consisted of well-defined operations, the productivity range is surprisingly large, with the least productive subjects producing 5 candles and the most productive 18. The top 10% produced two and half times the number of candles produced by the bottom 10%. Mean output is 9.8, median 9 and standard deviation 3.2.

There is no evidence of free riding in the pair treatment: mean output of subjects in the pair treatment (10 candles) is not significantly different from mean output of subjects in the individual treatment (9.7 candles). This result is in line with the experimental findings of van Dijk et al (2001) who find that individual and team payment induce the same effort levels. Although free riding occurred in some of their teams, this was compensated by many subjects providing *more* effort under team based pay than under individual pay. Possibly peer pressure and/or friendly competition among team members can explain this result. [17]In this study, the teams are very small (two subjects) and team members can easily monitor each other. In

[17] For an overview of the social psychology literature on free riding (or 'social loafing') see the meta-analysis by Karau and Williams (1993).

such a setting internal (guilt) and external (shame) peer pressure can easily mitigate the free-rider problem (Hamilton et al, 2003). [18]

Significant within-session productivity correlations are found. It appears that even within the context of a 20 minute task, small groups are able to develop work norms. 38 sessions were run. For only 2 of these sessions does the standard deviation exceed the total sample standard deviation. This strongly suggests that subjects regarded each other's speed as an anchor.

Figure 2 plots (for all treatments) subjects' output against the mean output of other subjects who were in the same session. The scatterplot shows a clear correlation between a subject's output and the session 'norm' i.e. the average output produced by other subjects in the same session.

Figure 3 plots the output of randomly chosen 'second' subject in the pair treatment against their partner's output. Output within pairs is highly correlated (r=0.79). The average standard deviation within pairs equals 0.7 whereas the standard deviation between pairs equals 6.1.

These scatter plots suggest that there are strong patterns in terms of within session and within pair output correlations. These norm effects are tested more formally in the rest of this section.

4.2. Output regressions

This section establishes the existence of norms by regressing individual output on mean output of other subjects in the session, controlling for treatment and personality effects.

[18] These findings are consistent with the Huck et al. (2003) model which shows how the free-rider problem under team incentives can be mitigated by work norms. This is illustrated by Knez and Simester (1999) who document an increase in productivity at Continental Airlines after the introduction (in 1995) of a team incentive scheme (bonus payments to all employees if a firm wide performance goal was reached). See also Hamilton et al. (2003) who use data from a garment plant that shifted from individual piece rates to group piece rates and found an improvement of 14% in average productivity.

Table 3 contains OLS regression results for individual output. These regression equations are subject to the reflection problem (Manski, 1993) in that output of worker i affects output of worker j which affects output of worker i, but the aim here is not to derive precise estimates of co-worker effects but merely to provide evidence of their existence.

The results in Table 3 show that the peer effects are robust to inclusion of treatment and personality effects.

In column (1) individual output (number of candles) is regressed on the mean output of **other** subjects in the same session.[19] It is evident that output levels within sessions are highly correlated. Whilst the coefficient on 'othersmean'in Table 3 cannot be interpreted as the effect of co-workers' productivity (because of endogeneity), these results clearly show that peer effects are very powerful in the experiment.

Financial incentive treatment dummies are included in column (2). The base (excluded) treatment is the pair treatment with piece rate £0.50 (pairat50). Standard economic theory predicts that the individual treatment at £0.50 piece rate (indat50) is equivalent to the pair treatment at £1 piece rate (pairat1) and that output in both of these treatments is lower than in the individual at £1 (indat1) treatment and higher than in the pair at £0.50 (pairat50) treatment. The findings do not correspond to these predictions: the average output levels for pairat50, pairat1, indat50 and indat1 are approximately 11, 8, 10 and 9, respectively.

[19] Guryan et al (2007) point out that one should expect a negative coefficient on average output of others in a group since for highly productive subjects the mean of others' output will tend to be low and vice versa. They suggest a correction which we haven't applied since this 'sampling without replacement' bias is negligible for the sample and the correction would only serve to increase the already highly significant peer effects.

Overall it seems that the magnitude of the piece rates and whether subjects worked for individual payment or for team payment (see column (3)) did not significantly affect their output. In the context of a 20 minute experimental task with relatively small amounts of money at stake, this is perhaps not surprising and can be explained by standard economic theory in that the cognitive cost of figuring out the optimal amount of effort to apply is too great relative to the range and magnitudes of possible payoffs. Column (4) shows that the number of subjects in the session has no systematic effect on output. Subjects in groups of 3 produce marginally less than subjects in groups of 2 or 4[20].

Column (5) includes personality measures (z-scores). Agreeableness has a significantly negative coefficient whereas, perhaps not surprisingly, the coefficient for conscientiousness is significantly positive. There is some evidence that agreeableness depresses earnings (see e.g. Mueller and Plug, 2006 and Heineck andAnger, 2008). Neuroticism has a large (although not significant) negative coefficient which is in line with the meta-analytic findings of Judge and Bono (2001) who report negative correlations of neuroticism with job performance (and job satisfaction). In another meta-analysis, Barrick and Mount (1991) find that conscientiousness is correlated with job performance in all occupational groups (professionals, police, managers, sales and skilled/semi-skilled). Salgado (1997), on the basis of a meta-analytical analysis of European studies, concludes that conscientiousness and emotional stability are valid predictors of job performance across criteria and occupational groups.

Age, gender[21] and nationality (by continent) are also included but are not significant. It also investigated whether time of day, a possible source of common productivity

[20] Interaction effects of group size and 'othersmean' were also included but they were not significant.
[21] Gender (and gender interaction effects) were included in all equations in the paper but did not find significance. Although laboratory experiments in competitive environments have found gender differences (e.g. Gneezy et al, 2003 ; Niederle and Vesterlund, 2007) recent studies in actual workplace

shocks, had a systematic effect (sessions were run at 1pm, 2pm, 3pm, 4pm and 5pm) but there are no significant differences in mean output.

These results confirm that individual output is strongly correlated with peer output and that some aspects of personality affect productivity. In particular, controlling for peer output, conscientious and non-agreeable subjects are more productive.

4.3. Peer effects exist in individual treatments as well as pair treatments

There are obvious reasons why outputs in the pair treatment are correlated. Subjects may have felt pressure not to let their partner down for example. It is less clear why outputs in the individual treatment, where there are no payoff externalities, are correlated. In this section it is argued that the results are not driven by the pair treatment alone.

4.3.1. Comparison of real pairs with nominal pairs

To show that the results are not driven by the pair treatment we calculated the correlation between outputs of artificial (nominal) pairs constructed by forming 'pairs' of individuals (from the same session) who participated in individual payment only sessions (16 sessions) and found a correlation of 0.70 whereas the correlation for 'real' pairs is 0.79 (not significantly different at 5%).

4.3.2. Comparison of output differences in pair and individual treatments

settings have failed to identify such differences (e.g. Blanes I Vidal, 2009; Antonovics et al.,2003; Lavy, 2013). Gender effects arise indirectly through the personality variables. In the sample, females have significantly higher scores than males on conscientiousness, neuroticism and openness.

The average absolute deviation from the other pair member's output in the pair treatment equals 1.4. The average absolute deviation from the mean of others in the same session for subjects in the individual treatment is 1.7. These means are not significantly different.

We conclude that peer effects are not limited to subjects allocated to the pair treatment.

4.4. A test for work norms: the List test

We devised a procedure to test whether allocation of a subject to a session which already contained at least one (other) high (low) productivity subject significantly increases the likelihood of the subject being high (low) productivity.

We divide the sample into high (H) and low (L) productivity subjects with H defined as output >9 (the median) and L defined as output ≤ 9. We then compile 2 lists of subjects, List(H) and List(L) containing subjects who were in the same session as the H and L subjects respectively. Table 4 presents the mean outputs for subjects on these lists and the proportions of H and L subjects on the lists.

In the absence of work norms, allocating a subject to a session which contains an L (H) subject should not affect the subject's probability of being L (H), namely 50%. However, List(L) contains 71% L subjects and List(H) contains 72% H subjects. A binomial test on these findings rejects the null hypothesis of no peer effects at less than 1% significance. The mean output levels on the two Lists should be equal in the absence of peer effects. However, we find a very significant difference ($t=5.3$, $p<0.0001$) with List(L) mean more than one candle below the sample average and List(H) mean more than 2 candles above the sample average. These results provide

strong evidence for peer effects; the presence of H (L) workers tends to increase (decrease) their co-workers' productivity.

4.5. A test for peer effects: the range test

The ranges of effort choices (as measured by output) for subjects in the *same* session are small whereas there are large differences between sessions. In this section we consider the likelihood of finding such small ranges in the absence of peer effects. We devised a test for peer effects on the basis of the output ranges for the sessions. We use the sample output distribution (see Table 2) to calculate the probability of finding a small range (≤ 2 for sessions with 2 subjects and ≤ 3 for sessions with 3 or 4 subjects) assuming subjects are randomly allocated to sessions. These results are presented in Table 5 with binomial test results on the number of sessions characterised by small ranges. This test again confirms that there is strong evidence for peer effects, the ranges of output levels within sessions are very small. The median ranges for group sizes 2, 3 and 4 are 1, 3 and 3, respectively.

4.6. A test for peer effects versus learning

The results so far indicate that there is conformity within sessions. A plausible cause of this conformity is that subjects take their cue on how hard to work from their co-workers. An alternative explanation, however, is that we are observing a learning effect. Although the candle production technology does not allow for deviations from the procedure explained in the training DVD, it is possible that we observe correlated outputs within sessions because subjects copy a successful co-worker (rather than conform to the speed of all co-workers). To address this concern, we start by noting that if learning causes correlated outputs then we would observe that the presence of

at least one high productivity worker (H) would generate a disproportionate fraction of H co-workers **and** that this proportion would not be affected by whether there was a low productivity (L) worker in the session.

We examined all sessions with at least one H subject (there are 23 such sessions) and for these sessions calculated the proportion of H co-workers of the chosen H subject. Out of the 50 co-workers, 31 are H. In line with results reported above, this fraction (31/50) is significantly above 0.5 ($z=1.75$, $p<0.05$). The mean output of co-workers of the chosen H subject equals 10.8 which is significantly ($p<0.05$) above the total sample average of 9.8.

Attention is limited to sessions which have at least one H and at least one L subject (11 sessions) and calculate the proportion of H co-workers of the chosen H and L subjects. Out of the 16 co-workers, 9 are H. This fraction (9/16) is not significantly different from 0.5 ($z=0.50$). The mean output of co-workers of the chosen H and L subjects equals 9.7, which is actually slightly below the average sample output.

These results provide some evidence that what we are observing is not merely a learning effect. Imitation of an H worker should not be hindered by the presence of an L worker. However, we find that the presence of an L worker (in addition to at least one H worker) tends to make it less likely that a subject is H and reduces the subject's output.

4.7. Susceptibility to norms

Given the existence of work norms, it may be interesting to check whether certain personality types are more driven to conform to the norm than others. If this is the

case then the likelihood of norms emerging may well vary by occupation since personality types vary by occupation. [22]

In this section we examine whether deviation from the norm (measured as the absolute value of the difference of the subject's output level and mean output of others in the same session) correlates with personality traits.

Columns (1) – (4) in Table 6 contain the regression results of deviation from the norm on the five personality dimensions (z-scores). For the pairs regressions in columns (1) and (3), the dependent variable is the absolute difference between the subject's output and his partner's output; for the regressions on the entire sample in columns (2) and (4) the dependent variable is the absolute difference between the subject's output and the mean output of others in the same session.

Subjects appear to be heterogeneous in how they respond to norms. There is some indication here that agreeableness correlates with susceptibility to work norms[23]. This result is particularly striking for the pair treatment where an increase of one standard deviation in agreeableness reduces the difference with pair member's output by about 0.7, a reduction of 50%.

Although not significant, there are also non-trivial effects for conscientiousness and openness in the pair treatment with the conscientious sticking closer to their partner's output and the open less close. In the peer pressure model of Kandel and Lazear (1992), deviations from equilibrium effort levels are disliked by co-workers and therefore bring disutility to the deviator. It is plausible that certain personality types

[22] Employrs select employees at least partly on the basis of psychological predictors of job performance. These predictors vary by occupation e.g. extraversion is valuable in salesmen whereas entrepreneurs score highly on conscientiousness and are less neurotic (Zhao et al., 2006).

[23] This result is not driven by the lower productivity of agreeable subjects or the skewness of the productivity distribution. The distribution of output-session norm is symmetrical and the mean positive deviation (1.99) is not significantly different from the mean negative deviation (1.83).

are more sensitive to disapproval and as a consequence are more likely to conform. These results are consistent with the conformity (measured as socially desirable responding) study by DeYoung et al. (2002) who find positive effects of agreeableness and conscientiousness (i.e. a negative coefficient in this setting) and negative effects of neuroticsm, extraversion and openness (i.e. a positive coefficient in this setting).

In columns (3) and (4) we include output and it appears that highly productive subjects are less susceptible to norm effects, as conjectured by Guryan et al. (2007). However, this finding may reflect the fact that it is more difficult to match co-workers' speeds accurately at higher output levels. We therefore check whether this result applies when relative deviation from the norm i.e. absolute value of (output-partner's output)/partner's output as a dependent variable, for the pairs in column (5), and absolute value of (output minus others' mean)/others' mean, for the entire sample in column (6). Although output is still positive it is no longer significant i.e. highly productive subjects are not significantly more inclined to deviate from norms.

The results in this section leave little room for doubt that subjects take cues from each other on how hard to work. As they stand, these results do not discriminate between the conformism and inequity aversion and competitiveness hypotheses. The next section analyses the post-task questionnaire which favour the competitiveness hypothesis.

5. ENJOYMENT RESULTS

'*The secret of joy in work is contained in one word - excellence. To know how to do something well is to enjoy it*.' [Pearl Buck]

The previous section presented evidence on the existence of peer effects. Now a psychological mechanism which can generate these peer effects is examined. People often enjoy what they are good at or what they think they are good at. [24] They may think they are good because they are better than their reference group. This section we provides evidence which is consistent with this hypothesis.

After completion of the experimental task, subjects were given a short questionnaire which included questions (on a scale of 1-5) about their stress level during the task (1= not stressed at all, 5=very stressed) and how much they enjoyed the task (1= did not enjoy it at all, 5=enjoyed it very much).

Whereas the responses to the stress question were largely correlated with personality factors, with conscientious individuals reporting less stress and neurotic more, the responses to the enjoyment question are rather revealing as they strongly suggest that subjects cared about their relative (to other group members) performance. Table 7 presents ordered probit results relating enjoyment to age, stress level and measures of relative (within session) performance. Gender and personality effects (and interaction of these with relative performance) were checked but these were not significant. Monetary incentive treatments, whether subjects were in pairs or group size similarly did not have an effect on enjoyment.

Older subjects appear to have enjoyed candle making most. Not surprisingly, those who experienced stress had lower enjoyment levels. Interestingly, the absolute level of output had no significant effect on enjoyment (see column (1) in Table 7). One might have speculated that, for intrinsic motivation reasons, subjects who enjoyed the

[24] "Golf and sex are about the only things you can enjoy without being good at." [Robin Green and Mitchell Burgess]

task would be more productive (and this would present a reverse causality problem) but there is no evidence for this. Enjoyment does though seem highly correlated with how well a subject performed relative to other subjects in the same session. Controlling for output, the ratio of subject's output to the average output of other subjects in his session is highly significant (see column (2) in Table 7). Subjects also reported significantly lower (higher) enjoyment levels if they were the least (most) productive subject in their session (see columns (3) and (4) in Table 7.)

In Table 8 reports this exercise for subjects in the pair treatment only. Age, gender and pay are not significant. As can be seen in column (3), relative output is again significant. This result allows us to conclude that what drives peer effects is indeed concern about relative performance rather than concern about relative financial payoff since both subjects get the same pay in the pair treatments.

For the paired subjects there is an additional variable which may influence enjoyment, namely whether subjects know each other. This variable turns out to be highly significant; subjects reported greater enjoyment of the task if they were paired with someone they knew. As in Table 7, individual output does not have a significant effect on enjoyment (see column (1) in Table 8) and the total pair output has a significant negative coefficient (see column (2) in Table 8). However, a subject's output relative to his partner's appears to affect enjoyment. So, it seems that even in the pair treatment where subjects were paid based on combined output, they were motivated by a desire to beat their partner.

Columns (5) and (6) in Table 8 include interaction variables and illustrate that the effect of relative performance on enjoyment is enhanced when subjects know their

partner. It appears to be more fun to outperform a co-worker you know. [25] Or possibly it is more depressing to be outperformed by someone you know. This suggests that work norms created by the ambition to do at least as well as your colleagues may be important in the workplace as workers usually know their co-workers.

Since subjective self-reported measurements on enjoyment are used, a caveat may be in order. To avoid cognitive dissonance, subjects may report attitudes consistent with their past behaviour: if they worked hard to outperform others they may conclude that they must have enjoyed the task or they would not have worked so hard. In addition, it is possible that subjects who feel driven to work harder than their peers have a tendency to over-report enjoyment of the task, perhaps in an attempt to please the experimenter. Future experiments could tackle this issue by, possibly in addition to gathering subjective measurements on enjoyment, offering subjects the opportunity to repeat the task. If those with high relative performance are more inclined to sign up for another round this would provide supporting evidence that enjoyment increases with relative performance.

The results in this section strongly suggest that subjects care about their relative performance. This casts doubt on the validity of the conformism hypothesis – if workers are motivated to comply with productivity norms they should be happiest when their output level is close to the norm. Subjects are happiest when they outperform their peers.

The results also suggest that inequity aversion is not likely to be a major motivating factor in the generation of peer effects. This is tested explicitly by constructing a measure of inequity and investigating its relationship with enjoyment. For the whole

[25] This is consistent with the literature on interpersonal comparisons which suggests that the effect on utility increases with proximity to the reference group (see e.g. Baron and Pfeffer, 1994).

sample inequity is defined for each session as the mean absolute deviation from session average output. For pairs, inequity is the output difference between the two subjects in the pair. As Table 9 shows, for the whole sample enjoyment actually *increases* in inequity and for the pair treatment (see Table 10) its effect is not significant.

6. CONCLUSIONS

This paper proposes a new underlying cause for peer effects: relative performance concerns. Workers may be motivated by the desire to outperform or at least keep up with their co-workers. This implies that in the presence of fast co-workers, individuals speed up whereas in the presence of relatively slow co-workers there is less pressure to work hard.

In the study, individual output does not appear sensitive to economic factors such as piece rates. Also, no evidence is found for free riding in the team treatments. However, psychology matters: less agreeable and more conscientious subjects are more productive. There is strong evidence for peer effects although there are no production interdependencies (as there would be in assembly lines for example). Adherence to work norms is at least partly determined by personality factors. To the extent that different work settings attract different personality types, it might be expected there would be mixed evidence on peer effects. The results indicate that agreeableness correlates with conformity to norms suggesting that peer effects may be stronger in caring professions such as teaching and nursing, where high agreeableness is overrepresented, and weaker in areas such as law and engineering where it is underrepresented (Caplan, 2003).

The finding that subjects' enjoyment of the task depends on their relative rather than absolute performance suggests that workers may be driven by pressure to outperform colleagues or at least to not be outperformed by colleagues. Whilst these results are certainly consistent with data on job satisfaction and relative pay, caution is appropriate in making predictions of behaviour in the work place. In the study, subjects report greater enjoyment of the task when their relative performance is higher but it is not clear whether in a repeated setting these findings would persist. Perhaps a temporary increase in satisfaction results from learning that you are more productive than your peers but this effect may be subject to habituation.

The hypothesis that norms arise because subjects are inequity averse is not supported. Job satisfaction does not appear to be negatively affected by inequality of outcomes. Further research is needed to explore whether inequity aversion is a likely cause of norms, perhaps using direct individual measures of inequity aversion (and competitiveness). One could argue that subjects could be motivated at least partly by the desire to conform but that they have insufficient control over their output to achieve conformity. However, in this setting whilst it may be difficult for slow workers to speed up it is eminently feasible for fast workers to slow down to the speed of their slower peers and they don't.

It must be conceded that the external validity of the findings is uncertain: the experiment involved small groups of subjects performing a menial, repetitive task for relatively small stakes and it is not clear the extent to which findings can be extrapolated to real workers in a real work place engaged in activities on which their livelihoods depend. However, within the context of the experiment we have illustrated that individuals respond to their co-workers' productivity levels. Work norms are important. Conformity to these norms appears to be particularly attractive to subjects

with an ‘agreeable’ personality. Overall, workers appear driven by competition and it is this concern about relative performance rather than inequity aversion which generates norms.

There are a number of interesting questions which deserve attention but which cannot be answered on the basis of the experimental data. Perhaps the most important one is whether heterogeneity of workers increases productivity i.e. in allocating workers to groups, should one aim for high or low heterogeneity? The experimental data show a positive correlation between heterogeneity in productivity and total output but it cannot be concluded from this that heterogeneity is desirable for at least two reasons. Firstly, ability is not directly measured; all the productivity data is in some sense contaminated by peer effects. Secondly, the sample productivity distribution is skewed which increases the likelihood of finding more variation in the higher output groups. Future research could tackle this question through an experimental design whereby individual worker’s ability can be measured and workers are rotated in different combinations to form groups. There is evidence from field data that more heterogeneous teams are more productive (see e.g. Hamilton et al., 2003). In Mas and Moretti (2009), the presence of high productivity workers benefits low productivity workers more than the detrimental effect the latter exert on the former so that a mix of workers which maximises diversity appears optimal.

The findings on the importance of relative performance concerns as a possible driver for peer effects is based on self-reports of task enjoyment. It would be desirable to design experiments in which the strength of relative performance as a motivator could be tested on the basis of subjects’ actions, perhaps in addition to subjective data. One could for example give subjects the opportunity to repeat the task and check whether

those who performed well with respect to their reference group are more likely to take up this offer.

Finally, it would be interesting to examine whether relative performance concerns remain important in a repeated setting with the same co-workers. Psychologists have known since Festinger (1954) that people adjust their performance to standards set by their social environment but it is not clear whether this effect only manifests itself in new or unknown tasks or whether it survives in familiar settings with familiar co-workers.

Acknowledgements

We thank David de Meza, Bernd Irlenbusch, Satoshi Kanazawa, Jordi Blanes i Vidal and Daniel Ferreira for helpful suggestions and Will Heron for assistance in the administration of the experiment.

REFERENCES

Antonovics, K., P. Arcidiacono and R.P. Walsh (2003) Competing against the opposite sex. *UCSD Economics Working Paper*, 2003-08.

Asch, S.E. (1951) Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (ed.) *Groups, leadership and men*. Pittsburgh, Carnegie Press, 177-190.

Bandiera, O., I Barankay, and I. Rasul (2006) The Evolution of cooperative norms: Evidence from a natural field experiment," *B.E. Journal of Economic Analysis and Policy* 6(2)

Bandiera, O., I. Barankay and I. Rasul (2010) Social incentives in the workplace. *Review of Economic Studies*, 77, 2, 417-458.

Baron, J.N. and J. Pfeffer (1994) The social psychology of organizations and inequality. *Social Psychology Quarterly*, 57, 3, 190-209.

Barrick, M.R. and M.K. Mount (1991) The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44, 1, 1-26.

Bellemare, C.; P. Lepage and B. Shearer (2010) Peer pressure, incentives, and gender: an experimental analysis of motivation in the workplace. *Labour Economics*, 17, 1, 276-83.

Bernheim, D. (1994) A theory of conformity. *Journal of Political Economy*, 102, 841-77.

Blanes I Vidal, J. and M. Nossol (2011) Tournaments without prizes: evidence from personnel records. *Management Science*, 57, 10, 1721-1736.

Brown, G., Gardner, J., Oswald, A.J. and Qian, J. (2008). Does wage rank affect employees' wellbeing? *Industrial Relations*, 47, 3, 355-389.

Caplan, B. (2003) Stigler-Becker versus Myers-Briggs: why preference-based explanations are scientifically meaningful and empirically important. *Journal of Economic Behavior and Organization*, 50, 4, 391- 405.

Charness, G., D. Masclet and M. C. Villeval (2010) Competitive preferences and status as an incentive: experimental evidence. *IZA Discussion Papers* 5034.

Clark, A.E. (1996). L'utilité est-elle relative? Analyse à l'aide de données sur les ménages. *Economie et Prévision*, vol.121, pp. 151-164.

Clark, A.E. and Oswald, A.J. (1996). Satisfaction and comparison income. *Journal of Public Economics,* vol. 61, pp. 359-81.

Chew, B., K. Clark and T. Bresnahan (1990), 'Measurement, coordination and learning in a multiplant network,' Chapter 5 in Robert Kaplan (ed.), *Measures for Manufacturing Excellence.* Boston, MA: Harvard Business School Press.

DeYoung, C.G., J.B. Peterson and D.M. Higgins (2002) Higher-order factors of the Big Five predict conformity: are these neuroses of health? *Personality and Individual Differences*, 33, 4, 533-52.

Dohmen, T.A. , A. Falk, K. Fliessbach, U. Sunde and B. Weber (2011) Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics*, 95, 3-4, 279-285.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In David, R. and R. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press.

Eriksson, T., A. Poulson and M.-C. Villeval (2009) Feedback and incentives: experimental evidence. *Labour Economics*, 16, 6, 679-688.

Falk, A., U. Fischbacher and S. Gachter (2002) Isolating social interaction effects. *Mimeo.* University of Zurich.

Falk, A. and A. Ichino (2006) Clean evidence on peer effects. *Journal of Labor Economics*, 24, 1, 39-57.

Festinger, L. (1954) A theory of social comparison processes. *Human Relations*, 7, 117-140.

Gneezy, U., M. Niederle and A. Rustichini (2003) Performance in competitive environments: gender differences. *Quarterly Journal of Economics*, 118, 3, 1049-74.

Graham, C. and Felton, A. (2006). Inequality and happiness: Insights from Latin America. *Journal of Economic Inequality*, 4, 107-122.

Guryan, J., K. Kroft and M. Notowidigdo (2009) Peer effects in the workplace: evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics,* 1, 4, 34-68.

Hamilton, B., J. Nickerson and H. Owan (2003) Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111, 3, 465-497.

Helliwell, J.F. and H. Huang, (2009). How's the job? Well-being and social capital in the workplace. *Industrial and Labor Relations Review*, 63, 2, 2.

Heineck, G. and S. Anger (2008) The returns to cognitive abilities and personality in Germany. Conference presentation ftp://ftp.zew.de/pub/zew-docs/veranstaltungen/NCS_Konferenz/presentations/AngerHeineck.pdf

Hertel, G; N.L. Kerr, M. Scheffler, S. Geister and L.A. Messe (2000) Instrumentality effects on motivation gains in groups: The role of impression management and spontaneous goal setting in promoting the Kohler effect. *Zeitschrift fur Sozialpsychologie*, 31, 204-220.

Hertel, G. ; G. Niemeyer and A. Clauss (2008) Social indispensibility or social comparison: the why and when of motivation gains of inferior group members. *Journal of Applied Social Psychology*, 38, 5, 1329-1363.

Huck, S., D. Kubler and J. Weibull (2012) Social norms and economic incentives in firms. *Journal of Economic Behavior and Organization*, 83, 2, 173-185 .

Ichino, A. and G. Maggi (2000) Work environment and individual background: explaining regional shirking differences in a large Italian firm. *Quarterly Journal of Economics*, 115, 1057-1090.

Judge, T.A. and J.E. Bono (2001) Relationship of core self-evaluations traits – Self-esteem, generalised self-efficacy, locus of control, and emotional stability – with job satisfaction and job performance: a meta-analysis. *Journal of Applied Psychology*, 86, 1, 80-92.

Kandel, E. and E. Lazear (1992) Peer pressure and partnerships. *Journal of Political Economy*, 100, 4, 801-17.

Karau, S.J. and K.D. Williams (1993) Social loafing: a meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65, 4, 681-706.

Kerr, N.L. and G. Hertel (2011) The Kohler group motivation gain: how to motivate the 'weak links' in a group. *Social and Personality Psychology Compass*, 5, 1, 43-55.

Kerr, N.L.; L.A. Messe, D.H. Seok, E. J. Sambolec, R.B. Lount Jr. and E.S. Park (2007) Psychological mechanisms underlying the Kohler motivation gain effect. *Personality and Social Psychology Bulletin*, 33, 828-841.

Knez, M. and D. Simester (2001) Firm-wide incentives and mutual monitoring at Continental Airlines. *Journal of Labor Economics*, 19, 4, 743-73.

Knight, J. and Song, L. (2009). Subjective well-being and its determinants in rural China. *China Economic Review*, 20, 4, 635-49.

Kohler, O. (1926) Kraftleistungen bei Einzel- und Gruppenarbeit [Physical performance in individual and group situations]. *Industrielle Psychotechnik*, 3, 274-282.

Kohler, O. (1927) Uber den Gruppenwirkungsgrad der menschlichen Korperarbeit und die Bedingung optimaler Kollectivkraftreaktion [On group efficiency of physical labor and the conditions of optimal collective performance]. *Industrielle Psychotechnik*, 4, 209-226.

Kuhnen, C.M. and A. Tymula (2012) Feedback, self-esteem, and performance in organizations. *Management Science*, 58, 1, 94-113.

Lavy, V. (2013) Gender differences in market competitiveness in a real workplace: evidence from performance-based pay tournaments among teachers. *Economic Journal,* 123, 540-573.

Lévy-Garboua, L. and Montmarquette, C. (2004). Reported job satisfaction: What does it mean? *Journal of Socio-Economics,* 33, 135-151.

Lount, R.B. Jr.; E.S. Park, N.L. Kerr, L.A. Messe and D.H. Seok (2008) Evaluation concerns and the Kohler effect: the impact of physical presence on motivation gains. *Small Group Research*, 39, 6, 795-812.

Luttmer, E. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics,* vol. 120, pp. 963-1002.

Mairesse, J. and Z. Griliches (1990), 'Heterogeneity in panel data: are there stable production functions?' in P. Champsaur et. al. (eds.), *Essays in Honor of Edmond Malinvaud, Volume 3: Empirical Economics*. MIT Press: Cambridge, MA.

Manski, C. (1993) Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, 60, 531-542.

Manski, C. F. (2000) Economic analysis of interactions. *Journal of Economic Perspectives*, 14, 3, 115-136.

Mas, A. and E. Moretti (2009) Peers at work. *American Economic Review*, 99, 1, 112-145.

McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. *Journal of Economic Behavior and Organization,* 45, 251-278.

McGahan, A. (1999), 'The performance of U.S. corporations: 1981–1994,' *Journal of Industrial Economics*, XLVII, 373–398.

Mohnen, A.; K. Pokorny and D. Sliwka (2008) Transparency, inequity aversion, and the dynamics of peer pressure in teams: theory and evidence. *Journal of Labor Economics*, 26, 4, 693-720.

Mueller, G. and E. Plug (2006) Estimating the effects of personality on male and female earnings. *Industrial and Labor Relations Review*, 60,1, 3-22.

Niederle, M. and L. Vesterlund (2007) Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122, 3, 1067-1101.

Peeters, M. , H. van Tuijl, C. Rutte and I. Reymen (2006) Personality and team performance: a meta-analysis. *European Journal of Personality*, 20, 377-96.

Rammstedt, B. and O. John (2007) Measuring personality in one minute or less: a 10 item short version of the Big Five Inventory in English and German, *Journal of Research in Personality*, 41, 203-12.

Salgado, J. F. (1997) The five factor model of personality and job performance in the European
Community. *Journal of Applied Psychology*, 82, 1, 30-43.

Sloane, P.J. and Williams, H. (2000). Job Satisfaction, comparison earnings, and gender. *Labour*, 14, 473-501.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*, 86(2): 534–50.

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

Weber, B. and G. Hertel (2007) Motivation gains of inferior group members: A meta-analytic review. *Journal of Personality and Social Psychology*, 93, 973-993.

Zajonc, R.B. (1965) Social facilitation. *Science*, 149, 269-274.

Zhao, H. and S.E. Scott (2006) The Big Five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, 91, 2, 259-271.

Figure 1: Candle Making Equipment

	I see myself as someone who…	Strongly Agree	Mostly Agree	Neither Agree nor Disagree	Mostly Disagree	Strongly Disagree
E	is reserved	1	2	3	4	5
A	is generally trusting	5	4	3	2	1
C	tends to be lazy	1	2	3	4	5
N	is relaxed, handles stress well	1	2	3	4	5
O	has few artistic talents	1	2	3	4	5
E	is outgoing, sociable	5	4	3	2	1
A	tends to find fault with others	1	2	3	4	5
C	does a thorough job	5	4	3	2	1
N	gets nervous easily	5	4	3	2	1
O	has an active imagination	5	4	3	2	1

Table 1: Personality inventory scoring

Output	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Frequency	2	10	17	21	10	19	6	9	4	2	3	4	3	4

Table 2: Output distribution

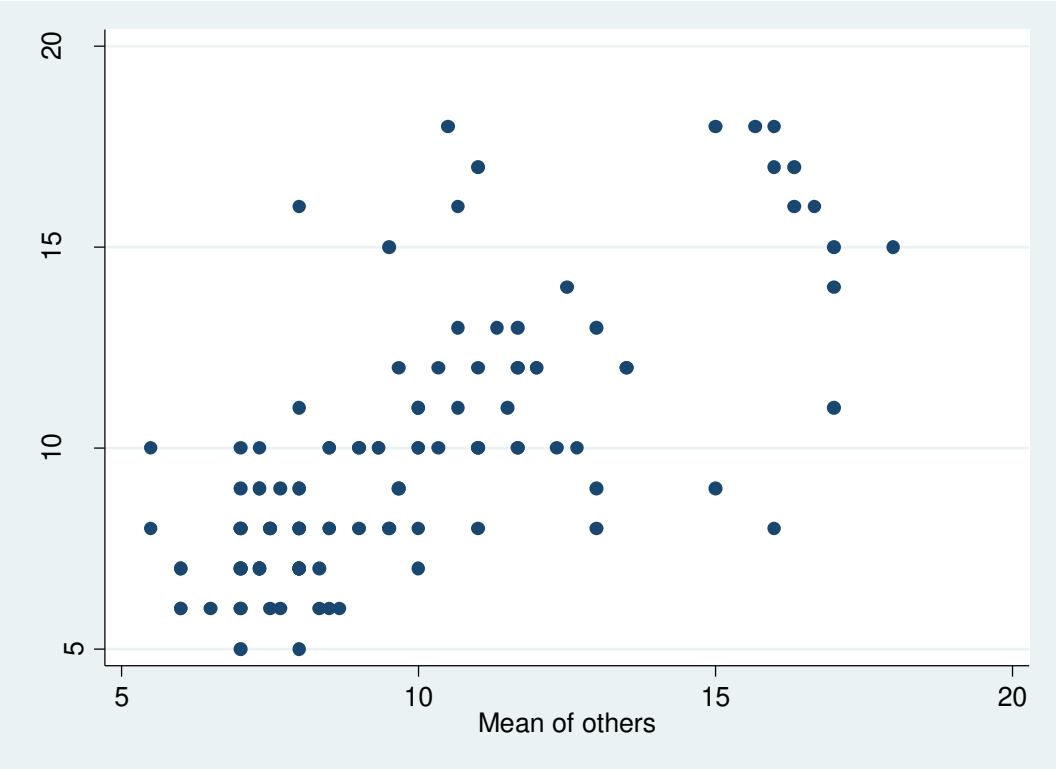

Figure 2: Individual output versus mean output of others in same session

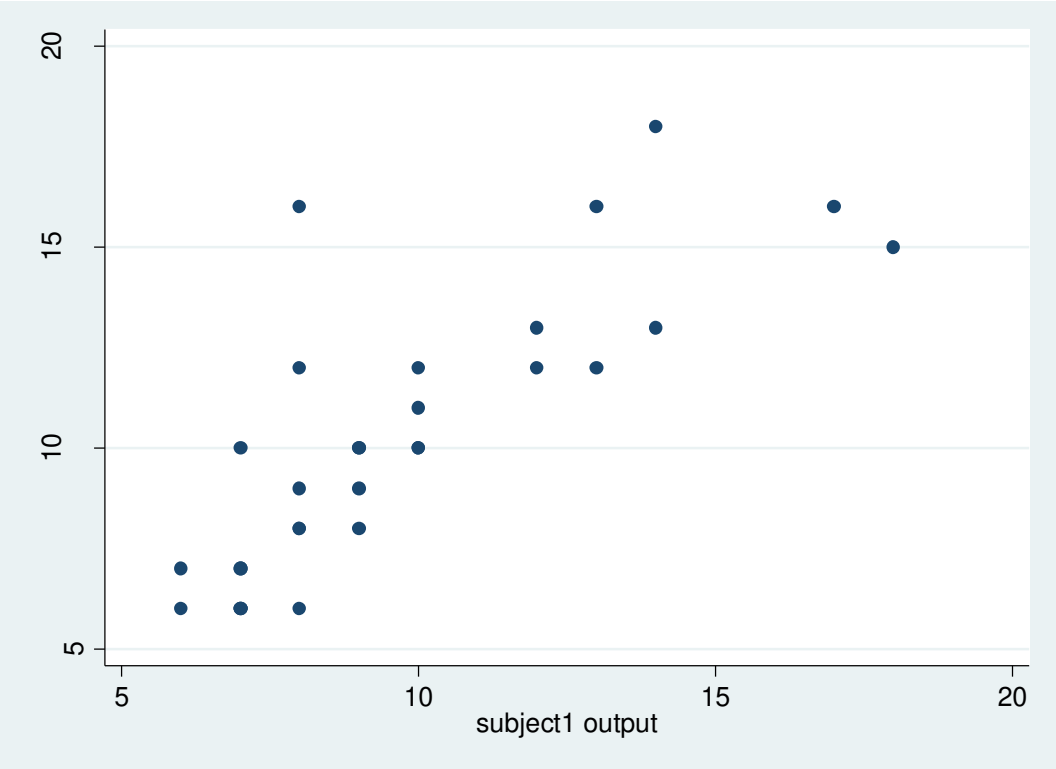

Figure 3: Outputs in pair treatment

Dependent variable: Output	(1)	(2)	(3)	(4)	(5)
Constant	2.302*** (2.73)	2.907*** (2.63)	2.388*** (2.76)	2.467*** (2.55)	2.495*** (2.74)
Othersmean	0.767*** (8.80)	0.738*** (7.38)	0.766*** (8.73)	0.754*** (8.57)	0.748*** (7.98)
Indat1		-0.500 (-0.96)			
Indat50		-0.337 (-1.15)			
Pairat1		-0.702* (-1.95)			
Individual			-0.170 (-0.72)		
Group size = 3				-0.263 (-0.84)	
Group size = 4				0.107 (0.34)	
Extraversion					0.220 (1.01)
Agreeableness					-0.456** (-2.07)
Conscientiousness					0.357** (2.14)
Neuroticism					-0.362 (-1.25)
Openness					-0.191 (-0.65)
R-squared	0.506	0.512	0.506	0.508	0.555
Prob>F	0.000	0.000	0.000	0.000	0.000

All OLS regressions on 114 observations robust standard errors clustered by session.
T-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 3: Output regressions

	List(L)	List(H)
Number of subjects	73	68
Mean output	8.6	11.2
Standard deviation	2.5	3.3
Proportion L	71%	28%
Proportion H	29%	72%

Table 4: List test data

Session size	# sessions	# sessions small range	P(small range)	Binomial p-value
2	11	8	0.47	0.079
3	13	10	0.40	0.008
4	13	9	0.28	0.002

Table 5: Range test data

Dependent variable:	Deviation from norm				Relative deviation from norm	
	(1) Pairs	(2) Entire sample	(3) Pairs	(4) Entire Sample	(5) Pairs	(6) Entire Sample
Extraversion	-0.008 (-0.04)	0.022 (0.15)	-0.109 (-0.55)	-0.039 (-0.24)	-0.002 (-0.10)	0.002 (0.16)
Agreeableness	-0.699** (-2.20)	-0.475* (-1.89)	-0.552* (-1.89)	-0.378* (-1.69)	-0.058* (-1.95)	-0.039 (-1.65)
Conscientiousness	-0.216 (-0.71)	-0.036 (-0.21)	-0.221 (-0.73)	-0.065 (-0.37)	-0.015 (-0.57)	0.003 (0.23)
Neuroticism	-0.012 (-0.09)	-0.044 (-0.28)	0.122 (0.92)	0.046 (0.26)	0.008 (0.70)	0.008 (0.39)
Openness	0.282 (1.09)	0.046 (0.32)	0.273 (1.19)	0.093 (0.69)	0.050 (1.59)	0.020 (0.97)
Output			0.170*** (2.76)	0.159** (2.15)	0.008 (1.22)	0.009 (1.12)
Constant	1.4*** (5.11)	1.649*** (7.53)	-0.305 (-0.48)	0.075 (0.11)	0.056 (0.83)	0.075 (0.99)
R-squared	0.206	0.078	0.300	0.161	0.259	0.107
Prob>F	0.150	0.525	0.002	0.099	0.169	0.453

OLS regressions with robust standard errors clustered by pair in (1), (3) and (5) and by session in (2). (4) and (6).
T-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 6: Personality and deviation from norm

Dependent variable: Enjoyment	(1)	(2)	(3)	(4)
Age	0.046** (2.19)	0.050** (2.34)	0.047** (2.27)	0.050** (2.40)
Stress	-0.215** (-2.08)	-0.227** (-2.16)	-0.233** (-2.17)	-0.212** (-2.02)
Output	0.020 (0.53)	-0.019 (-0.46)	-0.004 (-0.009)	0.012 (0.32)
Output/Othersmean		1.332*** (2.71)		
Lowest output			-0.400** (-2.09)	
Highest output				0.399** (2.20)
Pseudo R^2	0.035	0.056	0.043	0.045
Prob>χ^2	0.014	0.002	0.010	0.005

All ordered probits with robust standard errors clustered by session.
z-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 7: Ordered probit results for Enjoyment

Dependent variable: Enjoyment	(1)	(2)	(3)	(4)	(5)	(6)
Output	-0.078 (-1.60)		-0.109** (-2.10)	-0.082 (-1.62)	-0.112** (-2.15)	-0.081 (-1.61)
Pair total output		-0.056** (-2.16)				
Output/ partner's output			1.825** (2.04)		2.605** (1.98)	
Output<partner's				-0.556* (-1.78)		-0.644* (-1.80)
Don't know partner	-1.063*** (-2.62)	-1.121*** (-2.85)	-1.065*** (-2.62)	-1.000** (-2.40)	1.780 (1.11)	-1.149** (-2.14)
Don't know partner* Output/ partner's output					-2.913* (-1.82)	
Don't know partner* (Output<partner's)						0.326 (0.47)
Stress	-0.357** (-2.36)	-0.359** (-2.51)	-0.290* (-1.79)	-0.328** (-2.22)	-0.364** (-2.12)	-0.353** (-2.20)
Pseudo R^2	0.091	0.109	0.128	0.115	0.144	0.116
Prob>χ^2	0.044	0.011	0.002	0.027	0.013	0.048

All ordered probits with robust standard errors clustered by pair.
z-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 8: Ordered probit results for Enjoyment – Pair treatment only

Dependent variable: Enjoyment	(1)	(2)	(3)	(4)
Age		0.053** (2.40)	0.056** (2.52)	0.049** (2.24)
Stress		-0.244** (-2.26)	-0.226** (-2.09)	-0.238** (-2.22)
Output		-0.049 (-1.05)	-0.021 (-0.49)	-0.022 (-0.54)
Output/Othersmean		1.576*** (2.90)		
Lowest output				-0.368* (-1.90)
Highest output			0.567*** (3.05)	
Inequity	0.146* (1.68)	0.191* (1.95)	0.221** (2.25)	0.145 (1.49)
Pseudo R^2	0.014	0.001	0.001	0.010
Prob>χ^2	0.092	0.076	0.070	0.055

All ordered probits with robust standard errors clustered by session.
z-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 9: Ordered probit results for Enjoyment and Inequity

Dependent variable: Enjoyment	(1)	(2)	(3)
Output		-0.112** (-2.04)	-0.094* (-1.67)
Don't know partner		-1.068*** (-2.61)	-1.003** (-2.39)
Stress		-0.287* (-1.75)	-0.308* (-2.01)
Output/partner's output		1.850** (2.29)	
Output<partner's			-0.630* (-1.95)
Inequity	-0.016 (-0.12)	0.014 (0.11)	0.067 (0.56)
Pseudo R^2	0.000	0.123	0.118
Prob>χ^2	0.901	0.001	0.046

All ordered probits with robust standard errors clustered by pair.
z-values in brackets. * =10% sign., **=5%sign, ***=1% sign.

Table 10: Ordered probit results for Enjoyment and Inequity – Pair treatment only