



Munich Personal RePEc Archive

Reconciling economics and psychology on intrinsic motivation

Bruno, Bruna

Università di Salerno

November 2012

Online at <https://mpra.ub.uni-muenchen.de/42717/>

MPRA Paper No. 42717, posted 19 Nov 2012 11:30 UTC

Reconciling Economics and Psychology on Intrinsic Motivation

Bruna Bruno*

November 2012

Abstract

The paper analyzes how the debate on intrinsic motivation was imported from psychology into economics. The most important differences between the two disciplines are in the definition of intrinsic motivation and in the timing of the undermining effect of rewards. The economic framework of inter-temporal choices is proposed to reconcile the different empirical and theoretical results arising in the literature, and it is shown how rewards induce substitution and income effects depending on whether they are transitory or permanent. Furthermore, a distinction between input and output oriented intrinsic motivation is introduced.

Keywords: intrinsic motivation, rewards, crowding out, undermining.

*Department of Economics and Statistics - University of Salerno

Via Ponte don Melillo, 84084 Fisciano (SA) – Italy

brbruna@unisa.it

Introduction

In psychology, motivation is defined as a dynamic factor that directs behaviour toward an objective. According to Geen (1994), motivation refers to the initiation, direction, intensity and persistence of human behaviour. Extrinsic motivation operates when someone engages in a particular behaviour for purposes that are extrinsic to the behaviour itself, such as to receive praise, awards, good reviews or to avoid unpleasant situations or punishments. In Deci et al. (2008), behaviour motivated by extrinsic motivation 'entails doing an activity because it leads to some outcome that is operationally separable from the activity itself. That is, extrinsic motivation concerns activities enacted because they are instrumental rather than because one finds the actions satisfying in their own right' (p. 12). Intrinsic motivation, instead, operates when someone engages in behaviour because he finds the activity challenging and rewarding in itself, and derives satisfaction in enhancing his competence in that specific task.

Economics has devoted much attention to the role of incentives in economic relationships but, as early as 1971, Titmuss suggested that incentives for blood donors may have a countervailing effect. More recently, Frey (1992), importing from psychology the distinction between intrinsic and extrinsic motivation, raised new interest in this topic, stressing that economists must be cautious in prescribing economic incentives, because of the unexpected effects of extrinsic incentives on intrinsic motivation.

By contrast, Benabou and Tirole (2003) elaborated a model of informational signals to show that the perverse action of extrinsic rewards arises only in a specific setup of a principal-agent problem and that 'before worrying about the negative impact of rewards, one should first check that the reward provider has private information about the task or the agent's talent' (Benabou and Tirole, 2003, p.505). Their analysis is extended in Benabou and Tirole (2006) where the effects of social reputation are also considered.

The crucial point in economic debate is whether and when economic incentives induce the desired performance, with two opposing positions: one invoking a strong role for intrinsic motivation in describing some anomalous effects of incentives, the

other recalling the fundamental role of informational setup to explain why, in specific frameworks, the price signal sometimes shows unexpected effects.

Economic theorists apply their instruments of analysis to the psychological distinction between extrinsic and intrinsic motivations, but the two disciplines pursue different objectives and the reasoning does not completely overlap. The most important divergence between economic analyses and psychological research emerges when the effects of expected rewards on individual performance are discussed. While psychologists are interested in the 'evolution' of the behaviour of an individual engaged in a freely chosen activity, economists, following a consolidated tradition of comparative statics analysis, search for the 'simultaneous' effect on behaviour (and performance) of a reward supplied in a previously non monetary relationship. The issue becomes pivotal when empirical results are discussed: most results of the psychological research show undermining effects after the reward removal, whereas economics usually models the effects of rewards *during* the reward supply.

Psychology offers a large amount of empirical research on the effects of rewards on free-choice activities, well summarised in some meta-analytical studies (Deci et al., 1999; Rummel and Feinberg, 1988; Wiersma, 1992; Tang and Hall, 1995; Eisenberger and Cameron, 1996; Cameron and Pierce, 1994), while data in the economic field are more recent and less consolidated. Therefore, it is not infrequent for economic theorists to refer their working hypothesis to the results of psychological research, explaining dynamic results in a static framework. This ambiguity in the relation between economic theory and psychological results has two limiting implications. First, economics implicitly assumes that the effects of incentives on intrinsic motivation and performance do not change whether the incentives are temporary or permanent. Second, according to the two opposing positions, the price effect of an external incentive is reduced to its signalling power or to its income effect on individual behavior, whereas any intertemporal substitution effect is disregarded.

To reconcile economic theory with psychological research, the intertemporal substitution between the two periods of observation (during and post rewards) is considered. Substitution between two periods is relevant only for temporary policies (rewards), whereas for permanent policies only income effects can be predicted. The static economic analysis of the hidden costs of incentives implicitly assumes a permanent reward, whereas in psychology the rewards are often temporary. It can be

shown that economic models are interested in the income effect induced by a permanent reward, during its supply; on the other hand, psychology investigates the intertemporal substitution effect of a temporary reward in the post reward period. Therefore, similarities and differences among results proposed by economists and psychologists can be explained by adopting an intertemporal framework. Moreover, when economics searches for the perverse effects of incentives, it cannot invoke empirical results from psychology, usually related to a substitution effect. Dealing with income effects, economic analyses should concentrate on whether the free choice activity is considered by individuals to be an inferior good.

In the following section, some theoretical models on the role of intrinsic motivation in economic behaviour are discussed. Section 3 takes a deeper look into the question of the timing of the crowding out of intrinsic motivation, explaining how different results can be interpreted in terms of intertemporal choices. Section 4 analyses the different definitions of intrinsic motivations and how these definitions can be classified as input or output oriented motivations (Section 5). To test the robustness of the intertemporal choice approach to intrinsic motivation and crowding out, some empirical economic analyses are discussed in Section 6. Concluding remarks are in Section 7.

2. The economic models of intrinsic motivation

The economic literature has devoted some attention to the effect of intrinsic motivation on the performance of workers and students, to study the crowding out effect of pecuniary incentives (extrinsic motivations). Two effects of increasing rewards were distinguished: a relative price effect, which increases the supply of effort in the activity by lowering the opportunity cost of doing it; a crowding effect, which may increase the agent's performance (crowding in), or reduce the performance, if intervention undermines intrinsic motivation (crowding out) so much that the negative effect on performance is greater than the positive price effect (Frey and Goette, 1999). Both the relative price effect and the crowding effect are active, implying that the external intervention may be beneficial from the principal point of view depending on the relative size of the two countervailing effects. Frey and Jegen

(2001) formalised the crowding out effect in the Motivation crowding theory, but did not explain why derivatives representing crowding effects could differ from one person to another and from one situation to another (Harvey, 2005).

Many authors have stressed the relation between the crowding effect of rewards and the perceived control effect. Frey and Jegen (2001) identify two psychological processes that affect intrinsic motivations: impaired self determination and impaired self esteem. Benabou and Tirole (2003) use the ‘looking glass self’ (Cooley, 1902) to show that an agent takes the principal’s perspective in order to learn about himself. For the authors, the incentives enhance engagement in an activity only if they reveal hidden information to the agent, about the task or the agent’s talent, enhancing his confidence in himself. In the model of prosocial behaviour (Benabou and Tirole, 2006) rewards and punishments create doubt about the true motive for which good deeds are performed. In the signal-extraction problem, rewards crowd out reputational motivation, modifying the endogenous and unobservable mix of motivations (intrinsic, extrinsic and reputational) that is the source of pro social behaviour.

While Frey refers explicitly to the cognitive evaluation theory of Deci (1971) and to the fundamental role of intrinsic motivation in human behaviour, the models of Benabou and Tirole use the attributional approach to include the overjustification effect. Benabou and Tirole do not consider the role of intrinsic motivation as a crucial point, because they “show how the ‘overjustification effect’... can be understood as a signal-extraction problem in which *rewards amplify the noise*, leading observers... to attribute less of a role to intrinsic motivation in explaining variations in behavior” (Benabou and Tirole, 2006, p. 1660).

An asymmetric information framework also features in Sliwka model (2007), where the explanation of crowding out is explicitly ‘distinct from those proposed by psychologists’, and is based on the learning of the prevailing social norm, emerging from the incentive scheme proposed by the principal. In his model, asymmetric information is assumed about the type of agent and the distribution of types in the population, with selfish agents compared with fair agents, who care for the principal’s payoff, and conformist agents, who are alternatively selfish or fair, depending on how the majority of population behaves. Bolle and Otto (2010) assume a linear relation between individual utility and others’ utility. The intrinsic motivation towards the other’s welfare depends on the value of the good the other receives and not on the

good itself: the individual estimation of this value is assumed higher than the signal he extract from the reward and crowding out could occur if the signal is too low.

Most economic models are concerned with a backward shift of the standard supply curve, which predicts that effort increases with payments, or with a downward-sloping supply that totally reverses the sign of the relation between effort and payments. The coherence of these results with the psychological predictions on the effect of rewards will be discussed in the following section.

3. The timing of the crowding out

The control effect of rewards on behaviour is demonstrated by the fact that “when administered closely subsequent to a behaviour, rewards were reliably found to increase the likelihood that the behaviour would be emitted again, an effect that persisted as long as the reward contingency was operative. When rewards were terminated, the likelihood that the behavior would be emitted eventually returned to the prereward baseline” (Deci et al., 1999, p. 627). The undermining effect of rewards has a further implication when likelihoods following the reward removal are *below* the prereward baseline.

From the empirical point of view, results of empirical tests on intrinsic motivation generally refer to the degree to which participants return and persist at the target activity during a free-choice period subsequent to the experimental phase (Deci, 1971). The persistence can be measured by the amount of time spent in the free choice activity or by the number of trials or successes with the target activity. The undermining effect appears when a postreward response rate is below the response rate of a no-rewards comparison group (Deci et al., 1999, p. 658).

Most results from the psychological research show undermining effects after the reward removal, whereas during the reward supply a better performance can be found¹. The rationale for this is that, during the reward contingency, individual behaviour is the result of a mix of intrinsic and extrinsic motivation (Deci et al., 1999,

¹ Few exceptions can be found in psychological literature, as in Eisenstein (1985).

p.635). When importing into economics the evidence of the undermining effects of rewards on intrinsic motivation, this timing is less clear. The traditional attitude for comparative statics in economics led to modelling the effects of rewards *during* the reward supply. It is worth noting that the terminologies used in economics and psychology are also different: the “undermining effect” in psychology becomes the “crowding out effect” in economics, the latter being the conventional expression for countervailing and simultaneous effects. In what follows, the different meaning of these terminologies will be maintained.

In the words of Frey and Jegen (2001), an external reward, introduced in a previously non monetary relationship, produce two effects: the relative price effect occurs because external intervention (reward) raises performance by imposing a higher marginal cost of shirking, or increasing the marginal benefit of performing; and the crowding-out effect, which reduces the supply of effort by undermining the marginal utility deriving from the activity. As stated above, empirical psychological research stresses that, after the reward removal, a reduction in intrinsic motivation may be observed, under some specific circumstances. Both disciplines predict the same directions for each effect: intrinsic motivation enhances performance, rewards enhance performance, rewards (may) reduce intrinsic motivation, thereby reducing performance. The critical difference between the two disciplines is that psychology analyzes only the signs of the changes, measured after the reward removal, while economics is also interested in the size of the changes that occur during the reward supply. From this point of view, economic analysis focuses on a restrictive case of the undermining of intrinsic motivation. This occurs because economics is especially interested in the well functioning of economic incentives, and economic models try to predict under which circumstances the monetary incentive can be counterproductive.

3.1 Transitory and permanent rewards

A recent paper by Gneezy et al. (2011) addresses the problem of the timing of crowding out, distinguishing between crowding out when incentives are in place and crowding out after incentives are removed. To this they add the well-known distinction between short and long run, defining the first type of crowding out as a short run situation and the latter as a long run situation. However, these mixed

definitions can be misleading because it is not clear if the short/long run definition is an attribution of the reward or of the effect induced by the reward. To clarify the short/long run attribution of policies and of their effects, a separate discussion of undermining effects and crowding out effects will be conducted in the following paragraphs.

As to the undermining effect, the effect of reward is verified after the incentive is removed. Therefore, the reward is a transitory policy of the principal that can be labelled as short run policy. The undermining of intrinsic motivation is a permanent effect of this transitory policy that can be labelled as a long run effect, because it is in evidence after the reward removal but lasts from the period of reward supply until the period after reward removal onwards. The meta-analysis of Deci et al. (1999) supports this explanation when it explores explicitly the problem of the duration of the undermining effect of rewards, to answer the behaviourists' claim that undermining is merely a transitory effect. By comparing studies with different timings, the authors concluded that the timing of assessments after reward removal did not affect the results. This result confirms that undermining is a permanent effect after a temporary reward.

On the other hand, the crowding out measured when an incentive is in place has a static connotation without any relation to time. The only assessment that can be made is that the crowding out will appear while the reward is in place, other things being equal, whereas the post-reward effects are not investigated. From this point of view, both reward and its effect are considered as permanent, simply because the effects of temporary policies are disregarded. It should be noted that, based only on psychological data, economics should predict only that an incentive would be counterproductive if it is removed, signalling the inefficiency of temporary incentives.

3.2 Substitution and income effects

In economics, different behaviours are associated with temporary and permanent shocks, if they are expected. Facing a shock that is felt as temporary, agents are induced to an inter-temporal substitution: they try to benefit from the temporary shock during its contingency, knowing that in the following period the beneficial effect will end. This could be the case of a reward whose supply will end with the experiment. If

agents know (or perceive) that the reward is temporary, they will 'use' their motivation as much as possible during the reward supply to exploit the favourable period. They expect that in the future only intrinsic motivation will increase their utility, because the reward will disappear. The undermining effect can be interpreted as an intertemporal substitution effect: when facing a temporary reward, individuals choose more free choice activity when it is rewarded, reducing the future engagement in that activity. Compared with a no-reward situation, the undermining condition will exhibit a reduction in the free choice activity in the post reward removal period, whereas the unrewarded individual shows a constant level of activity and motivation. This explanation needs the assumption that agents can freely choose the intertemporal allocation of their intrinsic motivation, as occurs for consumption. Furthermore, it should be noted that the substitution effect is at work when the reward is expected or announced before or during the performance. This condition is supported by empirical psychological research showing that for unexpected rewards no detrimental effect is present.

When the reward is perceived or announced as permanent, agents will choose actions that incorporate a never-ending reward. Agents leave the distribution of the free choice activity unchanged among different periods at a constant level, because each period is equally rewarded. The intertemporal effect predicted is only an income effect (no substitution takes place): individuals facing a reward that will last for each period feel richer. The income effect is associated with an increase in the good consumption (the free choice activity) if the good is a normal good, whereas the free choice activity is reduced if it is an inferior good. If the motivated activity is an inferior good, then the effort in that activity will be reduced during the reward supply, showing a crowding out situation. If the motivated activity is a normal good, the crowding in will appear. Note that the classical distinction between inferior and normal goods in economics is not a quality statement of different goods but concerns their different affordability. In the scarcity approach to individual choices, some goods are inferior simply because they are easily affordable with a low income. This explanation therefore needs the assumption that the free choice activity is a source of individual utility easily affordable at a low-income level. When income increases, individuals can consume a wider variety of goods and the consumption of the (inferior and generally low price) good decreases.

If this distinction reflects the actual difference between economic and psychological results, the ‘undermining’ effect of post reward removal is produced by the ex ante perception of a temporary reward, which induces an intertemporal substitution in the choices related to the motivated activity. The ‘crowding-out’ effect is the consequence of a reward perceived as permanent, bringing an income effect that reduces the motivated activity if it is an inferior good. In a two periods framework, the ‘undermining’ is usually observed in the second period, whereas the claim for the existence of ‘crowding out’ concerns the first period. Note that, when the reward is temporary, in the first period both effects can be in place. With normal goods, the two effects would reinforce each other, enhancing the motivated activity both for a substitution and an income effect. When the motivated activity is classified as an inferior good, the substitution effect will increase performance whereas the income effect will decrease it. Therefore, the net effect in the first period is ambiguous.

4. Intrinsic motivations classification

Some further differences between economics and psychology are in their definition of intrinsic motivation. The distinction between intrinsic and extrinsic motivation is based on the relation between motivation and the activity performed and/or the individual. Starting from the idea that motivation is a factor that activates and directs human behaviour, one can easily classify economic incentives in the category of extrinsic motivations, because they are exogenous both to the activity and to the individual.

By comparison, any motivation that is endogenous to the individual and/or to his behaviour is intrinsic. When motivation is not provided by someone else, is formalised into a feeling and/or is strictly linked to the activity performed, it is intrinsic. The effect of the intrinsic motivation is easy to recognize when economic theory predicts a very different behaviour: volunteering with a zero wage is an example. Nevertheless, in many other cases it is difficult to ‘differentiate between different sources of motivation, which in the economic view are just manifestations of underlying preferences (for the task itself, or for the reward that is associated with performing the task)’ (Frey and Jegen, 2001, p.591), and between the two polar cases

of purely extrinsically and purely intrinsically induced individuals there is a continuum of combination of the two motivations.

To better explain how intrinsic and extrinsic motivations compete in behaviour orientation it is useful to consider the effect of more than one goal at a time, as suggested by Lindenberg (2001). The following classification of goals of human behaviour relevant for intrinsic motivation is from Meier and Stutzer (2008). Their original classification is extended in this paper with a fourth category (n. 2) to better include the Benabou and Tirole (2003) framework. These categories are considered relevant for economic studies: a brief discussion of the differences to psychological approaches will follow.

1. Pleasure in performing the task. Strictly following the Deci definition, ‘to be intrinsically motivated means to engage in an activity because the activity itself is interesting and enjoyable’ (Deci et al. 2008, p. 11). Intrinsic motivation directs behaviour towards the pleasure of doing something and the opportunity to let one’s own competence increase. Being linked to individual preferences and to the activity, motivation is intrinsic both to the individual and to the activity. No one else is involved.
2. Desire to succeed in performing the task. From the Benabou and Tirole perspective, an ‘agent will undertake the task only if he has sufficient confidence in his own ability to succeed’ (Benabou and Tirole, 2003, p. 491). Intrinsic motivation is identified with the probability of success, which in turn depends on the agent’s self-confidence, his ability and the difficulty of the task. It is intrinsic both to the individual and to the activity. Through the ‘looking glass self’, the probability of success could be reinforced or crowded out by a principal that, supplying an extrinsic motivation, provides information on personal ability and/or the difficulty of the task. The extrinsic motivation may be an explicit reward or a public acknowledgement.
3. The warm glow. An impure form of altruism is what Andreoni (1990) defined the warm glow, to point out that people are often ‘motivated by a desire to win prestige, respect, friendship, and other social and psychological objectives’ (Olson, 1965). In this respect, motivation is a feeling that needs an ex post social approval to be reinforced. Though warm glow is provided by the

surrounding society, it is intrinsic to the individual and could be intrinsic to the activity.

4. Social preferences. Social preferences imply that an individual has as an objective not only his welfare but also other people's welfare (Fehr and Fischbacher, 2002) and can be interpreted as a category of intrinsic motivation: individual choices are directed by an internal emotional goal that is related to the welfare of others. Motivation is intrinsic to the individual, for it is embedded in his preferences. It could be not intrinsic to the activity.

Moving from category 1 to 4, one can observe an increasing social involvement of the individual, from pure individual gratification, to the need to be accepted at a social level, to the other regarding preferences.

Some criticism can be expressed towards the category related to the "warm glow", included by Meier and Stutzer as a kind of intrinsic motivation. Psychological research would classify this motivation as extrinsic because it involves "taking a regulation but not fully accepting it as one's own" (Deci, Ryan 2000 p. 72). These introjected behaviours, in the words of Deci and Ryan, "although internally driven ... are not really experienced as a part of the self" because they are directed to win prestige or friendship in the external environment and so would be better classified as extrinsically motivated.

This criticism is relevant to an analysis of the model in Benabou and Tirole (2006), where the authors investigate pro social behaviour through a preference function including intrinsic, extrinsic and reputational motivations. The authors describe intrinsic motivation to behave pro-socially as stemming from two sources. The first is the *pure* form of altruism of an agent who "may care about the overall level of a public good to which his action contributes, such as air quality". This part of intrinsic motivation can reflect a genuine interest in social welfare, but it is later also defined as "the marginal utility of a public good generated by total contribution" (p. 1670), which is the selfish agent's marginal utility deriving from the total public good supply. The second is a 'joy of giving' (independent of social – or self esteem concerns) that makes him value his own contribution more than someone else's, which they call *impure* altruism (p.1657). If we exclude the *impure* altruism from the

category of intrinsic motivation, the second component cannot be labelled as intrinsic motivation.

The relevant point is that, in the model offered by Benabou and Tirole, the first term becomes vanishingly small in large groups because it depends inversely on the size of the group. Consequently, the only intrinsic motivation that is eventually crowded out in the model is the “warm glow”, that is the desire to gain social approval.

5. Motivations and rewards: an input and output classification

Starting from the above categories, a new distinction among goals that are relevant for intrinsic motivation can be proposed, which could be especially useful for economic modelling, because it uses the classical distinction between input and output concepts. Intrinsic motivation is *input oriented* if it is not directed toward the output dimension of the activity. On the other hand, when the goal of intrinsic motivation is the output of the activity, the intrinsic motivation is *output oriented*. It is easy to show that in the first category the intrinsic motivation is not related to the output dimension of the activity, and the intrinsic motivation is input oriented. In the last three categories, the objective of intrinsic motivation is the output of the activity, rather than the activity itself, and the intrinsic motivation is output oriented.

The distinction is also relevant because of the different role of self esteem and self determination mechanisms, which are referred to as relevant mechanisms in cognitive evaluation theory. Self determination is related to autonomy and competence of individual effort and affects the input dimension of intrinsic motivation. Self esteem arises from the comparison between results (real self) and expectations (ideal self) and affects the output dimension of intrinsic motivation. It plays a role only when the intrinsic motivation is output oriented, and is not relevant when an activity is engaged just for the pleasure of performing the activity itself and the agent doesn't care about the results of his performance. Note that social psychologists and sociologists describe the two sources of intrinsic motivation more or less in a similar way. Galbraith (1977) and Staw (1989) consider intrinsic motivation as deriving from task involvement or goal identification.

The input/output classification, implying traditional economic terminology, can be more easily reconciled with the psychological analysis and its classification of reward types. The coding proposed and reported in Deci et al (1999) can be summarised as follows.

1. Task – noncontingent rewards do not require the person to engage in the target activity (e.g. rewards contingent on waiting or for being in the study);
2. Engagement – contingent rewards require that some effort in a specific activity must be supplied in order to obtain the reward.
3. Completion – contingent rewards are offered only when the task is completed.
4. Performance – contingent rewards depend on the performance level the person exhibits in completing the task.

The argument proposed in this paper to link the instruments for economic modelling to the psychological analysis is the following. The first two categories of rewards are linked to the input dimension of the participant, whereas the latter two categories are linked to the output dimension of the activity: the first with a discrete evaluation (yes or no), the second with a continuous measure (the quantity/quality level of performance).

In economic modelling, it can be useful to compare the orientation of the motivation (input/output) described by a specific preference structure with the incentive policy designed by the principal. Different situations and equilibrium values emerge from each combination of preference and incentives. Simplifying, four different situations can arise from the combination of motivations and rewards:

1. input oriented motivation and input related reward
2. input oriented motivation and output related reward
3. output oriented motivation and input related reward
4. output oriented motivation and output related reward

Intuitively, the mixed combinations with input (output) objectives and output (input) measures are not relevant, whereas the direct measure of the objective of motivation in the input/input and output/output combinations are candidates for undesired effects. Unfortunately, this intuition cannot be easily tested because it is difficult to discern if

individuals are driven by input or output oriented motivations. Moreover, psychological literature finds undermining effects not only for output related rewards but also for the engagement-contingent rewards, an input related type of reward.

6. Evidence in economics on crowding out

To verify if economic analysis is based on evidence dealing with permanent rewards and income effects, a brief discussion of empirical tests in the economic field will be useful. The quantity of economic evidence on the crowding out is not comparable to the great bulk of empirical research that psychology has accumulated in the last thirty years. Some experimental and empirical studies support the crowding out of intrinsic motivation, through the observation of a reduction in performance induced by the introduction of a reward. All these studies measure the performance during the reward supply.

Note that the simplest way to verify that a good is an inferior good is its relation with income: when income increases the consumption of an inferior good will decrease.

Therefore, data arising from economic studies can be discussed in two different ways:

- a) when incentives are permanent, no undermining (substitution) effect takes place; if the objective of motivation is an inferior good, crowding out (income effect) can appear. Testing crowding out requires data on the same individual experiencing different levels of income (or reward);
- b) when incentives are temporary, an undermining (substitution effect) emerges in individuals observed before and after reward. A crowding out (income effect) can be tested when comparing individuals with different levels of rewards in the same period.

A first set of experiments on motivation crowding out is in Gneezy and Rustichini (2000a). Comparing their analysis with the standard psychological studies, they state the difference between the two as follows: “We study the behavioural response to different rewards in a single - stage setup. The comparison is across individuals, not across successive periods for the same individual following the reward” (p. 795). It is worth noting that the meta-analysis by Deci et al. (1999) considers only the studies

that had appropriate no-reward control groups. The comparisons exerted by psychological research are across individuals analysed across successive periods, whereas the experiments by Gneezy and Rustichini compare different individuals in one period. The psychological studies analyse the difference in behaviour between a reward group and a no-reward (control group), after reward removal, whereas the Gneezy and Rustichini experiments analyse the difference in behaviour between a reward group and a no-reward (control group), during the reward supply. This different set up is coherent with the different objective of analysis.

The first experiment reported in Gneezy and Rustichini (2000a) involved a group of students paid to perform an IQ test. All groups of students were paid for participating in the test and three of the four groups had additional payments – with different rewards – for questions they answered correctly. Rewards are all announced before starting the test for each group. Following the rewards classification above, the control group experienced an engagement-contingent reward, whereas the others an additional performance-contingent reward. Both kinds of reward were found to be undermining in the psychological literature (Deci et al., 1999), while in the Gneezy and Rustichini experiment the crowding out emerges only in the group with a very high performance pay, whereas the low paid group increased their performance, compared with the control group rewarded for the engagement.

It is difficult to compare these results with the results from psychology because the control group would also be classified as a rewarded group by the psychological literature: it is a group that faces an engagement-contingent reward. Without a control group, it could not be argued that incentives undermine intrinsic motivation but only that different kind of incentive have different effects on performance. Moreover, in the light of the previous explanations, the rewards for the IQ test are transitory and can induce a substitution effect, but we cannot observe the behaviour after the reward removal. The crowding out of the high pay group supports the idea that the performance in the IQ test is an inferior good, whose consumption is reduced when income increases.

In the second experiment in Gneezy and Rustichini (2000a), students going out from house to house collecting monetary donations were divided in three groups: one control group with no reward beside a brief speech on the importance of the activity; two groups with different rewards, consisting of a percentage of the amount collected.

In the meta-analysis in Deci et al. (1999), the verbal rewards (the brief speech in the experiment) can also be considered as (non-tangible) rewards, but were found to enhance intrinsic motivation and not be undermining. Therefore, simplifying the framework, the group receiving only the brief speech can be considered as a control group.

The experiment results report that the tangible-rewarded groups exhibited worse performances than the control group (verbally rewarded) and that the high-rewarded group performed better than the low-rewarded group, but still worse than the control group, leading the authors to choose the title “Pay enough or don’t pay at all”. In the authors’ description, the donation experiment is based on an Israeli tradition where a few ‘donation days’ take place every year and students go from door to door to collect donations. The existence of a ‘tradition’ of door to door collecting is enough to confirm that the reward is temporary, because it is associated with that specific experiment and not with any ‘donation day’. Again, no information is available on the subsequent performances of students collecting donations and it is impossible to test the undermining effect of rewards. In line with the ‘inferior good hypothesis’, the comparison between the rewarded groups and the not rewarded group shows a crowding out, where the negative income effect of the rewards is greater than the positive substitution effect. When the two rewarded groups are compared, the substitution effect seems to prevail. It could be argued that the income effect is stronger when a previously unrewarded activity is paid than when subjects engaged in a rewarded activity receive a pay increase. It seems reasonable that agents consuming free activities drastically change their choices, by reducing the motivated activity, when other expensive goods become available through the payment gained. It is also reasonable that the reduction in free activity is smaller for further income growth when the expensive goods are still available. Further analyses are needed to support this intuition.

A cross sectional analysis among individuals in the same period is also carried out in Frey and Goette (1999). The authors use an econometric approach to the volunteer labour supply to investigate the effect of compensations on the quantity of hours volunteered. They compare a control group (without rewards) and two rewarded groups of volunteers with different payments, finding that rewarded volunteers supply less hours than non rewarded volunteers. A standard problem in volunteering is that it

is not easy to recognise if individuals are volunteers or low paid workers. Estimates do not include variables such as employment status or income, so it is impossible to discern if the rewarded volunteers are simply low paid workers. If they are low paid workers, the reward is the way to attract less motivated individuals and is probably a permanent compensation².

A different design is found in the Frey and Oberholzer-Gee (1997) experiment on the willingness to accept a repository for radioactive waste in the hometown of respondents. The survey was conducted by posing a series of questions to the same sample of individuals. First they were asked if they would be willing to permit the construction of a nuclear waste repository. Secondly, they were asked whether they would be willing to accept the construction if the government had decided to compensate all residents with an annual monetary compensation. The percentage of acceptance of the repository when a reward was offered was lower than in the first question, when no reward was offered. In this study there was no control group, because the same individuals answered all the questions and it is difficult to assess the psychological effect of that specific sequence. Nevertheless, it is clear that the reward is offered permanently and that no inter-temporal substitution effect can be advocated. The undermining effect cannot be observed, whereas the crowding out could be the consequence of the permanent (expected) income increase, which reduces the consumption of the inferior good, in this specific experiment defined as civic duty.

Other empirical/experimental tests often advocated as support for the motivation crowding theory concentrate on dependent variables that are further from the narrow definition of intrinsic motivation. Fehr and Gächter (2000) and Bruni et al. (2009) test reciprocity, Gneezy and Rustichini (2000b) verify the persistence to norm adherence, Frohlich and Oppenheimer (1998) analyse other regarding preferences while Bohnet, Frey and Huck (2001) analyse trustworthiness. In the light of the previous classification of input/output oriented motivations, the latter studies concentrate on output oriented motivations, whereas the first studies discussed refer to input oriented motivations.

² A discussion on the difficulties to test intrinsic motivation in volunteering is in Bruno and Fiorillo (2012).

This brief discussion of some widely cited tests of motivation crowding out shows that evidence from economics is not completely comparable to that found in the psychological literature. More evidence is needed to support the idea that crowding out occurs during the supply reward. In spite of the mixed evidence of empirical economic tests, there is some support for the idea that crowding out emerges because activity performed for intrinsic motives is an inferior good.

7. Concluding remarks

The most important divergence between economic analyses and psychological research on intrinsic motivation emerges when the effects of expected rewards on individual performance are discussed. Most results from psychological research show undermining effects on intrinsic motivation after the reward removal, whereas during the reward supply a better performance can be found. Economics usually studies the effects of rewards during the reward supply, because of its interest in the well functioning of economic incentives, but evidence on crowding out of intrinsic motivation during the reward supply is still limited.

If crowding out is confirmed by further tests, it could be reconciled with evidence on the undermining effects of rewards in psychological research when interpreted in an intertemporal choices framework. The reduced intrinsic motivation that emerges after the reward removal (undermining of intrinsic motivation) can be induced by an intertemporal substitution between two periods: in the first period, the agent exploits his intrinsic motivation in order to gain as much reward as he can, moving the motivated activity from the second to the first period. This substitution effect is the permanent effect of a reward that will be in place if the reward is offered temporarily. This explanation needs the assumption that agents can freely choose the intertemporal allocation of their intrinsic motivation, as occurs for consumption.

On the other hand, intrinsic motivation can be crowded out by a reward, and the performance reduced during the reward supply, if the permanent reward is offered for an activity valued as an inferior good. Individuals who feel richer will decrease their consumption of inferior goods and increase their consumption of normal goods. The income effect will produce a decrease in the activity freely chosen for intrinsic

motives if it is valued as an inferior good. This explanation needs the assumption that the free choice activity is a source of individual utility easily affordable at a low-income level. Discussion on the evidence on crowding out in economic science supports the idea that motivated activities can often be interpreted as inferior goods, whose consumption is reduced when income increases because a wider range of goods becomes available.

More research is needed to empirically test the inferior good hypothesis with data on the same individual experiencing different levels of income (or reward), compared with adequate control groups.

REFERENCES

- Andreoni, J., 1990. Impure Altruism and Donations to Public Goods: a Theory of Warm-Glow Giving, *Economic Journal*, 100, 464-77.
- Bénabou, R. and Tirole, J., 2003. Intrinsic and Extrinsic Motivation, *Review of Economic Studies*, 70, 489-520.
- Bénabou, R. and Tirole, J., 2006. Incentive and Prosocial Behavior, *The American Economic Review*, 96 (5), 1652- 1678.
- Bohnet, I., Frey, B. S. and Huck, S., 2001. More Order with Less Law: On Contract Enforcement, Trust and Crowding, *American Political Science Review*, 95, 131-134.
- Bolle, F. and Otto, P. E., 2010. A Price Is a Signal: on Intrinsic Motivation, Crowding-out, and Crowding-in, *Kyklos*, 63(1), 9-22.
- Bruni, L., Corazzini, L., and Stanca, L., 2009. Testing theories of reciprocity: Do motivation matter?, *Journal of Economic Behaviour & Organization*, 71(2), 233-245.
- Bruno, B. and Fiorillo, D., 2012. Why Without Pay? Intrinsic Motivation in Unpaid Labour Supply, *Journal of Socio Economics*, 41, 659-669.
- Cameron, J., Pierce, W. D., 1994. Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64, 363-423.
- Cooley, C., 1902. *Human Nature and the Social Order*. New York: Scribner's.
- Deci E. L., 1971. Effects of Externally Mediated Rewards on Intrinsic Motivation, *Journal of Personality and Social Psychology*, 18, 105-115.
- Deci, E. L., Koestner, R., Ryan, R. M., 1999. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation, *Psychological Bulletin*, 125 (6), 627-668.
- Deci, E. L., Ryan, R. M., 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being, *American Psychologist*, 55(1), 68-78.

- Deci, E. L., Ryan, R. M., van Steenkiste M., 2008. "Self Determination theory and the explanatory role of psychological needs in human well-being", in Bruni L. et al. (Eds) *Capabilities and Happiness*. Oxford University Press.
- Eisenberger, R., Cameron, J., 1996. Detrimental effects of rewards: Reality or myth?, *American Psychologist*, 51, 1153-1166.
- Eisenstein, N., 1985. Effects of contractual, endogenous, or unexpected rewards on high and low interest preschoolers. *Psychological Record*, 35, 29-39.
- Fehr, E. and Fischbacher, U., 2002. Why Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives, *The Economic Journal*, vol. 112, n.478, C1-C33.
- Fehr, E. and Gächter, S., 2000. Do Incentive Contracts Crowd Out Voluntary Cooperation?, *Working Paper n.34. Institute for Empirical Research in Economics*. Working Paper Series. Zurich University.
- Frey, B. S., 1992. Tertium Datur: Pricing, Regulating and Intrinsic Motivation, *Kyklos*, 45 (2), 161-184.
- Frey, B. S. and Goette, L., 1999. Does Pay Motivate Volunteers? Unpublished Manuscript, Institute for Empirical Research in Economics. Zurich University.
- Frey, B. S. and Jegen, R., 2001. Motivation Crowding Theory: A Survey of Empirical Evidence, *Journal of Economic Surveys*, 15(5), 589-611.
- Frey, B. S., Oberholzer-Gee F., 1997. The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out, *The American Economic Review*, 87 (4), 746-755.
- Frohlich, N. and Oppenheimer, J., 1998. Optimal Policies and Ethical behaviour: Some Unexpected Contradiction. Unpublished Manuscript. Meeting of the preferences Network Research Group.
- Galbraith, J. R., 1977. *Organization Design*. Reading, Addison Wesley.
- Geen, R., 1994. *Human motivation: A psychological approach*. Wadsworth Publishing.
- Gneezy, U., Rustichini A., 2000a. Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, 115 (3), 791-810.

- Gneezy, U., Rustichini A., 2000b. A Fine is a Price, *Journal of Legal Studies*, 29 (1), 1-18.
- Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and Why Incentives (Don't) Work to Modify Behavior, *The Journal of Economic Perspectives*, 25 (4), 191-209.
- Harvey, S. J. Jr., 2005. Why did you do that? An Economic Examination of the Effect of Extrinsic Compensation on Intrinsic Motivation and Performance, *Journal of Economic Psychology*, 26, 549- 566.
- Lindenberg, S., 2001. Intrinsic Motivation in a New Light, *Kyklos*, 53, n.2/3, 317-342.
- Meier, S. and Stutzer, A., 2008. Is Volunteering Rewarding in Itself?, *Economica*, 75, 39-59.
- Olson, M., 1965. *The logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.
- Rummel, A., Feinberg, R., 1988. Cognitive Evaluation Theory: A meta-analytic review of the literature. *Social Behavior and Personality*, 16, 147-164.
- Sliwka, D., 2007. Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes, *American Economic Review*, vol. 97, 3.
- Staw, B., 1989. Intrinsic and Extrinsic Motivation, in H. J. Leavitt, L. R. Pondy, and D. M. Boye (eds) *Reading in Managerial Psychology*, University of Chicago Press.
- Tang, S. H., Hall, V. C., 1995. The overjustification effect: A meta-analysis, *Applied Cognitive Psychology*, 9, 365-404.
- Wiersma, U. J., 1992. The effects of extrinsic rewards in intrinsic motivation: A meta-analysis, *Journal of Occupational and Organizational Psychology*, 65, 101-114.