On the Wondrous Behaviour of Rats and Researchers

Antonio Villar

Universidad Pablo de Olavide & Ivie, Seville, Spain

Email: avillar@upo.es

Abstract

This paper deals with the connection between effort and rewards, regarding the optimal behaviour of the agents involved. The context is that in which a subject is required to perform a costly task and is rewarded with a prize when the task is performed properly. We provide here a simple model that explains why reducing the reward associated with a costly task may induce a higher effort. This is an experimental observation of the behaviour of rats in the lab, but it also appears in other mammals (e.g. researchers).

Keywords

Effort, Costly Task, Reward, Behavioural Patterns

1. The Wonder

After a stormy meeting with the politicians that decide on research funding in my Region, I was complaining with some colleagues about the erratic nature of the research policy. I mentioned that it was really surprising to observe that researchers kept complying with the rules and making substantive efforts in spite of the random nature of some funding policies. Even when research funds were curtailed. An experimental psychologist told me that the same wondrous behaviour could be observed in the lab when rats were subject to the manipulation of their rewards. Reducing the frequency of the rewards, associated with the performance of some task, yielded the unexpected result that some experimental subjects worked harder, exhibiting a compulsive and aggressive behaviour matched with unhappiness. "Just like us", he concluded. So there seemed to be some common driving force between rats and researchers in the presence of random rewards.

Back home I kept thinking about this comparison, trying to figure out how to uncover that common behavioural pattern. The model below is the answer I

propose: a simple analytic framework that is able to reproduce the bizarre pattern of behaviour of rats and researchers. The bottom line is actually one of the most elementary psychological principles: the repetition of the stimulus reduces the intensity of the response. Or, put more formally, the concavity of the objective function (actually the degree of concavity) is what drives this behaviour of the experimental subject. Let us see how.

2. The Explanation

Consider a situation in which an experimental subject (a rat, say) is required to exert an effort or perform a given task that is costly. He is granted a reward proportional to the number of times he does it properly.[1] The experiment is repeated continuously for a given time span, that we normalize to one for the sake of simplicity. The controller decides on the frequency π of correct actions that will be rewarded with a fixed prize, of size 1 (that can be interpreted as one unit of food). The behaviour of the subject responds to an objective function that incorporates effort and rewards as the conditioning variables. More specifically, the subject makes an effort decision $e \in [0,1]$, to be interpreted as the number of times (actually the fraction) the subject does what is required. The reward that the subject receives along the experiment is proportional to his effort, with a degree of proportionality given by the frequency of the prizes set by the controller. In other words, an effort e yields a reward πe, where $\pi \in [0,1]$ is decided by the controller. That is, π is the expected reward per unit of effort.

The experimental subject derives satisfaction from the rewards and dissatisfaction from the effort. This can be formulated in terms of the following function:

$$U(e,\pi) = f\big((1-e), \pi e\big) \tag{1}$$

where e denotes effort, $(1 - e)$ is the satisfaction derived of the effort avoided, and πe the expected reward. We assume that function f is increasing in the expected reward, πe, but that the increase of satisfaction grows at a decreasing rate. That is,

$$\frac{\partial f}{\partial (\pi e)} > 0, \ \frac{\partial^2 f}{\partial (\pi e)^2} \leq 0$$

(in other words, f is increasing in and concave in the expected reward).

Note that the effort variable has two different effects. On the one hand, it is a source of dissatisfaction. On the other hand, it is positively related to the expected reward. The relationship between effort and satisfaction is described by the derivative of $U(.)$ with respect to e. That is,

$$\frac{\partial U}{\partial e} = -\frac{\partial f}{\partial (1-e)} + \frac{\partial f}{\partial (\pi e)} \pi \tag{2}$$

---

[1]We do not consider here any punishment (electroshocks, say), even though they can be easily accommodated within the model. One can interpret that the subject has been deprived from food for a relevant while, so that no food is already a punishment.

11

Consequently, satisfaction is positively related to effort whenever $\dfrac{\partial f}{\partial(\pi e)}\pi > \dfrac{\partial f}{\partial(1-e)}$ and will be negatively related otherwise. So it will depend on the specific shape of function f and on the particular point we consider. Note that the concavity of f in the expected rewards suggests that this relationship will be positive for low values of the effort and negative for high ones. Be as it may, the agent will choose the optimal effort, e^{*}, which is the value that satisfies the following equation:

$$\frac{\partial f}{\partial(1-e)} = \frac{\partial f}{\partial \pi e}\pi \qquad (3)$$

that is, the value of the effort such that the subject's incremental satisfaction due to the prize obtained equals the dissatisfaction derived from the effort. This equalization or marginal effects is the standard requirement for optimal actions.

The relationship between optimal effort and the frequency of the prizes is a subtle one, because larger rewards induce two opposite effects. On the one hand, there is a tendency to increase effort due to the fact that each unit of effort becomes more rewarding. On the other hand, there is a tendency to reduce effort because now with less effort the subject may achieve the same prize.

Which effect eventually dominates depends on the shape of f, in particular on its curvature (the degree of concavity of the function). Recall on this point that the curvature of a function is controlled by its second derivative and that it can be expressed in terms of the elasticity of its first derivative.[2] In our case that elasticity measures the relative change in the marginal satisfaction of the subject due to a change in the expected reward.

By letting ρ denote such an elasticity, we would have:

$$\rho = -\frac{\partial^2 f}{\partial(\pi e)^2}\frac{\pi e}{\dfrac{\partial f}{\partial(\pi e)}} \qquad (4)$$

When $\rho < 1$ we say that the incremental satisfaction is inelastic (an increase of 1% in the expected reward makes the incremental satisfaction change by less than 1%). Values of $\rho > 1$ indicate that the derivative of f varies more than proportionally with respect to the change in the rewards.

This elasticity corresponds to the Arrow-Pratt coefficient of relative risk aversion, in the context of expected utility theory when dealing with monetary lotteries [1]. The meaning of this measure is simple: it tells us the rate at which marginal satisfaction decreases when there is a 1% increase in utility due to an increase in the probability of getting a given prize. When $\rho > 1$ the individual's marginal utility decreases more than 1% when utility increases in 1%. And

---

[2]The elasticity ε of a function g at a given point x is simply the ratio of the relative variations of the function and the variable (usually with a minus sign when the function is decreasing and taking limits when $\Delta x \to 0$). That is,

$$\varepsilon_{\Delta x \to 0} = -\frac{\Delta g(x)/g(x)}{\Delta x/x}$$

vice-versa.

The following result is thus obtained:

Proposition: *Optimal effort responds positively (*resp. negatively*) to a reduction in the frequency* π *of the rewards if and only if* $\rho > 1$ *(*resp. $\rho < 1$*), where* ρ *is the elasticity of the incremental satisfaction of the rewards.*

(The proof is given in the **Appendix**.)

This result establishes that, for all values of $\rho > 1$, effort increases with the reduction in the frequency of the rewards. The larger this coefficient, the larger the increase in effort exerted in response to a reduction in the expected prize. This result explains why we may observe that experimental subjects work harder when the prize is given with a smaller frequency. The more responsive the subject, the larger the effort increase associated with a reduction in the prize. That may also explain compulsive behaviour in some subjects.

Note that the reduction of the frequency of the prize makes those subjects with $\rho > 1$ unhappy in a twofold way. On the one hand, they get on average smaller rewards. On the other hand, they exert a higher effort. Yet working harder is their best response!

The simplest case of a utility function that permits one discussing the role of this coefficient ρ is that in which it is constant. The family of functions with Constant Elasticity of Substitution, CES, yields the following formula in our case:

$$U(.) = \left[a(1-e)^{\frac{\rho-1}{\rho}} + b(\pi e)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}$$

The derivative of U with respect to e in this case is given by:

$$\frac{\partial U}{\partial e} = \frac{\rho}{\rho-1} \left[a(1-e)^{\frac{\rho-1}{\rho}} + b(\pi e)^{\frac{\rho-1}{\rho}} \right]^{\left(\frac{-1}{\rho}\right)} \left[-a\frac{\rho-1}{\rho}(1-e)^{\left(\frac{-1}{\rho}\right)} + b\frac{\rho-1}{\rho}(\pi e)^{\left(\frac{-1}{\rho}\right)}\pi \right]$$

The sign of this expression depends on the sign of the second term of the right hand side. The optimal level of effort is obtained when that term is zero. That is,

$$b(\pi e)^{\frac{-1}{\rho}}\pi = a(1-e)^{\frac{-1}{\rho}}$$

$$\Rightarrow e^* = \frac{a^{-\rho}}{\left(b^{-\rho}\pi^{1-\rho} + a^{-\rho}\right)}$$

From this it follows immediately that:

$$\frac{de^*}{d\pi} \begin{cases} > 0 & \text{if } \rho > 1 \\ = 0 & \text{if } \rho = 1 \\ < 0 & \text{if } \rho < 1 \end{cases}$$

The case in which $\rho = 1$ (which corresponds to a Cobb-Douglas utility function) yields:

$$e^* = \frac{b}{a+b}$$

a value that is independent on the frequency of the rewards. In this case those two opposite effects derived from a change in the frequency are exactly of the same magnitude, so that one cancels the other.

3. Discussion: What to Expect?

It seems to follow logically from the result above that starvation maximizes the subject's willingness to cooperate, when $\rho > 1$. Therefore a path of reductions in the research funds would induce researchers to achieve the highest possible production levels.[3] Yet this need not be the case because the result in the Proposition holds under two implicit assumptions: *pattern recognition* and *no recall*. Pattern recognition means that the experiment involves a sufficient number of rounds so as to give the experimental subject the opportunity to recognize the pattern of the rewards (frequency). No recall refers to the fact that differences in the effort decisions associated with different frequencies are to be interpreted as corresponding to the outcomes of different experiments with identical subjects for which the controller applies different rates of rewards. Otherwise past rewards will most likely enter the objective function *U*.

The results might actually be different when those implicit assumptions are violated. On the one hand, if the subject cannot recognize the pattern of the rewards (*i.e.* rewards are perceived as independent on the action), he will end up by making zero effort. This is so because in that case the behaviour is governed by the equation:

$$U(e,\pi) = f\big((1-e),K\big)$$

for some constant *K*. In this case the optimal decision is clearly $e^* = 0$. This outcome is reminiscent of Seligman [2] theory of *learned helplessness*. On the other hand the behaviour will be different when there is recall, as shown in the Morris [3] water maze experiment. Putting past rewards in the objective function opens a new set of possibilities and the degree of convexity of the function relative to that variable will again play a role in the determination of the behaviour (e.g. taken as a benchmark may induce frustration when the frequency is reduced and hence reduce effort or, alternatively, still more effort is exerted in order to try to achieve previous outcome).

4. Final Remarks

We have analysed here a behavioural pattern observed both in the lab and in some humans (rats and researches, in our reference model) that seems rather counter-intuitive. It refers to the response derived from reducing the expected reward associated with performing a costly task. In some cases, reducing the rewards results in a higher effort exerted by the subjects. We have shown that

---

[3]That statement should be modulated when we introduce some additional elements, which are relevant in real life, such as the existence of minimal nutritional requirements per unit of effort (or minimal funds to set a lab or doing research, if we think of researchers). In that case we would find a threshold below which the correlation established fails.

there are particular circumstances in which the bizarre experimental behaviour of rats and researchers can be rationally explained in terms of agents that try to maximise their achievements (nourishment, satisfaction, welfare …). The key element for that behaviour is the sensitivity of the marginal response to changes in the reward, which is reflected in the degree of concavity of the objective function. Yet increasing effort when the frequency of prizes is reduced for those sensitive subjects is not a universal law, as their behaviour may be also affected by some other aspects, such as pattern recognition and recall.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Pratt, J.W. (1964) Risk Aversion in the Small and in the Large. *Econometrica*, **32**, 122-136. https://doi.org/10.2307/1913738

[2] Seligman, M.E. (1972) Learned Helplessness. *Annual Review of Medicine*, **23**, 407-412. https://doi.org/10.1146/annurev.me.23.020172.002203

[3] Morris, R. (1984) Developments of a Water-Maze Procedure for Studying Spatial Learning in the Rat. *Journal of Neuroscience Methods*, **11**, 47-60. https://doi.org/10.1016/0165-0270(84)90007-4

Appendix. Proof of the Proposition

The first order conditions of a maximum of function U are given by:

$$\frac{\mathrm{d}U}{\mathrm{d}e} = 0 \Rightarrow -\frac{\partial f}{\partial(1-e)} + \frac{\partial f}{\partial \pi e}\pi = 0 = h(\pi, e)$$

This condition is satisfied for some particular values (π^*, e^*). Then, in a neighbourhood of (π^*, e^*) the function $h(\pi, e) = 0$ defines e as an implicit function of π, that is, $e = g(\pi)$ (assuming the necessary conditions for the implicit function theorem). Then we have:

$$\frac{\partial h}{\partial e} = \frac{\partial^2 f}{\partial(1-e)^2} + \frac{\partial^2 f}{\partial(\pi e)^2}\pi^2 \le 0$$

This derivative is negative because function f is concave in πe and increasing in e. Then we can write:

$$g'(\pi) = -\frac{\dfrac{\partial h}{\partial \pi}}{\dfrac{\partial h}{\partial e}} = -\frac{\dfrac{\partial^2 f}{\partial(\pi e)^2}e\pi + \dfrac{\partial f}{\partial(\pi e)}}{\dfrac{\partial^2 f}{\partial(1-e)^2} + \dfrac{\partial^2 f}{\partial(\pi e)^2}\pi^2}$$

The sign of this derivative coincides with the sign of the numerator:

$$\frac{\partial^2 f}{\partial(\pi e)^2}e\pi + \frac{\partial f}{\partial(\pi e)}$$

The first term of this equation is negative, due to the concavity of f in πe, whereas the second term is positive, as f is increasing in πe. Now observe that:

$$\frac{\partial^2 f}{\partial(\pi e)^2}e\pi + \frac{\partial f}{\partial(\pi e)} = \left(\frac{\dfrac{\partial^2 f}{\partial(\pi e)^2}}{\dfrac{\partial f}{\partial(\pi e)}}e\pi + 1\right)\frac{\partial f}{\partial(\pi e)} = (1-\rho)\frac{\partial f}{\partial(\pi e)}$$

where,

$$\rho = -\frac{\dfrac{\partial^2 f}{\partial(\pi e)^2}}{\dfrac{\partial f}{\partial(\pi e)}}e\pi$$

Therefore, the derivative $\mathrm{d}e^*/\mathrm{d}\pi$ turns out to be negative (resp. positive) if and only if $\rho > 1$ (resp. $\rho < 1$).

Q.e.d.