THE PAT ON THE BACK PARADOX AND HUMAN ADJUSTMENT TO ECONOMIC INCENTIVES[1]

Ido Erev
Columbia Business School
Ernan Haruvy
Harvard Business School
Orit Perry
Technion, Dept of Industrial Engineering and Management

Abstract

Informal rewards, also known as pats on the back, are shown to be extremely effective rewards in some settings, prompting some managers to declare them as more "valuable" to employees than monetary rewards. On the other hand, salary and benefits continue to be the main motivators used by firms, and are clearly not substitutable by pats on the back. We name this puzzling dichotomy "the pat on the back paradox" and seek behavioral regularities that could explain this puzzle.

[1] Very preliminary. Do not quote or circulate.

1. Introduction

Extant studies of the effect of monetary and informal rewards reveal two robust but inconsistent regularities. First, questionnaire-based studies show that informal rewards, also known as "pats on the back" appear to be far more effective than pay raises and promotions. For example, in a survey of American workers, 63% indicated a "pat on the back" to be an effective incentive (Lovio-George, 1992). In other survey-based studies (Graham and Unruh, 1990), pat-on-the-back incentives are shown to be more effective than monetary rewards. Such findings have been attributed to the recognition bestowed by the pat on the back and have prompted statements such as: "There are two things people want more than sex and money ... recognition and praise." (Nelson, 1994, quoting Mary Kay Ash, founder of Mary Kay Cosmetics).

In sharp contrast to the informal reward view, economists attribute effort solely to wage. The wage-effort relationship has been recognized as a fundamental building block of labor economics from its early days as a discipline. One of the undisputed fathers of modern economics, Alfred Marshall, was among the first to note that worker effort was not fixed (Marshall, 1936, book 6, Ch. 2) and depended on pay. This assertion was demonstrated early on with extensive studies (e.g., Baldamus, 1957; Lytle, 1942; Ryan, 1947). The wage-effort relationship is crucial in explaining cross-section wage differentials and unemployment patterns (e.g., Akerlof and Yellen, 1990).

This wage-effort dependence is also noted by human resource management experts, who agree that salary and benefits are crucial in hiring and retaining employees and in keeping morale high (e.g., Messmer and Taylor, 2001). Recently, firms in specific industries (see e.g., Berta, 2001; Lazar, 2001) have found themselves unable to avoid raising salaries and bonuses to keep employees motivated and retain them in the firm. The apparent dichotomy in the human resource literature between situations where informal rewards are most effective and situations where monetary rewards are most effective is puzzling. If the "recognition and praise" explanation to informal rewards' effectiveness is correct, it is unclear why such "recognition and praise," valued more than money and sex in one realm, cannot substitute for money in other realms. In fact, some of the most contentious salary disputes in the past decade have been in professional sports

and more recently in the entertainment industry, both of which are saturated with recognition, praise and other pats on the back. We refer to this apparent inconsistency as the pat on the back paradox.

The main goal of the current research is to examine if this paradox may be due to the basic properties of human adjustment to economic incentives. To address this goal we review the relevant known properties and explore the conditions under which they are expected to lead to a pat on the back paradox. The results of this analysis show that the conditions that lead adaptive agents to be more sensitive to pats on back than to monetary incentives are wide but have clear boundaries. A positive effect is predicted when the pat on the back signals a probabilistic future value, which in expectation makes the reinforced behavior attractive. Surprisingly, in these settings a pat on the back is more effective than immediate play of the lottery. Thus, the paradox can be explained with the assertion that the positive effect of a pat on the back would be best observed in organizations where it signals future monetary rewards, whereas monetary rewards would be more effective where a pat on the back cannot reliably signal a lottery.

The paper is organized as follows. Section 2 reviews the evidence for the existence of a pat on the back paradox. Section 3 presents the relevant properties of human adaptation to economic incentives and derives (using a descriptive model that summarizes these properties) the conditions under which pats are expected to be effective. Section 4 presents a simple experiment that demonstrates that the factors captured by the model are, indeed, sufficient to produce a pat on the back paradox. Section 5 discusses the relationship of the current explanation/reason for the paradox to alternative explanations.

2. The pat on the back paradox

Informal rewards, often referred to collectively as pats on the back are low cost or no cost, often verbal, rewards that have virtually no monetary market value. Nelson (1994) notes that such rewards are most effective with an immediate on-the-spot recognition of a job well done by their manager. They can be congratulatory remarks, praise, or symbolic items. Nelson (1994) relates the story of an incident that occurred at the Hewlett-Packard Company. An engineer solved a difficult problem and, overcome

with excitement, burst into his manager's office with the news. The manager offered the only reward he could find at the moment—a banana left over from his lunch[2].

As noted in the previous section, the suggestion that informal rewards can be more effective than monetary incentives comes from questionnaire-based research. This line of research shows that 81% of workers and 60% of managers do not associate monetary compensation with productivity increases (Schneier, 1989). There is also evidence that employees who receive cash bonuses for good performance tend to come to expect them, whether they deserve them or not (Hayes, 1999).

In a survey of American workers, 63% indicated a "pat on the back" to be an effective incentive (Lovio-George, 1992). Graham and Unruh (1990), in a study on effective incentives, report the following four incentives as ranked in the top five out of 65 potential incentives: (1) A verbal congratulation by the manager following a job well done, (2) a personal congratulatory note from the manager following a job well done, (3) a public recognition following a job well done, and (4) a morale-building meeting between manager and worker following a job well done. It appears that companies recognize the value of such reinforcements. For example, Tektronix, Inc., a manufacturer of oscilloscopes and other electronic instruments located in Beaverton, Oregon issues "You Done Good" notecards for managers to award deserving employees in the company (Nelson, 1999).

An outsider to the field of human resource management could falsely conclude from the above research findings that companies would save a great deal of money by shifting resources from monetary incentives to pat-on-the-back type of incentives. A possible retort to such a suggestion would be to point at the pat-on-the-back rich entertainment industry, where recent labor disputes (Writers Guild of America, Screen Actors Guild, American Federation of Television and Radio Artists) have shown pats-on-the-back to not be an effective substitute to monetary rewards. Alternatively, a cursory look at newspaper want ads reveals a lack of ads advertising pats on the back or congratulatory notes.

[2] The Banana Award is now a coveted employee prize at Hewlett-Packard.

3. A simplified Principal-Agent environment

Consider a simple "Principal-Agent" environment in which the agent can select between two alternatives: "Low" or "High" effort levels. The principal can observe each choice with probability $p < 1$ and reinforce each observed high effort choice with probability $q < 1$. Assume that in the absence of a bonus system, the agent gains one money-equivalent unit from each low effort choice and nothing from a high effort choice. The value of the productivity loss from low effort relative to high effort to the principal is $C > 1$. Thus, the principal is motivated to change the agent's incentive structure by reinforcing each of the observed high choices with a bonus of B, such that $1 < pqB < C$. In other words, the expected bonus must be greater than the worker's opportunity cost yet smaller than the firm's expected gain from high effort. In addition, the principal in this model has the ability to reinforce observed high effort choices with a pat on the back. The pat costs C(pat) to the principal and has no direct value to the agent. Assume that the principal can select among four bonus systems:

Delayed monetary rewards
Immediate monetary rewards
Immediate pat on the back
Informative pats: Immediate pats that signal a delayed reward with probability q.

It is easy to see that under the assumption that the agent is rational, the principal should avoid the pat on the back. Under the opposite assumption, that the agent ignores monetary incentives, the principal should avoid monetary rewards. The pat on the back paradox implies that these extreme predictions are not descriptive. It appears that monetary incentives are important but pats on the back can nevertheless be effective. That is, immediate monetary rewards may be more effective than pats without monetary rewards, but less effective than informative pats.

3.1 Three relevant behavioral regularities

Three of the known characteristics of human adaptation to economic incentives appear to be relevant to the "pat on the back" paradox. The best known regularity is the

Law of Effect (Thorndike, 1898), the findings that animals (including humans) tend to respond in an adaptive fashion to immediate feedback. The importance of the immediacy of feedback led Herrnstein and his co-authors (Herrnstein, 1990; Herrnstein and Prelec, 1991; Herrnestein, Lowenstein, Prelec, and Vaughan, 1993) to abstract this regularity as melioration: an attempt to maximize immediate payoffs.

A second relevant characteristic is the sensitivity to secondary reinforcement. Studies of human behavior show that the reinforcing value of a lottery ticket that gives a probability for a future consequential outcome is close to the expected value of the ticket[3]. In a recent research, Perry, Erev and Haruvy (in press) observed that the effect of paying decision-makers with unresolved lotteries (lotteries that will be played at the end of the experiment) is practically identical to the effect of immediate payoff that equal to the expected value of the lottery. For example, the outcome "you got the lottery that gives –200 with probability 0.03; 0 otherwise" had the same effect as the outcome "you got –6 for sure." We refer to this observation as the *EV-like secondary reinforcement effect.*

A final and least known *relevant property is the payoff rank effect* (Barron & Erev, 2001). Experimental study of repeated choice behavior among gambles shows that decision makers are very sensitive to the expected ranking of the gambles. Thus, when expected ranking is incongruent with the expected payoff, experience tends to move decision-makers away from maximization. For example, the average participant in Barron and Erev's study learned (during 400 trials) to prefer "3 with certainty" over the gamble "32 with probability 0.1 and 0 otherwise." Similar pattern was observed in choice among losses: the average participant learned to prefer the gamble "-32 with probability 0.1 and 0 otherwise" over the sure payoff "-3."

3.2 Implications

Under the assumption that the agent is likely to display the three behavioral regularities summarized above, a consultant to the principal can make three obvious recommendations. First, immediate rewards should be preferred to delayed monetary rewards (see Mowrer and Ullman, 1945; Banks and Vogel-Sprott, 1965; Cohen, 1968;

[3] With risk neutrality and odds clearly communicated.

Kamin, 1959; Setterington and Walters, 1964; and Walters, 1964). Second, informative pats should be preferred to uninformative pats. And, when the pat is costless, it is likely to have a positive effect on the principle's payoff.

To derive more precise predictions the consultant should quantify the joint effect of the relevant behavioral regularities. The current analysis uses the quantification proposed by Erev and Barron (2001). This model, referred to as "reinforcement learning among cognitive strategies" (RELACS) was found to provide relatively accurate prediction of choice behavior. For example, after estimating the model's two parameters on 46 experimental conditions, the correlation between its ex ante prediction and observed aggregated choice probabilities in 27 other conditions (similar to the simplified problem analyzed here) was 0.98. The basic assumptions of the model are presented in Appendix 1.

Table 1 presents the expected effect of the three immediate reward rules on the principal's expected payoff for various combinations of parameter values (p = 0.1, 0.3, 0.5; q = 0.01, 0.03, 0.10; B = 100, 200, 400; C = 2, 4, 6) under the assumption of RELACS agents. The results show normalized average payoff per trial to the principal in a 1000 period repeated interaction. The normalization involved setting the firm's net-of-cost payoff from low employee effort to 0. The results reveal that unless C(pat) is very high, the pat-on-the-back paradox emerges: That is, the principal's expected payoff increases with the addition of pats on the back.

4. Experimental demonstration

4.1. Experimental Design

To evaluate the descriptive value and highlight the implications of the present analysis, the current section presents a simple experimental demonstration of the effect of informal rewards. Recall that the principal can observe effort with probability p and conditional on observing high effort can reward it with probability q. We examine the behavior of agents in an environment in which p = 1/3, q = 0.03 and B = 200. In addition, base payoffs to each of the agent's effort choices, previously 1 for Low and 0 for High, are both incremented by a gamble yielding 1 with probability 0.5 and 2

otherwise. Thus, in all the conditions we examined the participants had to select among the following two gambles:

Low effort: (2 with prob 0.5; 3 otherwise)
High effort: (1 with prob 0.5; 2 otherwise) + (200 with probability 0.01; 0 otherwise)

Informal rewards were simulated by a yellow smiley face on the screen. Three bonus conditions were compared:

No smiley: The selected gambles were played after each choice.

Uninformative smiley: In this condition, high effort choices were rewarded with probability 1/3 with a presentation of the yellow smiley face. The presentation of the smiley was not correlated with the outcome of the selected gamble. Both types of feedback (the smiley and the outcome of the selected gamble) were presented in parallel.

Informative smiley: In this condition participants were informed that each presentation of a smiley implied they had earned a lottery ticket that would pay 200 tokens with probability 0.003 at the end of the experiment. As in the previous condition, the smiley faces were presented after high effort choices with probability 1/3. Note that the expected value of the two alternatives is the same in all three conditions.

Twenty-five subjects (engineering students at the Technion) participated in each of two immediate monetary incentive schemes (No smiley and uninformative smiley). Fifty subjects participated in the informative smiley condition. Each treatment involved subjects facing the same two choices repeatedly for 400 periods. Choices were represented by two virtual buttons on a computer screen. The buttons were labeled "Right" and "Left" and subjects were told to select a button in each repetition. Once pressed, the button selected would display the token payoff for the corresponding choice. The number of tokens earned in a given round would be added to the cumulative payoffs displayed at the bottom of the screen. Although the buttons were presented to the subjects in neutral terms, we find it convenient, for the purpose of continuity of the discussion from the last section, to refer to the choices as 'High' and 'Low.' Similarly, though payoffs were in tokens (conversion rate of 40 tokens = 1 Shekel = $0.25), we may think

of payoffs as net reward: monetary reward minus the money-equivalent energy value expended working.

4.2. Predictions and Results

The right hand column of Figure 1 presents the prediction of RELACS (with the parameters estimated by Erev and Barron, 2001) in eight blocks of 50 trials. The payoff rank effect, modeled in RELACS, implies that without the informative smiley the maximization rate (proportion of high choices) will be below 50%. The addition of informative smiley is expected to move behavior toward maximization.

Consistent with the RELACS prediction, the addition of the informative smiley improves the maximization rate from 45% to 55%. This effect is significant at the 5% level[4] (one-tail t-test p-value = 0.041). The effect of uniformative smile is insignificant (2-tail t-test p-value of 0.86).

5. Conclusions

The current analysis shows the two of the robust properties of human adaptation to economic incentives provide sufficient conditions to the emergence of the pat on the back paradox. Since human agents behave as if they (1) underweight of rare events, and (2) treat unresolved lotteries like reward with similar EV, pats on the back that signal possible reward can be very effective. In addition, a third property on human adaptation, underweighting of delayed payoffs, is expected to heightened the effect of pats on the back

Notice that this explanation has two optimistic implications. First, the apparent disagreement between the survey-based research and the empirical economic research is smaller than it seems. Both lines of research capture accurate regularities. Under the current experimental-based explanation pats on the back are truly useful, and monetary rewards are necessary. We believe that this convergence is good news.

A second optimistic implication concerns the possibility of eliminating costly violations of maximization. In the current setting, pats on the back appear to work

[4] If we were to ignore DMs who never change their behavior in response to feedback the significance would increase (p-value < 0.02).

because they move boundedly rational agents toward the efficient behavior predicted under the rationality assumption.

References

Abramowitz, A.J., and S.G. O'Leary (1990), "Effectiveness of Delayed Punishment in an Applied Setting," *Behavior Therapy*, 21, 231-239.

Akerlof, G. A. and J. L. Yellen (1990), The Fair Wage-Effort Hypothesis and Unemployment, *Quarterly Journal of Economics*, Vol. 105, No. 2., pp. 255-283.

Ammons, R.B. (1956), "Effects of knowledge of performance: A survey and tentative theoretical formation," *Journal of General Psychology* 54, 279-299.

Anthony, R. N. and J. Dearden (1989), *Management Control Systems*, Norton M. Bedford ed., 5th Ed., Irwin, Homewood, Ill., p 57.

Baldamus, W. (1957), The Relationship Between Wage and Effort, *Journal of Industrial Economics*, Vol. 5, No. 3, pp. 192-201.

Banks, R.K. and M.D. Vogel-Sprott (1965), "The effect of delayed punishment on an immediately rewarded response in humans," *Journal of Experimental Psychology* 70, 357-359.

Becker, G.S. (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76(2), 169-217.

Barron, G. and I. Erev (2000), "Toward a general descriptive model of one-shot and repeated decision making under risk and uncertainty," Technion working paper.

Berta, Dina (2001), Survey says salaries, bonuses and benefits are on rise, Nation's Restaurant News 35(21), 12-13.

Boyle, D. C. (1987), "Ideas for Action--The 100 Club," *Harvard Business Review*, March-April 1987, p 27.

Cohen, P. S. (1968) "The interactive effects of delay and intensity of shock," *Journal of the Experimental analysis of behavior* 11, 789-799.

Cox, J.C. and R.L. Oaxaca (1995), "Inducing risk-neutral preferences: further analysis of the data," *Journal of Risk and Uncertainty* 11: 65-79.

Cox, J.C., V.L. Smith, and J.M. Walker (1985), "Experimental development of sealed-bid auction theory; calibrating controls for risk aversion," *American Economic Association Papers and Proceedings* 75: 160-165.

Cox, J.C., V.L. Smith, and J.M. Walker (1988), "Theory and individual behavior of first-price auction," *Journal of Risk and Uncertainty* 1: 61-99.

Daniels, Aubrey (1999), "Incentives, Safety and Performance Management," [Online]. Available: http://www.p-management .com/articles/9908.htm.

Ehrlich, I., (1974) "Participation in Illegitimate Activities – An Economic Analysis," in Becker, G.S. and W.M. Landes, eds., *The Economics of Crime and Punishment*, New York: Columbia University Press, 68-134.

Erev, I., and Barron, G. (2001), "On adaptation, maximization, and reinforcement learning among cognitive strategies," Columbia Business School working paper.

Erev, I. and A. Roth (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria," *American Economic Review* **88**, 848-881.

Erev, I., Y. Bereby-Meyer, and A. Roth (1999), "The Effect of Adding a Constant to All Payoffs: Experimental Investigation, and Implications for Reinforcement Learning Models," *Journal of Economic Behavior and Organization* 39, 111-128.

Friedman, M., and L. J. Savage (1948). "The Utility Analysis of Choice Involving Risk," *J. of Political Economy* 56, 279-304.

Graham, G. H. and J. Unruh (1990), "The Motivational Impact of Non-financial Employee Appreciation Practices on Medical Technologists," *Health Care Supervisor*, 8(3), pp 9-17.

Greenberg, J. and M. Liebman (1990), "Incentives: The Missing ink in Strategic Performance," *Journal of Business Strategy*, July-August 1990, p 9.

Grogger, J. (1991), "Certainty vs. Severity of Punishment," *Economic Inquiry*, 29(2), 297-309.

Haruvy, E. and I. Erev (2000), "When to Pursue Variable Pricing: The Relationship Between Price Format and Quality," Working paper, Technion, Israel.

Hayes, S. (1999), Clashing views on cash rewards, Workforce 78(9), 104.

Herrnstein, R. J. (1990), Rational Choice Theory: Necessary But Not Sufficient, American Psychologist, 45, 356-367.

Herrnstein, R. J., G. F. Lowenstein, D. Prelec and W. Vaughan, Jr. (1993), Utility Maximization and Melioration: Internalities in Individual Choice, Journal of Behavior Decision Making 6, 149-185.

Herrnstein, R. J. and Prelec, D. (1991), Melioration: A theory of distributed choice, Journal of Economic Perspectives, 5, 137-156.

Hill, F. C. (1989), "Generating Ideas That Lower Costs and Boost Productivity," *National Productivity Review*, Vol. 8, No. 2, p 161.

Ilgen, D.R., Fisher, D.C. and Taylor, M.S. (1979), "Consequences of individual feedback on behavior in organizations," *Journal of Applied Psychology* 64, 349-371.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decisions Under Certainty," *Econometrica*, 47, 263-291.

Kamin, L.J. (1959) "The delay of punishment gradient," *Journal of Comparative and Physiological Psychology* 52, 434-437.

Kanter, R. M. (1986), "Kanter on Management--Holiday Gifts: Celebrating Employee Achievements," *Management Review*, December 1986, pp 19, 21.

Lazar, Gerald (2001), Keep your key people, *Electronic Business* 27(5), 86-94.

Lovio-George, C. (1992), "What Motivates Best?" *Sales & Marketing Management*, April 1992.

Lytle, Charles Walter (1942*), Wage incentive methods, their selection, installation and operation*, New York: Ronald Press.

Maciariello, J. A. and C. J. Kirby (1994), Management Control Systems: Using Adaptive Systems to Attain Control, Prentice-Hall, Englewood Cliffs, NJ.

Marshall, Alfred (1936), *Principles of Economics*, Book Six: The Distribution of the National Income, 8th edition, London.

McCormick, Michael (2000), "Performance Management Through Behavior Modification," [Online]. Available: http://www.p-management .com/articles/2013.htm.

Messmer, M. and Taylor, L. E. (2001), Motivating Employees for Dummies, New York: Hungry Minds.

Mowrer, O. H. and A. D. Ullman (1945), "Time as a determinant of integrative learning, Psychological Review 52, 61-90.

Nelson, B. (1994), *1001 Ways to Energize Employees*, New York: Workman Publishing Company.

Nelson, B. (1999), "The Use of Informal Rewards in Recognizing Performance" [Online]. Available: http://www.p-management .com/articles/9902.htm.

Nord, W. R. (1974), "Beyond the Teaching Machine; Operant Conditioning in Management," From Henry L. Tosi and W. Clay Hammer's, *Organizational Behavior and Management: A Contingency Approach*, St. Clair Press, Chicago, 1974, pp 385, 387, 395, 397, 398.

Ng., Y. K. (1965). "Why Do People Buy Lottery Tickets? Choice Involving Risk and the Indivisibility of Expenditure," *Journal of Political Economy* 73, 530-535.

Retting, R. A., A. F. Williams, D.F. Preusser and H.B. Weinstein (1995), "Classifying Urban Crashes for Countermeasure Development." *Accident Analysis and Prevention* 27, 283-294.

Retting, R. A., A. F. Williams, C.M. Farmer and A.F. Feldman (1999a), "Evaluation of Red Light Camera Enforcement in Oxnard, California" *Accident Analysis and Prevention* 31, 169-174.

Retting, R. A., A. F. Williams, C.M. Farmer and A.F. Feldman (1999b), "Evaluation of Red Light Camera Enforcement in Fairfax, Va., USA" *ITE Journal*, August, 30-34.

Robson, A. (1996). "The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth*," Games and Economic Behavior* 14, 190-207.

Roth, A. E. and Malouf, M.W.K. (1979), "Game theoretic models and the role of information in bargaining," *Psychological Review* 86: 574-594.

Ryan, Thomas A. (1947), *Work and effort*, New York: Ronald Press Co.

Schneier, C. E. (1989) "Capitalizing on Performance Management,

Recognition, and Rewards Systems," *Compensation and Benefits Review*, March-April 1989, p 23.

Smith, A. B. C. (1961), "Consistency in Statistical Inference and Decision," *Journal of the Royal Statistical Society* 23:1-25.

Thaler, R., Tversky, A., Kahneman, D., and A. Schwartz (1997): "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test," *Quarterly Journal of Economics*, 112, 647-661.

Tversky, A., and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 9, 195-230.

Vroom, V. H. (1976) "Leader," in M.D. Dunnette (ed.) Handbook of Industrial and Organizational Psychology, Rand McNally, Chicago, 1976, pp 1527-51.

Walker, J.M., V.L. Smith, and J.C. Cox (1990), "Inducing Risk Neutral Preferences: An Examination in a Controlled Market Environment," *Journal of Risk and Uncertainty* 3: 5-24.

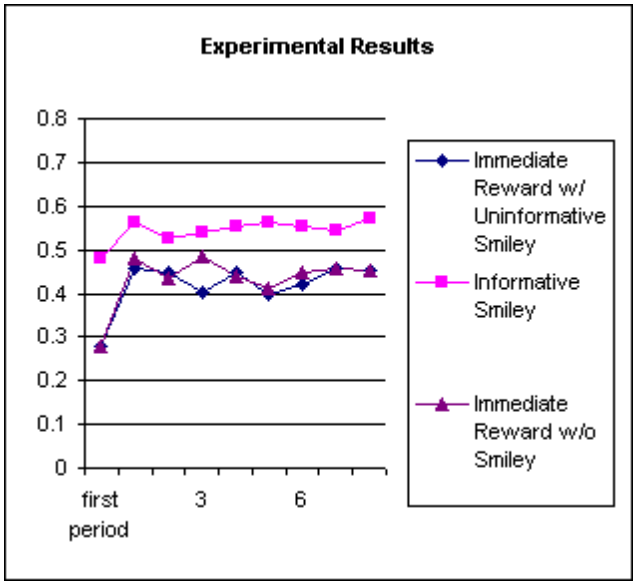
Walters, R. H. (1964), "Delay of reinforcement gradients in children's learning," *Psychonomic Science* 1, 307-308.

Walters, R. H. and L. Demkow (1963), "Timing of Punishment as a determinant of response inhibition," *Child Development* 34, 207-214.

Figure 1

Experiment

Work	Shirk
1 (0.5) 2 (0.5) + 200 (0.01)	2 (0.5) 3 (0.5)



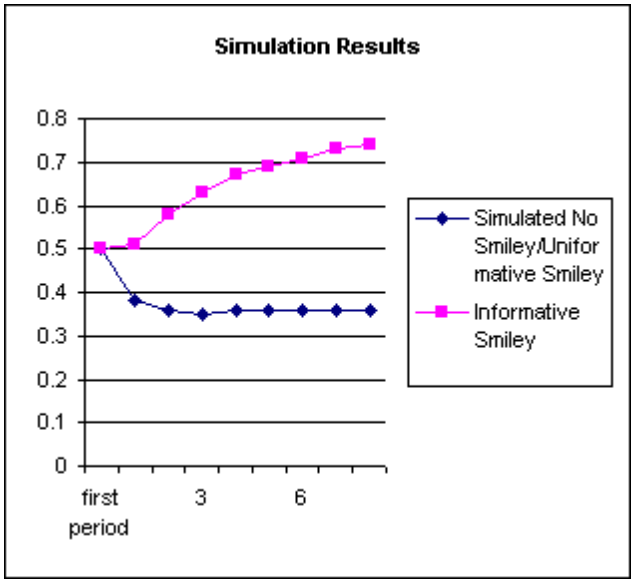


Table 1. The expected effect of different reward rules in different environments under the assumption of RELACS agents. The expected payoff was normalized by setting the payoff from low effort to 0. Thus, C is the maximal possible payoff.

				Expected payoff to the firm		
p	q	B	C	No reward	Monetary reward	Informative pat
0.1	0.03	400	2	0.02	0.06	0.34
0.1	0.03	400	4	0.03	0.06	1.05
0.1	0.03	400	6	0.05	0.20	2.06
0.1	0.10	200	4	0.03	0.61	1.28
0.1	0.10	200	6	0.05	1.10	2.49
0.1	0.10	400	6	0.05	1.06	1.47
0.3	0.01	400	2	0.02	0.04	0.43
0.3	0.01	400	4	0.04	0.13	1.51
0.3	0.01	400	6	0.05	0.08	2.56
0.3	0.03	200	2	0.02	0.05	0.14
0.3	0.03	200	4	0.04	0.57	1.52
0.3	0.03	200	6	0.04	1.08	2.87
0.3	0.03	400	4	0.04	0.18	0.31
0.3	0.03	400	6	0.05	1.24	1.87
0.3	0.10	100	4	0.03	0.61	0.77
0.3	0.10	100	6	0.05	1.83	2.29
0.5	0.01	400	4	0.03	0.32	1.53
0.5	0.01	400	6	0.05	0.73	3.09
0.5	0.03	100	2	0.02	0.12	0.35
0.5	0.03	100	4	0.04	0.72	1.76
0.5	0.03	100	6	0.05	1.11	3.15
0.5	0.03	200	4	0.03	0.56	0.82
0.5	0.03	200	6	0.06	1.59	2.45
0.5	0.10	100	6	0.06	0.72	0.85

Appendix 1: RELACS

RELACS, the model used here is an extension of the model proposed by Erev, Bereby-Meyer and Roth [1999]. It can be summarized by the following assumptions:

Strategies and decisions:

The decision makers are assumed to consider two sets of strategies: Direct (e.g., Low or High), and Cognitive. The cognitive strategies are rules that condition the action in trial t + 1 on the outcome of the first t trials. Two cognitive strategies are assumed: (1) hill climbing and (2) loss avoidance.

In the current, binary choice, setting the hill climbing strategy implies a selection of the action with the highest recent obtained payoff (and a random selection in the case of a tie). To follow this strategy the subjects are assumed to recall the last payoff obtained from each strategy. The initial value of the last payoff record is A(1) – the expected payoff from random choice.

Since all the payoffs in the present problems are positive, the loss avoidance strategy implies indifference. The indifference is resolved by a selection of one of the direct strategy according to the rule described below.

At each trial, the agent is assumed to select one set of strategies, Direct or Cognitive, and then to select one of the strategies in the selected set. Both selection phases are assumed to obey the same choice rule.

Reinforcement learning: The probability $p_j(t)$ that a DM selects option j in decision d (d=1 involves the decision among the sets, d=2 is the choice among the strategies in the selected set) at time t is given by the following variant of Luce's (1959) choice rule,

$$p_j(t) = e^{q_j(t)\lambda(t)} \Big/ \sum_{k=1}^{n_d} \left(e^{q_k(t)\lambda(t)} \right) \qquad (1)$$

where $\lambda(t)$ is a payoff sensitivity (defined below), $q_j(t)$ is the propensity to select option j, and n_d is the number of relevant options ($n_d = 2$ in the choice among the two sets, and the number of strategies in the selected set in the second choice).

If option j is selected at trial t, the propensity to select it in trial t+1, $q_j(t+1)$, is a weighted average of the propensity in t and the obtained payoff R(t).

$$q_j(t+1) = q_j(t)[1 - w(t,d)] + R(t)w(t,d) \qquad (2)$$

The weight of the new reinforcement, $w(t,d)=1/[\eta/[(d)(n_d)]+C_j(t)]$ where η is a parameter that captures the "strength" of the initial value $q_j(1)$, $C_j(t)$ is number of times j was selected in the first t trials. The initial propensity $q_j(1)$ is assumed to equal A (the expected payoff from random choice). On each trial two propensities are updated, one for the chosen set and one for the chosen act or strategy. Propensities of options that were not selected are not updated.

Payoff sensitivity. The payoff sensitivity at trial t, $\lambda(t) = \lambda/S(t)$ where λ is a payoff sensitivity parameter and S(t) is an measure of observed payoff variability. When payoff of unselected alternative are not known (the current case):

$$S(t+1) = S(t)[1 - W'(t)] + |R(t) - A(t)|W'(t) \qquad (3)$$

where $W'(t) = 1/(t+\eta)$ and A(t), a measure of payoff average, is updated like the payoff variability term:

$$A(t+1) = A(t)[1 - W'(t)] + R_j(t)W'(t). \qquad (5)$$

Altogether, the model has two learning parameters. Erev and Barron's estimation yields the values: λ =4.5 and η = 100. The prediction of the model for the current settings were derived by running computer simulations in which virtual agents that behave according to the models assumptions play each of the problems. A simulator can be found in