TI 2006-092/1
Tinbergen Institute Discussion Paper

The Self-Perception Theory versus a Dynamic Learning Model

Otto H. Swank

Erasmus Universiteit Rotterdam, amd Tinbergen Institute.

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

The Self-Perception Theory Versus a Dynamic Learning Model*

Otto H. Swank†

Erasmus University Rotterdam and Tinbergen Institute

October 10, 2006

Abstract

Several economists have directed our attention to a finding in the social psychological literature that extrinsic motivation may undermine intrinsic motivation. The self-perception (SP) theory developed by Bem (1972) explains this finding. The crux of this theory is that people remember their past decisions and the extrinsic rewards they received, but they do not recall their intrinsic motives. In this paper I show that the SP theory can be modeled as a variant of a conventional dynamic learning (DL) model. A comparison between the assumptions underlying the SP model and the DL model shows that the SP model could be relevant in a wide variety of educational contexts. However, the SP model seems less relevant than the DL model in other contexts.

JEL Classification: D82, D83

KEYWORDS: Self-Perception Theory, Dynamic Learning, Incentives, Intrinsic Motivation.

*I would like to thank Josse Delfgaauw, Robert Dur, Chaim Fehrsman and Hein Roelfsema for valuable comments and suggestions.

†Contact: Otto Swank, Erasmus University Rotterdam, H 8 – 13, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Email: swank@few.eur.nl.

1 Introduction

In the last couple of years, "Earning by Learning" programs have been launched, in which children are offered $2 for every book they read. In Dallas, for example, 38,103 students earned $688,012 by reading 344,006 books in this way.[1] A similar program is run by Pizza Hut, a chain of pizza restaurants. In that program, by reading books elementary school students can receive a certificate at school. They can take this certificate to a participating restaurant for an award and a free special pizza.[2] The stated goal of these programs is helping children to develop a love for reading. Will this goal be attained?[3]

In the "Earning by Learning" programs, one hopes for a long run effect (a love for reading) of a temporary reward (a pizza). To gain insight into the long-run effects of temporary rewards one could employ the dynamic learning model developed by Grossman, Kilhstrom and Mirman (1977). In that model a person has to make a decision on an activity in the present and in the future. By experimenting now the person learns about his preferences with respect to the activity, and makes a more informed decision in the future. The model gives an affirmative answer to the question whether "Earning by Learning" programs have positive long-run effects. A reward encourages people to read in the present. As a consequence, more children discover that reading is fun. Hence, more children keep on reading.

I took the example of motivating children to read by offering rewards from a well-known textbook on social psychology (Aronson, Wilson, and Akert, 2005). The example is discussed in the context of a treatment of intrinsic versus extrinsic motivation. Until the early seventies, the dominant view in the psychological literature was that rewards are always positive reinforcers. In the early seventies, a couple

[1] See http://www.eblofdallas.org/index.html.

[2] See http://www.bookitprogram.com/default.asp.

[3] Understanding the long-run effects of temporary rewards is not only important if one wants children to develop a love for reading. It is also key in marketing. For example, by means of sales promotion (money off, free samples, bonus packs etc.), firms sometimes try to raise trial rates, in the hope that new buyers like the product and keep purchasing it (Brown, 1974, Rothschild and Gaidis, 1981). The typical effect of a sales promotion is as follows (Ehrenberg, Hammond, and Goodhart, 1994; Jobber, 2004, p. 595). Sales increase sharply during the promotion. Just after the promotion, sales drop because people have stocked up the product. The long-run effects of sales promotion can be positive, neutral or negative. For well-known brands the long-run effect is usually neutral, but for new brands the long-run effect may be positive. Sometimes, sales promotion devalues a product in the eyes of buyers, leading to a negative long-run effect.

of studies challenged this view (Deci, 1971; Kruglanski, Friedman and Zeevi, 1971; Lepper, Greene and Nisbett, 1973). In several experiments, it was found that "when a person engages in an intrinsically interesting activity, under certain conditions, the imposition of superfluous extrinsic rewards may have detrimental effects on his or her intrinsic motivation" (Lepper and Greene, 1978). The self-perception theory developed by Bem (1972) offers an explanation for these findings. Point of departure of this theory is that people do not completely know themselves. For example, a person does not know whether or not he likes reading. The theory describes how people learn about their selves. Essentially, people can infer information about themselves from their past behavior (did I read in the past?) and the surrounding circumstances (was I rewarded for reading books?). Whether or not a reward was present may influence how people evaluate their past behavior. In the absence of a reward, reading in the past can only be explained by a positive attitude towards reading. By contrast, when a reward was present, a person may infer that he read for the reward rather than for fun. The implication is that a reward for reading in the past may lead a person to abstain from reading today.

The above discussion suggests that we have at least two competing theories that may help us to predict the long-run consequences of temporary rewards. The main objective of the present paper is to identify the conditions under which each theory is valid. To this end, I first present a simple dynamic learning (DL) model. Next, I show that Bem's self-perception SP theory can be formalized as a variation of the simple DL model. A comparison between the assumptions underlying the DL model and the SP model reveals that the two models essentially differ in one respect: they differ in what a person remembers from the past. In the DL model, it is assumed that a person recalls his decision on the activity, how much he enjoyed the activity, and why he undertook the activity. In the SP model, it is assumed that a person remembers his decision and sometimes the environment in which the decision was taken, but he does not remember how much he enjoyed the activity nor his past attitude towards the activity. Having pinpointed the assumptions underlying the two models, we also have identified the conditions under which each model is valid. To answer the question whether offering a child a pizza for reading books encourages this child to read when he is grown up, it is key to know what this child will remember. Does he remember the pizza or does he remember how much he enjoyed reading?

Formalizing Bem's (SP) theory was not only useful for making a comparison between the SP model and the DL model. It also enabled me to disentangle the effects of temporary rewards. In the short-run, rewards encourage persons to undertake activities. At this point, the DL model and the SP model do not differ. It is perhaps good to point out that this is the most commonly held view by social psychologists. As an illustration, consider the following quote from the earlier mentioned textbook on social psychology:

> There is no doubt that rewards are powerful motivators and that pizzas and money will get kids to read more. One of the oldest and most fundamental psychological principles is that giving a reward each time a behavior occurs will increase the frequency of that behavior. Whether it be a rat pressing a bar in order to obtain food pellet or a child reading to get a free pizza, rewards can change behavior (Aronson, Wilson, and Akert, 2005, p. 147).

The long-run effect of a reward on future decisions can be decomposed into two components. First, a person infers from his past decision information about his attitude towards the activity. As a reward is a short-run motivator, this effect is positive. Second, a person tries to infer from the size of the reward information why he opted for the activity (or why not). The higher is the reward, the smaller is the probability that he had a positive attitude towards the activity. It is this partial effect that social psychologists have highlighted in their experiments (see discussion under Proposition 2). However, it is worth emphasizing that also in the SP model, the total long-run effect of a temporary reward may be positive. This requires that the first component dominates the second one.

Our paper contributes to the literature that tries to explain why rewards are sometimes counterproductive (see, for instance, Deci, 1971; Frey, 1997; Gneezy and Rustichini, 2000; Lepper, Greene and Nisbett, 1973). One explanation builds on the idea that people are concerned about the impression they make on others or on themselves (Benabou and Tirole, forthcoming; Janssen and Mendys-Kamphorst, 2004). These concerns induce people to demonstrate their attractive traits (for example, their generosity or their sense of civic duty). Rewards make it harder for people to demonstrate their attractive characteristics and consequently may be counterproductive. Another explanation for why rewards may backfire is that a reward sometimes contains information about the environment in which the agent operates (see, for example, Benabou and Tirole, 2003, and Ellingsen and Johannesson, 2005). For instance, an agent may infer from the existence of a reward that a task is difficult to perform or that it is boring. A reward may also signal distrust. In such cases, a reward may keep an agent from undertaking the task. The present paper does not focus on situations in which agents are concerned about their reputation, or rewards contain relevant information about the environment. I consider situations where rewards make it more difficult for people to assess how much they enjoy an activity. One reason for this choice is that I want my model to describe the setting in the pioneering experiments conducted by Deci (1971) and Lepper, Greene and Nisbett (1973). These experiments initiated the discussion on intrinsic and extrinsic motivation among social psychologists. Another reason for my choice is the belief that some explanations for the observation that rewards may be counterproductive are more relevant than others in *economic* contexts. In fact, the analysis below suggests that extrinsic motivation undermining intrinsic motivation may be relevant in educational contexts, but are less relevant in economic contexts. Of course, when rewards contain information about the environment, or when reputational concerns are important, rewards may also backfire in economic contexts.

This paper is organized as follows. Section 2 presents the DL model. Section 3 and Section 4 discuss two versions of the SP model. Section 5 concludes with a discussion on the question to what extent the SP model is relevant in economic contexts.

2 The Dynamic Learning Model

2.1 The Model

Consider a two-period model, with $t = \{1, 2\}$. In each period t, an agent has to make a decision on an activity $X_t = \{0, 1\}$, with $X_t = 1$ denoting undertaking the activity and $X_t = 0$ denoting not undertaking the activity. In period 1, the decision

X_1 yields payoffs to the agent equal to

$$U_1(X_1 = 1) = p + \mu \tag{1}$$

$$U_1(X_1 = 0) = 0 \tag{2}$$

where $p \geq 0$ is a reward, and μ is a stochastic term representing the (stable) taste of the agent. We assume that μ is uniformly distributed on the interval $[-h, h]$. The agent does not know to what extent he likes activity X. Before deciding on X, he receives a signal s_1 about μ. With probability π, s_1 reveals μ, with probability $1 - \pi$, s_1 does not contain information about μ. The signal s_1 can be interpreted as the outcome of introspection or it can be a judgmental heuristic. The assumption that the signal is noisy is consistent with the psychological literature that suggests that introspection gives only limited information about own's motives.

In period 2, the decision X_2 yields payoffs to the agent equal to

$$U_2(X_2 = 1) = \mu \tag{3}$$

$$U_2(X_2 = 0) = 0 \tag{4}$$

The difference between period 1 and 2 is the presence of a reward in period 1. In this section, we assume, in line which much of the economics literature, that if the agent chooses $X_1 = 1$, he learns his taste, μ. Moreover, he will recall μ in period 2. If, in contrast, the agent chooses $X_1 = 0$, then he does not learn μ.

2.2 Solving the Model

To ensure a time-consistent solution, we solve the model by backward induction. So, first consider the agent's decision problem in period 2. Two cases should be distinguished. First, if $X_1 = 1$, then when deciding on X_2 the agent knows μ, and chooses $X_2 = 1$ if and only if $\mu \geq 0$. Second, if $X_1 = 0$, then $X_2 = 0$. The reason is that the incentive to choose $X_1 = 1$ is stronger in period 1 than in period 2.

Now turn to period 1. The agent anticipates his behavior in period 2. The agent's behavior in period 1 is characterized by a value of s_1, $\bar{s}$, for which he is indifferent between $X_1 = 0$ and $X_1 = 1$. For signal $s_1 = \bar{s}$, $X_1 = 1$ yields a payoff

$$p + \pi \bar{s} + (1 - \pi) \frac{1}{4} h \tag{5}$$

The third term of (5) gives the expected payoff in period 2. As the incentive to undertake the activity in period 1 is stronger than in period 2, $X_2 = 1$ requires that $\mu \geq 0$ and that s_1 is uninformative. In that case, the expected payoff for period 2 equals $\Pr(\mu > 0) E(\mu \mid \mu > 0) = \frac{1}{4} h$. Since $X_1 = 0$ yields a payoff equal to zero, the agent chooses $X_1 = 1$ if (5) is greater than zero:

$$s_1 \geq \bar{s} = -\frac{p}{\pi} - \frac{(1 - \pi)}{\pi} \frac{1}{4} h \tag{6}$$

The first term of (6) shows that a reward encourages the agent to undertake the task in period 1. The last term of (6) captures learning by doing. This term shows that the possibility of learning encourages the agent to undertake the activity in period 1 (see Prescott, 1972, and Grossman, Kihlstrom and Mirman, 1977). Notice that the higher is π, the smaller is the scope for learning by doing. Of course, the reason is that learning by introspection (through s_1) and learning by doing are substitutes.

What are the consequences of the reward for activity engagement in period 2? To answer this question, suppose that the uniform distribution on $[-h, h]$ captures the tastes of all agents in society. What is the proportion of agents, Π_e choosing $X_2 = 1$? At an aggregated level, half of the agents who received a correct signal will choose $X_2 = 1$. Moreover, agents for which $\mu > 0$ and $X_1 = 1$ will choose $X_2 = 1$. It is now easy to see that

$$\Pi_e = \frac{1}{2} \pi + \frac{1}{2} (1 - \pi) \frac{1}{2h} \left(h + \frac{p}{\pi} + \frac{(1 - \pi)}{\pi} \frac{1}{4} h \right) \tag{7}$$

Equation (7) shows that a temporary reward has a long-run effect. The reason is that encouraging an activity in the short-run leads to learning. More agents learn that they have a taste for the activity. The next proposition summarizes the above discussion.

Proposition 1 *In a learning by doing model, a temporary reward is a short-run as well as a long-run motivator.*

3 Self-perception Theory: The Myopic Case

3.1 The Model

In this section, I adjust the model of the previous section to formalize the SP theory developed by Bem (1972). An essential feature of Bem's theory is that agents form beliefs about themselves on the basis of (i) their past decisions, and (ii) the environment in which those decisions were made. To account for this feature, I assume that in period 2, the agent recalls neither s_1 nor μ. However, the agent does remember his decision X_1 and the reward p. Moreover, I assume that, as in period 1, in period 2 the agent receives a signal, s_2, about his taste for the activity. Again, with probability π, this signal is fully informative, $s_2 = \mu$, and with probability $1 - \pi$, this signal does not contain information about μ. Unless $s_1 = s_2$, s_1 and s_2 are independent of each other. Concerning period 1, the current model is identical to the model of the previous section.

One of the objectives of this section, is to explain the outcome of the experiment by Lepper, Greene and Nisbett (1973) that superfluous rewards may work counterproductive. In this experiment, the subjects were probably not aware of the multi-period setting (see below). In fact, in Bem's theory, the emphasis is put on how a person's self-perception is shaped by his past decisions. The emphasis is not on how a person's future self-perception depends on his current decisions. For this reason, I assume in this section that the agent is myopic. In period 1, he is unaware of the future. In the next section, I relax this assumption.

3.2 Solution of the Model

Consider period 1. The implication of the assumption that the agent is myopic is that he chooses $X_1 = 1$ if and only if

$$s_1 > -\frac{p}{\pi} \tag{8}$$

Throughout, we assume that $h \geq |-\frac{p}{\pi}|$. Equation (8) shows that the higher is the reward (p), the wider is the range of s_1 for which the agent chooses $X_1 = 1$.

Lemma 1 *In the SP model, a temporary reward is a short-run motivator.*

On the basis of the information the agent possesses in period 2, four situations can be distinguished.

1. $X_1 = 1$ and $s_2 > -\frac{p}{\pi}$. In this case, the signal s_2 and the decision $X_1 = 1$ are consistent. The expected value of μ, conditional on the available information, equals
$$E\left(\mu \mid X_1 = 1, s_2 > -\frac{p}{\pi}\right) = \pi s_2 + \pi (1-\pi) \frac{1}{2}\left(h - \frac{p}{\pi}\right) \tag{9}$$
The last term of (9) captures that the agent infers from his past decision that he had a positive attitude towards X in the past. The implication is a positive bias towards X in period 2. The magnitude of this bias depends on p. The reason is that the higher is p, the less information the past decision on X contains about μ. From (9), it directly follows that, conditional on $X_1 = 1$ and $s_2 > -\frac{p}{\pi}$, the agent chooses $X_2 = 1$ if
$$s_2 > -\frac{1}{2}(1-\pi)\left(h - \frac{p}{\pi}\right) \tag{10}$$
One can verify that for low values of p, $p < p_L = \frac{\pi(1-\pi)h}{3-\pi}$, condition (10) is redundant. That is, $s_2 > -\frac{p}{\pi}$ implies (10) for $p < p_L$.

2. $X_1 = 1$ and $s_2 \leq -\frac{p}{\pi}$. Now the signal s_2 is inconsistent with the agent's decision in period 1. Either his current signal is incorrect or his past perception of X was incorrect. Using that
$$E\left(\mu \mid X_1 = 1, s_2 \leq -\frac{p}{\pi}\right) = \pi(1-\pi) s_2 + \pi(1-\pi)\frac{1}{2}\left(h - \frac{p}{\pi}\right)$$
it easy to check that the agent chooses $X_2 = 1$ if
$$s_2 > -\frac{1}{2}\left(h - \frac{p}{\pi}\right) \tag{11}$$
It is worth noticing that (11) is less restrictive than (10). The implication is that a higher value of s_2 does not always make the decision $X_2 = 1$ more likely. Suppose, for example, that $p = \frac{1}{2}$, $\pi = \frac{1}{2}$, and $h = 4$. Then, in the ranges $s_2 \in \left[-1\frac{1}{2}, -1\right]$ and $s_2 \in \left[-\frac{3}{4}, 4\right]$, the agent chooses $X_2 = 1$. However, in the range $\left(-1, -\frac{3}{4}\right)$ he chooses $X_2 = 0$. To understand this finding, first suppose

that $s_2 = -\frac{7}{8}$. Then, the signal is consistent with the agent's decision in the first period. This increases the probability that $\mu = -\frac{7}{8}$ and that the agent will suffer from choosing $X_2 = 1$. Now suppose that $s_2 = -\frac{5}{4}$. In that case, s_2 and the past decision are not consistent. The agent is therefore inclined to suspect signal $s_2 = -\frac{5}{4}$. Accordingly, the agent puts lower weight to signal $s_2 = -\frac{5}{4}$ than to signal $s_2 = -\frac{7}{8}$.

Another implication of (11) is that for high values of p, $p \geq p_H = \frac{1}{3}\pi h$, $s_2 \leq -\frac{p}{\pi}$ excludes that (11) holds. For $p < p_H$, and given $X_1 = 1$ and $s_2 \leq -\frac{p}{\pi}$, a higher reward narrows the range for which the agent chooses $X_2 = 1$.

3. $X_1 = 0$ and $s_2 > -\frac{p}{\pi}$. As in the previous case the past decision and the current signal are not consistent. Using

$$E\left(\mu \mid X_1 = 0, s_2 > -\frac{p}{\pi}\right) = \pi (1-\pi) s_2 + \pi (1-\pi) \frac{1}{2}\left(-h - \frac{p}{\pi}\right)$$

one can show that the agent chooses $X_2 = 1$ if

$$s_2 > \frac{1}{2}\left(h + \frac{p}{\pi}\right) \tag{12}$$

The right-hand side of (12) is positive, implying that in period 2 the agent has a bias against the activity. The reason is that $X_1 = 0$ signals that the agent has no taste for activity X. This bias against X must be compensated with a favorable signal in period 2. As in the previous case, a rise in p makes $X_2 = 1$ less likely.

4. $X_1 = 0$ and $s_2 \leq -\frac{p}{\pi}$. The signal is consistent with the past decision. As

$$E\left(\mu \mid X_1 = 0, s_2 \leq -\frac{p}{\pi}\right) = \pi s_2 + \pi (1-\pi) \frac{1}{2}\left(-h - \frac{p}{\pi}\right)$$

The agent chooses $X_2 = 1$ if

$$s_2 > \frac{1}{2}(1-\pi)\left(h + \frac{p}{\pi}\right) \tag{13}$$

This case does not need much further discussion. As $s_2 \leq -\frac{p}{\pi}$ excludes that (13) holds, the agent always opts for $X_2 = 0$.

Equations (10-13) describe the agent's behavior in period 2. A striking finding is that, given the decision in period 1, for $p_L < p < p_H$ a rise in p narrows the range of s_2 for which the agent chooses $X_2 = 1$. This brings me to the following proposition.

Proposition 2 *Given the choice of X_1, the probability that $X_2 = 1$ is a non-increasing function of p.*

To grasp the intuition behind Proposition 2 recall that an important feature of the model of this section is that the agent does not remember the grounds for his choice in period 1. However, the agent knows that in period 1 the reward was a positive reinforcer. As a consequence, the reward affects the agent's evaluation of his past decision. Suppose, for instance, a high reward. Then, the agent attributes $X_1 = 1$ to the reward rather than to his attitude. Moreover, from $X_1 = 0$ the agent infers that his attitude must have been very negative.

Proposition 2 is consistent with the findings of experiments by social psychologist. For example, take the Lepper, Greene and Nisbett (1973) experiment. In this study, the experimenters introduced an attractive drawing activity into preschool classrooms during "free-play" periods. The new activity was one of the alternative activities. Children showing an initial interest for the new activity were selected for the experiment. The sample was therefore restricted to subjects for which $X_1 = 1$. In the experiment, the children were divided in three groups. In the first group, the children were asked to engage in the drawing activity in order to win an award (an extrinsic reward). In the second group, the children were asked to engage in the activity and could unexpectedly win a reward. Finally, in the third group, the children were simply asked to engage in the activity. Two weeks later, to measure the interest of the children in the drawing activity, it was again placed in the classrooms. The experimenters found that children of the first group showed a decrease in interest in the activity, whereas no changes were found for children of the two other groups. This pattern of results was found in many subsequent experiments (see Deci, Koestner and Ryan, 1999; Lepper, Henderlong and Gingras, 1999).

So far, I have focused on the effect of p on the decision on X_2, given the choice of X_1. Let me now incorporate the effect of p on X_2 through X_1 into the analysis. In Lemma 1 I have concluded that the effect of p on the probability that $X_1 = 1$ is positive. In period 2, the agent infers from his decision in period 1 information about his attitude towards the activity: $X_1 = 1$ suggests a favorable signal, while $X_1 = 0$ suggests an unfavorable signal. Therefore, as an increase in p encourages $X_1 = 1$, it also encourages $X_2 = 1$.

The upshot of the above discussion is that the total effect of p on the probability that $X_2 = 1$ can be decomposed into two parts. First, the effect of p on $X_2 = 1$, given X_1. As I have shown above Proposition 2, this effect is non-positive. Second, the effect of p on X_2 through the effect of p on X_1. This effect is positive. To establish the total long-run effect of p, three ranges of p should be distinguished.

First, $p \leq p_L$. As argued earlier, for low values of p the condition in (10) is redundant. The implication is that the effect of p on X_2, given X_1 is small. In the appendix, I show that for $p \leq p_L$, an increase in p has a positive effect on the probability that $X_2 = 1$. The reason for this result is that for low rewards an agent does not attribute his past decision to the reward. The main effect of p on X_2 is through X_1.

Second, $p_L < p < p_H$. If p is in this interval, then the effect of p on X_2, given X_1 is large. When analyzing his past decision, an agent takes seriously the possibility that the reward induced a positive decision. Of course, the higher was the reward, the more the agent is inclined to attribute a (past) positive decision to the reward. In the appendix, I show that for $p_L < p < p_H$ an increase in p has a negative effect on the probability that the agent chooses $X_2 = 1$.

Third, $p > p_H$. In this case, the effect of p on $X_2 = 1$, given $X_1 = 1$, is limited because (11) is redundant. For high values of p, the past decision does not contain much information on the agent's taste for the activity. It was the reward that drove the decision on X_1. If $p = \pi h$, then the decision on X_2 is made independent of what happened in period 1.

Proposition 3 *Let $Z(p)$ denote the probability that the agent chooses $X_2 = 1$. The effect of the reward on Z is for $p \leq p_L$, $\frac{dZ}{dp} = \frac{\pi}{2h} > 0$; for $p_L < p < p_H$, $\frac{dZ}{dp} = \frac{1}{4h^2\pi^2}\left(\left(3-\pi\right)\left(\pi^2-1\right)p + h\pi\left(\pi^3-1\right)\right) < 0$; for $p > p_H$, $\frac{dZ}{dp} = \frac{1}{4h^2\pi}\left(p\left(1-\pi^2\right) + h\pi^2\left(\pi-1\right)\right)$. Moreover, $Z(0) = Z(\pi h) = \frac{1}{2}$.*

Proof: Appendix.

Figure 1 depicts the effect of p on the probability that $X_2 = 1$ for the case that $h = 4$ and $\pi = \frac{1}{2}$. A small reward increases the probability of $X_2 = 1$. This

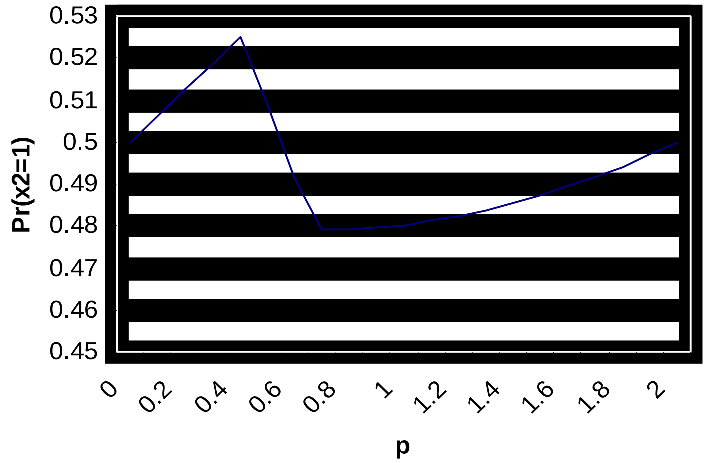


Figure 1: The relationship between p and Pr(X2=1)

probability rises until $p = 0.4$. For $0.4 < p < \frac{2}{3}$, the probability of $X_2 = 1$ decreases in p. Notice that if p is sufficiently large, then the probability that $X_2 = 1$ is smaller than $\frac{1}{2}$. The long-rung effect of the reward is negative. Finally, for $p \geq \frac{2}{3}$ a rise in p increases the probability that $X_2 = 1$.

4 Self-perception Theory: the time-consistent case

4.1 The Model

The model is identical to the model of the previous section, except that in the present model the agent is aware of the future. That is, when making a decision on X_1 he realizes that this decision will affect his decision in the next period.

4.2 Solution

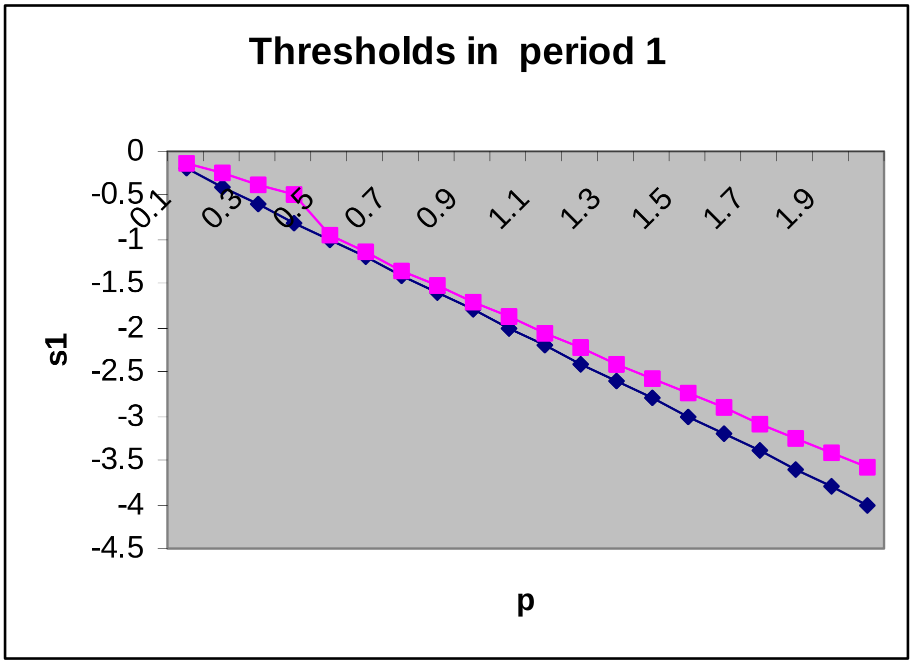
We first consider period 2, assuming that in period 1 the agent chose $X_1 = 1$ if and only if $s_1 \geq s$. Analogous to the previous section, four cases can be distinguished. By replacing $-\frac{p}{\pi}$ with s in (10-13), one gets the decision the agent takes in each case. Thus, the agent chooses $X_2 = 1$ in case 1 if $s_2 > -\frac{1}{2}(1-\pi)(h+s)$; in case 2 if $s_2 > -\frac{1}{2}(h+s)$; in case 3 if $s_2 > \frac{1}{2}(h-s)$; and in case 4 if $s_2 > \frac{1}{2}(1-\pi)(h-s)$.

Now consider period 1. As a benchmark assume that $p = 0$. Is it optimal for the agent to choose $X_1 = 1$ if and only if $s_1 \geq 0$? Suppose $s_1 = 0$. The cost of $X_1 = 1$ is an overestimation of s_1 in period 2: $E(s_1 \mid X_1 = 1, s = 0) = \frac{1}{2}h$. The problem is that if the agent observes $0 < s_2 < -\frac{1}{2}h$, he chooses $X_2 = 1$. By contrast, the cost of $X_1 = 0$ is an underestimation of s_1 in period 2: $E(\mu \mid X_1 = 0, s = 0) = -\frac{1}{2}h$. In this situation, the agent chooses $X_2 = 0$ if $\frac{1}{2}h > s_2 > 0$. Because of symmetry and risk-neutrality, the benefit of choosing $X_1 = 1$ rather than $X_1 = 0$ if $s_1 = 0$ is equal to the cost. Hence, for decision-making in period 2, $s = 0$ is optimal.

Now suppose that $p > 0$. A direct implication is that $X_1 = 1$ becomes more appealing. From a period 1 perspective, the agent would choose $X_1 = 1$ if and only if $s_1 \geq -\frac{p}{\pi}$. How would such a strategy affect the agent's decision in period 2. Assume that $p \in [0, p_L]$, with $p_L = \frac{\pi(1-\pi)}{3-\pi}$. As discussed before, this assumption about p implies that if $X_1 = 1$, then for $s_2 \geq -\frac{p}{\pi}$ (10) is satisfied. Now suppose that $s_1 = -\frac{p}{\pi}$. The agent anticipates in period 1 that he would suffer from $X_1 = 1$ in period 2 if his signals were correct: $X_1 = 1$ implies $X_2 = 1$ if $s_1 = s_2 = -\frac{p}{\pi}$. This is the first cost of choosing $X_1 = 1$ rather than $X_1 = 0$. Moreover, $X_1 = 1$ implies that the agent is expected to be overly optimistic about s_1 in period 2: $E(s_1 \mid X_1 = 1) = \frac{1}{2}\left(h - \frac{p}{\pi}\right) > -\frac{p}{\pi}$. This is the second cost of choosing $X_1 = 1$ rather than $X_1 = 0$. By choosing $X_1 = 0$ rather than $X_1 = 1$ when $s_1 = -\frac{p}{\pi}$, the agent would underestimate s_1 in period 2: $E(s_1 \mid X_1 = 0) = -\frac{1}{2}\left(h + \frac{p}{\pi}\right)$. As $-\frac{1}{2}\left(h + \frac{p}{\pi}\right)$ is closer to $-\frac{p}{\pi}$ than $\frac{1}{2}\left(h - \frac{p}{\pi}\right)$, this cost of choosing $X_1 = 0$ is smaller than the second cost of choosing $X_1 = 1$. A comparison of the various costs shows that choosing $X_1 = 1$ if and only if $s_1 \geq -\frac{p}{\pi}$ cannot be part of an equilibrium. If $s_1 = -\frac{p}{\pi}$, then the agent strictly prefers $X_1 = 0$ to $X_1 = 1$ in order to avoid an overly optimistic perception of s_1 in period 2. The implication is that the agent is indifferent between $X_1 = 0$ and $X_1 = 1$ for $0 < s < -\frac{p}{\pi}$.

Now suppose that $p > p_L$. Then, the first cost of $X_1 = 1$ when $s_1 = -\frac{p}{\pi}$ is not present anymore, as for $s_2 = s$, the agent chooses $X_2 = 0$ (see 10). The second cost of $X_1 = 1$ remains. Consequently, the agent is still inclined to choose $X_1 = 0$ when $s_1 = -\frac{p}{\pi}$ although $X_1 = 1$ is optimal from a period 1 perspective. This inclination, however, is weaker for $p > p_L$ than for $p \leq p_L$.

Figure 1 depicts the values of s_1 for different values of p. The straight line is simply $-\frac{p}{\pi}$ (with $\pi = \frac{1}{2}$). The line above this line give the thresholds for different



values of p when the agent is forward-looking. The agent is less inclined to choose $X_1 = 1$. After the kink (around $p = .5$), (10) is not always satisfied anymore.

The above discussion leads to the following proposition.

Proposition 4 *Suppose that in the psychological model the agent is forward-looking. Then, in period 1 the agent responds less to a reward in order to avoid an overly optimistic view in period 2.*

Proposition 4 explains why decisionmakers in organizations are often afraid of "creating precedents." Creating a precident refers to a situation in which a decision on a case or person today may affect a decision on a similar case or decision tomorrow. For example, a request of a person to work home may be rejected on account of a fear that permission may affect decisions on future requests. In the above model, the agent may choose $X_1 = 0$ because he fears that $X_1 = 1$ gives too strong an incentive to choose $X_2 = 1$ later.

5 Discussion

In the economics literature, the stereotypical decision maker is a person who knows himself. Uncertainty may exist, but it relates to the environment rather than to a person himself. Social psychologists take a different perspective. They assume that a person holds a view of himself, but that this view is not always correct.

In this paper, we have adopted the view of social psychologists. We have analyzed two models of how a person may learn about his preferences concerning an activity. The first model is based on the dynamic learning DL model developed by Grossman et al. (1977). In the DL model, an agent learns his preferences by doing. The second model formalizes Bem's self-perception SP theory (Bem, 1972). Recently, this theory has attracted the attention of economists. Especially, its prediction that extrinsic motivation may crowd out intrinsic motivation.

Our analysis of the two models shows that both predict that a temporary reward is a short-run motivator. However, the predictions of the model with respect to the long-run diverge. In the DL model, a temporary reward is also a long-run motivator, but in the SP model the effect of a temporary reward is ambiguous. Small rewards are long-run motivators, but large rewards may be counterproductive. A comparison between the assumptions underlying the two models shows that they differ in one important respect: in the DL model the agent remembers his payoff, his decision on the activity and the reward, while in the SP model the agent does not recall his payoff nor his past attitude towards the activity at stake. The difference between the assumptions underlying the models can be used to establish for which environment each model is relevant.

As an illustration of an environment, let me discuss a recent field experiment conducted by Charness and Gneezy (2006). In this experiment, a group of 120 students were divided into three groups of equal size. The first group was the control group. A student in the second group would receive \$25 if he visited the gym at least once during that week. A student in the third group would receive \$100 if he attended the gym at least eight times during the next four weeks. The experimenters kept records of the students' visits to the gym during the year. Notice that the set-up of this experiment fits well with the two models discussed in the previous sections. Crucial is that a distinction can be made between the short-run and long-run effects of a temporary reward. The results of the experiment are fully consistent with the predictions of the DL model. The reward appeared to be a clear short-run motivator. Moreover, the higher was the reward, the larger was its short-run effect. Once the incentives were removed, the visits to the gym declined but the average number of visits to the gym was higher than the initial average number of visits. So, the rewards had also a positive long-run effect.[4]

One of the objectives of the gym experiment was to test a possible prediction of the SP model that a reward crowds out intrinsic motivation. As argued above, this prediction requires that when deciding whether or not to visit the gym a student does not recall how much he liked (or disliked) the gym the last time. As the time spans between visits to the gym are usually not very long, it seems fairly unlikely that this requirement is met. Therefore, it is not very surprising that the experimenters did not find that in the gym context extrinsic motivation crowds out intrinsic motivation.

As an illustration of another environment, let me now return to "Earning by Learning" programs discussed in the introduction. Is it possible that giving money or pizza's for reading books to children may have negative long-run effects? Again, the crucial question is what do the children remember later? If the reward has induced a child to read four books every week for years, it seems highly implausible that this child forgets how much he likes reading. So let us focus on a person who stopped reading for a while (for example, because of a study or raising children), and considers the possibility of reading books again. In that case, it is very well possible that the person does not recall his past attitude towards reading but does remember the money or the pizza. Especially, a pizza as a reward for reading is somewhat remarkable. So, we cannot exclude that the SP model is the appropriate model for describing the long-run effects of "Earning by Learning" programs for some children.

As the SP model requires sufficient time between an agent's decisions - otherwise the assumption of imperfect recall is implausible - the model seems to be most relevant in educational settings. Indeed it is not very hard to imagine situations in which rewarding or punishing childrer for certain kinds of behavior might backfire. Pushing your children to go to church may eventually discourage them to go to church. Likewise, rewarding your children to save money may have adverse consequences for their saving behavior when they are grown up.[5]

---

[4] These results do not conflict with the prediction so the SP model. Also in that model, both the short-run and long run effects of a reward can be positive.

[5] I thank Hein Roelfsema for pointing to these examples.

Apart from an educational context, are there other contexts for which the SP model could be relevant? It is hard to come up with many contexts for which the SP model is relevant when we stick to a strict version of the model. Things become different when we allow for situations in which the decision maker in the past is not the same person as the decision maker today. Public policy is a point in case. Stiglitz (2000) suggests that an analysis of a public program should begin by investigating its history and the circumstances under which it arose. As in the SP model, decision makers today can learn from decisions made in the past. As an example, Stiglitz discusses the U.S. social security program that was launched in the midst of the Great Depression. In that time, Stiglitz argued, the need for insurance was that urgent that society had to make some provision for it. In the Netherlands, economists tend to refer to the mid-fifties rather than to the thirties when discussing the history of the Dutch social security program (Wolfson, 1988). Instead of an urgent need for insurance, a fair income distribution and solidarity were emphasized as rationales for social security. The SP model shows that the environment in which past decisions were made may affect current decisions. Moreover, the requirement that a person does not know the attitude of the person who made the past decision is more likely to be met when the two persons are not the same. Of course, similar situations exist in the private sector. A new manager of a firm may observe past decisions. Moreover, if he has worked in the same sector before, he may know the environment in which past decisions were made. However, he may have less information about the characteristics of the firm itself. In such a situation, the new manager may attribute a past decision to the environment in which the firm operated, while actually the decision was more inspired by some comparative advantage of the firm.

The upshot of the above discussion is that the requirements under which a strict version of the SP model is relevant are rarely met except in educational settings. This does not make the SP model useless for economists, however. In many organizations, decisions today and the environment in which these decisions are made will affect decisions in the future. As today's decision makers will be replaced by other ones, future decision makers may have more information about the environment under which decisions were made than about the exact considerations of current decision makers.

6 References

Aronson, E., T.D. Wilson, and R.M. Akert, 2005. Social Psychology. Englewood Cliffs: Prentice Hall, 5th edition.

Bem, D.J., 1972. Self-Perception Theory. In: L. Berkowitz (Ed.), Advances in experimental social psychology Vol. 6. New York: Academic Press.

Bénabou, R., and J. Tirole, 2003. Intrinsic and Extrinsic Motivation. Review of Economic Studies, 70, 489-520.

Bénabou, R., and J. Tirole, forthcoming. Incentives and Prosocial Behavior. American Economic Review.

Brown, R.G., 1974. Sales Response to Promotions and Advertising. Journal of Advertising Research, 14, 33-39.

Charness, G. and U. Gneezy, 2006. Incentives and Habits, Mimeo.

Deci, E.L., 1971. Effects of Externally Mediated Rewards on Intrinsic Motivation. Journal of Personality and Social Psychology 18, 105-115.

Ehrenberg, A.S.C., K. Hammond and G.J. Goodhardt, 1994. The After-Effects of Price-Related Consumer Promotions. Journal of Advertising Research, 34, 11-21.

Ellingsen, T., and Magnus Johannesson, 2005, Trust as an Incentive, mimeo.

Frey, B.S., 1997, Not just for the Money. An Economic Theory of Personal Motivation. Cheltenham. Edward Elgar.

Frey, B.S., and F. Oberholzer-Gee, 1997, The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. American Economic Review, 87, 746-755.

Gneezy U. and A. Rustichini, 2000, A Fine is a Prize, Journal of Legal Studies, 29, 1-17.

Grossman, Sanford J., Richard E. Kihlstrom, and Leonard J. Mirman, 1977. A Bayesian Approach to the Production of Information and Learning By Doing. Review of Economic Studies 44(138), 533-547.

Janssen, M.C.W., and E. Mendys-Kamphorst, 2004. The Price of a Price: On the Crowding Out of Social Norms. Journal of Economic Behavior and Organization, 55, 377-395.

Jobber, D., 2004. Principles and Practice of Marketing, McGraw-Hill.

Kruglanski, Arie W., Irith Friedman, and Gabriella Zeevi, 1971. The Effects of Extrinsic Incentive on Some Qualitative Aspects of Task Performance. Journal of Personality 39(4), 606-617.

Lepper, M.R., and and David Greene, 1978, Overjustification Research and Beyond: Towrda a Means-Ends Analysis of Intrinsic and Extrinsic Motivation, in: The Hidden Cost of Reward: New Perspectives on the Psychology of Human Motivation, New York: John Willey.

Lepper, M.R., D. Greene, and R.E. Nisbett, 1973. Undermining Children's Intrinsic Interest with Extrinsic Reward. Journal of Personality and Social Psychology 28, 129–37.

Prescott, E.C., 1972. The Multi-Period Control Problem Under Uncertainty. Econometrica, 40, 1043-1058.

Rothschild, M.L. and W.C. Gaidis, 1981. Behavioral Learning Theory: Its Relevance To Marketing and Promotions. Journal of Marketing, 45, 70-78.

Stiglitz, J.E., 2000. Economics of the Public Sector, Norton.

Wolfson, D.J., 1988, Publieke Sector en Economische Orde, Wolters-Noordhoff.

7 Appendix

Proof of Proposition 3

Define: $a = -\frac{p}{\pi}$

$$A = \Pr(\mu > -\tfrac{p}{\pi}) = \int_{\left(-\frac{w}{p}\right)}^{h} \frac{1}{2h} du$$

$$B = \Pr\left(X_1 = 1 \mid \mu > -\tfrac{p}{\pi}\right) = p + (1-p)\int_{\left(-\frac{w}{p}\right)}^{h} \frac{1}{2h} du,$$

$$C = \Pr\left(X_2 = 1 \mid A \wedge B\right) = p\int_{-\frac{1}{2}(1-p)\left(h+\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{h-\left(-\frac{w}{p}\right)} du$$

$$+ (1-p)\left(\int_{-\frac{1}{2}(1-p)\left(h+\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{2h} du + \int_{-\frac{1}{2}\left(h+\left(-\frac{w}{p}\right)\right)}^{\left(-\frac{w}{p}\right)} \frac{1}{2h} du\right)$$

$$D = \Pr\left(X_2 = 1 \mid A \wedge \not{B}\right) = p \int_{\frac{1}{2}\left(h-\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{h-\left(-\frac{w}{p}\right)} du + (1-p) \int_{\frac{1}{2}\left(h-\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{2h} du$$

$$E = \Pr\left(s_2 > \frac{1}{2}(1-\pi)(h-a) \mid X_1 = 0\right) = (1-p)\left(\int_{-h}^{a} \frac{1}{2h} du\right)$$

$$\left(p \int_{\frac{1}{2}(h-a)}^{h} \frac{1}{h-a} du + (1-p) \int_{\frac{1}{2}(h-a)}^{h} \frac{1}{2h} du\right)$$

$$F = \Pr(X_1 = 1 \mid \mu \leq -\tfrac{p}{\pi}) = (1-p) \int_{\left(-\frac{w}{p}\right)}^{h} \frac{1}{2h} du$$

$$G = \Pr\left(X_2 = 1 \mid \not{A} \wedge B\right) = p \int_{-\frac{1}{2}\left(h+\left(-\frac{w}{p}\right)\right)}^{\left(-\frac{w}{p}\right)} \frac{1}{h+\left(-\frac{w}{p}\right)} du +$$

$$(1-p)\left(\int_{-\frac{1}{2}(1-p)\left(h+\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{2h} du + \int_{-\frac{1}{2}\left(h+\left(-\frac{w}{p}\right)\right)}^{\left(-\frac{w}{p}\right)} \frac{1}{2h} du\right)$$

$$H = \Pr\left(X_1 = 1 \mid \not{A} \wedge \not{B}\right) = (1-p) \int_{\frac{1}{2}\left(h-\left(-\frac{w}{p}\right)\right)}^{h} \frac{1}{2h} du$$

Total expected utility can now be written as:

$$U = A(BC + (1-B)D) + (1-A)(FG + (1-F)H)$$

The effect of p on $\Pr(X_2 = 1)$ if $p < p_L$

If $p < p_L$, then in C and G, $-\frac{1}{2}(1-p)\left(h + \left(-\frac{w}{p}\right)\right) = (-\frac{w}{p})$, implying $U = \frac{1}{4h^2 p}\left(h + 2w - 2pw + hp^2 + 2pw - 2w - h + hp^2\right) = \frac{1}{2h} p$.

The effect of p on $\Pr(X_2 = 1)$ if $p_L < p < p_H$

Expected utility equals

$\frac{1}{8h^2p^2}\left(pw^2 - 3w^2 - 2hpw + 2hp^4 w + 5h^2p^2 - h^2p^3 + h^2p^4 - h^2p^5 + 3p^2w^2 - p^3w^2\right)$.

Differentiating with respect to w equals

$\frac{\partial\left(\frac{1}{8h^2p^2}\left(pw^2 - 3w^2 - 2hpw + 2hp^4 w + 5h^2p^2 - h^2p^3 + h^2p^4 - h^2p^5 + 3p^2w^2 - p^3w^2\right)\right)}{\partial w} =$

$\frac{1}{4h^2p^2}\left(pw - hp - 3w + hp^4 + 3p^2w - p^3w\right) =$

$\frac{1}{4h^2p^2}\left((3-p)(p^2-1)w + hp(p^3-1)\right) < 0$.

The effect of p on $\Pr(X_2 = 1)$ if $p > p_H$

If $p > p_H$, then in C and G, $-\frac{1}{2}\left(h + \left(-\frac{w}{p}\right)\right) = (-\frac{w}{p})$, implying that total expected utility equals

$\frac{1}{8h^2p}\left(w^2 + 4h^2p - 2hp^2w + 2hp^3w - h^2p^2 + 2h^2p^3 - h^2p^4 - p^2w^2\right)$.

Differentiating with respect to w yields

$\frac{\partial\left(\frac{1}{8h^2p}\left(w^2 + 4h^2p - 2hp^2w + 2hp^3w - h^2p^2 + 2h^2p^3 - h^2p^4 - p^2w^2\right)\right)}{\partial w} =$

$\frac{1}{4h^2p}\left(w - hp^2 + hp^3 - p^2w\right) = \frac{1}{4h^2p}\left(w(1-p^2) + hp^2(p-1)\right)$.