

Is the evidence for hyperbolic discounting in humans just an experimental artefact?

Glenn W. Harrison^a and Morten Igel Lau^b

^aDepartment of Economics, College of Business Administration, University of Central Florida, Orlando, Florida 32816-1400; ^bCentre for Economic and Business Research, Copenhagen Business School, DK-2000 Frederiksberg, Denmark. Glenn.Harrison@bus.ucf.edu mol@cebr.dk
http://www.bus.ucf.edu/gharrison http://www.cebr.dk/mol

Abstract: We question the behavioral premise underlying Ainslie's claims about hyperbolic discounting theory. The alleged evidence for humans can be easily explained as an artefact of experimental procedures that do not control for the credibility of payment over different time horizons. In appropriately controlled and financially motivated settings, human behavior is consistent with conventional exponential preferences.

Ainslie's (2001) book, *Breakdown of Will*, is based on hyperbolic discounting theory. This theory predicts that the individual could behave in a dynamically inconsistent manner, by holding and acting on preferences at one point in time that contradict the preferences of the same individual at a later date. However, before worrying about ways that the individual could address possible dynamic inconsistencies, we need to be sure that the behavioral premise is valid.

A critical design feature in the empirical literature on hyperbolic discounting is the use of a time delay to the *early* payment option in order to control for any confounding effects from fixed premia due to transactions costs. The use of this front end delay (FED) means that one cannot differentiate between "quasi-hyperbolic preferences" and "exponential preferences," and we do not believe that any credible design can do so.

Ainslie concludes his discussion of the empirical evidence on hyperbolic discounting with the following passage:

There is extensive evidence that both people and lower animals spontaneously value future vents in inverse proportion to their expected delays. The resulting hyperbolic discount curve is seen over all time ranges, from seconds to decades. Because a hyperbolic curve is more bowed than the exponential curve that most utility theories go by, it describes a preference pattern that these theories would call irrational: It predicts temporary preferences for the poorer but earlier of the two alternative goals during the time right before the poorer alternative becomes available. (Ainslie 2001, p. 47)

This passage confounds three things. The first is whether the discount rate varies with the length of the time horizon over which it is being elicited, such as it does with continuously hyperbolic preferences. The second is whether the discount rate for a given horizon and elicited with a FED is different than the discount rate for the same horizon and elicited with no FED. For an experimenter, and for subjects evaluating the credibility of being paid, these are very different questions. The potential importance of this distinction seems to have been first noticed by Benzion et al. (1989).¹ It was also highlighted by Roberts (1991, p. 344), in the context of comments on Ainslie and Haendel (1983) and Winston and Woodbury (1991). The third issue is whether nonexponential preferences imply dynamic inconsistency when one relaxes the restrictive assumption of temporally separable preferences (Machina 1989; McClelland 1990).

The FED design was introduced into discount rate experiments to address concerns about differential credibility. Although it may not completely solve the potential credibility problem, it arguably mitigates it. The FED also serves to equalize any other unspecified differences subjects may perceive between the two payment options. For example, if subjects have a "passion for the present," they demand a premium in order to accept a delay of any length. In a choice between immediate payment and delayed payment, this premium is attached only to the delayed payment. Thus, the subject is being asked to compare "good apples today" with "bad apples tomorrow," confounding the discount rate with the credibility of receiving the commodity. However, if both payments are

delayed, the premium applies to both choices and thus becomes irrelevant to a choice between them. Harrison et al. (2002) used a FED in a major field experiment in Denmark, and found that elicited discount rates are proximately invariant with respect to horizon.

There are, however, many field settings in which the relevant issue is what the discount rate is for "money today" versus "money in the future."² Even if the experimenter faces the inferential problem of having to then tease apart the effects of time horizon from credibility, transactions, or other subjective costs, it is entirely appropriate that experiments with no FED be considered. If there is a finding that discount rates are not constant when there is no FED, then it is a matter for interpretation as to whether this is a subjective differential cost effect or a time-inconsistency effect (or both).

Evidence for the behavioral importance of a 30-day FED was provided by Coller and Williams (1999). In one of their experimental treatments they had no such delay, and the results from those experiments can be directly compared to their other experiments. After some minor modifications to their statistical analysis, Coller and Williams's results provide evidence that the use of a FED decreases elicited rates by a large amount. The average effect of having no FED is to increase elicited rates by 28 percentage points, with a 95% confidence interval between 52 percentage points and 3 percentage points. Coller et al. (2003) provide additional laboratory evidence on the role of the FED, and show that a 7-day FED is sufficient to overcome the effects of subjective transactions costs.

Finally, there have been no direct tests of the implication of dynamically inconsistent choice behavior using real rewards. Such longitudinal tests require that one allow for possible changes in the states of nature that the subject faces, since they may confound any in-sample comparisons of discount rate functions at different points in time. Harrison et al. (2005) have reported the results of a large-scale panel experiment undertaken in the field to examine this issue and found evidence strikingly consistent with dynamic consistency.

ACKNOWLEDGMENT

We thank the Danish Social Science Research Council for research support under project No. 24-02-0124.

NOTES

1. Holcomb and Nelson (1992) reexamined the role of a FED with monetary payoffs, motivated by a concern that Benzion et al. (1989) only studied hypothetical choices. Their FED was only one day long, so it is not obvious that the subjects viewed this as substantially different from there being no FED. They observed no apparent effect of the one-day FED on behavior.

2. Such settings might include individual decisions of whether to consume now or save for future consumption, or to purchase a more expensive but energy efficient appliance. We believe that individual decisions involving more significant sums of money or public policy decisions are better characterized as having a FED.

Shaping your past selves

Jeanne Peijnenburg

Faculty of Philosophy, University of Groningen, 9712 GL Groningen, The Netherlands. Jeanne.Peijnenburg@rug.nl

Abstract: I propose to complement Ainslie's idea of "bargaining with your future selves" with that of "shaping your past selves." The result of such a complementation is that an action can work in two ways: (1) as a precedent for future behavior and (2) as a shaper of past behavior. I argue that this diminishes the unwanted effects of hyperbolic discounting even further.

Weakness of will, or akrasia, comes in two different forms. Broad, apparent, or diachronic akrasia covers cases where an agent fails

to stand by a previous decision about what he will do; strict, clear-eyed, or synchronic akrasia comprises actions that go against an overall judgement that the agent still considers the best at the time of the action. Broad akrasia constitutes the easy problem: How does one explain that an agent changes his mind? Strict akrasia presents the hard problem: How does one understand that an agent believes at time t_1 that action A is the best, all things considered, and yet performs not- A at t_1 ? Some have given short shrift to the hard problem by declaring that strict akrasia is an illusion (Socrates, notably); others have tried hard to solve it (e.g., Donald Davidson, whose seminal work [Davidson 1969] spawned dozens of papers and books on the subject).

Be that as it may, Ainslie's (2001) book deals with the easy problem (where "easy" should of course be read tongue-in-cheek). For Ainslie locates akrasia in reversals of preference that occur whenever an agent comes close to a tempting, lesser reward. Ainslie's explanation of this phenomenon is very original in that it is based on the idea that broad akrasia is the rule, whereas its opposite – enkrateia or strength of will – is the exception. Hence, the problem is not why people do not stick to their guns, but rather, why they often do. Ainslie's solution to *that* problem lies in the view of an agent, P , as being a collection of agents P_1, P_2 , and so on, at different times, t_1, t_2 , and so on. These P_1, P_2, \dots have different and often competing interests, but there are also interests that they all have in common. By cleverly bargaining together in the intrapersonal version of a repeated Prisoner's Dilemma, they might succeed in letting the common interests prevail, thereby accounting for P 's strong will.

Ainslie's explanation of enkrateia is ingenious, and if couched in less technical jargon, it might well become a useful therapeutic instrument. However, it has a questionable implication as well. For it presupposes that an earlier P_j must be interested in a later P_k , and if P is a pitiable alcoholic, this presupposition is doubtful. The hallmark of an addict suffering from weakness of will is that he does not care how he will feel tomorrow or next year. (Ainslie denies this on pp. 17–18, where he argues that a "rational addict" *does* care about the future, because she "wouldn't even try to kick her habit." What Ainslie means, of course, is that she would not even try to kick her habit *now*. But this illustrates, contrary to what Ainslie suggests, precisely her *carelessness* about the future. In the usual sense of "caring about the future" the agent is able to see and reason further than the present moment – something that an addict *qua* addict is unable to do.)

If an akrates were really to care about the future (in the usual sense), the first step towards his recovery would have been taken. Ainslie's happy thought is to model this first step as a decision that functions as a precedent for future decisions, and hence keeps pace with a "personal rule" that, if followed, will generate a greater reward in the end. Nonetheless, each P_i can still fall prey to hyperbolic discounting by choosing the earlier, sooner reward over the larger, later one; and if he does, he can always logically claim that this was a special case and not a violation of the general rule. However, I think we can make this problem less pressing.

Imagine that I am a happily married mother of five. One day I go to a party, where I dance and drink exuberantly, only to wake up the following morning in a hotel bed next to an attractive man whom I cannot recall having seen before. Although it seems all too clear what happened, I still have some latitude in determining what I have already done. In particular, I can make it the case, through my future actions, that this adventure becomes either a mere incident or the beginning of a long and secret affair.

This example shows that sometimes I can, to a certain extent, determine my past actions. Moreover, my knowledge of the fact that I have this possibility, and hence my understanding that I am at a bifurcation point, might motivate me to pursue the one rather than the other course. Thus, we have here another way of evading the effect of hyperbolic discounting. For if choosing the larger, later reward (continue a happy family life) simultaneously means determining a past action (make my adventure a mere incident), then the smaller, sooner reward (date the attractive stranger again) loses much of its temptation. The reason for this is, of course, that

the shaping in retrospect of a past action is already very rewarding in itself. Similarly, when an alcoholic realizes that, through his future actions, he can make a recent lapse become an exception rather than a precedent for his future behavior, he might feel relieved. Very likely this knowledge will diminish his feelings of fatalism and hopelessness, and make him more motivated to contemplate bargaining with his future selves in order to obtain the larger, later reward.

I therefore propose that Ainslie's idea of "bargaining with your future selves" should be complemented with the idea of "shaping your past selves." The result of such a complementation is that an action can work in two ways at the same time, that is, as a precedent for future behavior and as a shaper of past behavior. This means, to use Ainslie's terms, that the behavior in question is not pushed, but pulled (pp. 19, 69). However, it is now pulled more strongly, for two forces are operating simultaneously. In Ainslie's metaphor, a future reward is pulling my present behavior into the future. To this I have added the metaphor of a current reward that is pulling my past behavior into the present. The resultant force is greater than either of its components, and it may well recruit strengthened motivation (cf. Peijnenburg 2004; forthcoming).

Problems with internalization

Howard Rachlin

Psychology Department, State University of New York at Stony Brook, Stony Brook, NY 11794-2500. howard.rachlin@sunysb.edu

Abstract: Ainslie's *Breakdown of Will* contains important insights into real world self-control problems, but it loses testability to the extent that it internalizes concepts whose meaning lies in overt behavior and its consequences.

Most psychologists who think about self-control tend to stop when they have postulated two forces: a primary impulsive tendency to consume an immediate reinforcer, and a more far-seeing tendency ("the will," in Ainslie's terms) to resist such consumption when it interferes with long-term goals. *Breakdown of Will* (Ainslie 2001) shows conclusively that such a two-force conceptual scheme is totally insufficient to describe almost any real-life motivational dilemma. In our society of plenty, the far-seeing tendency itself needs to be controlled. Otherwise, as Ainslie clearly points out, we will be just as badly off as we would have been if we simply gave in to all our impulses in the first place; indeed, we might be worse off. This fascinating book contains a rich analysis of human motivation and many deep and insightful descriptions of motivational dilemmas.

Having said this, it might sound churlish to complain. Yet I do. Although Ainslie takes care to relate the phenomena he discusses to hyperbolic discounting – a fundamentally behavioral conception – he tends to treat hyperbolic discounting itself as an internal, nonbehavioral (or at least non-overtly behavioral) process. Consequently, some of the discussion takes the form of a literary essay (albeit finely wrought), rather than a scientific analysis. (See particularly the discussion of indirection, pp. 187–96.)

At the root of this problem is Ainslie's attitude towards mental life in general; it is not behavioral enough. (I daresay most of the other commentaries will complain that it is too behavioral.) There is a paucity of empirical research described or cited and few suggestions about how such research could be conducted, especially in the later chapters. Instead, an internal arena is imagined with behaviors, discriminative stimuli, and rewards – all concepts originally constructed to describe the interaction of the behavior of whole organisms with their environments – interacting and competing over time. This internalization of fundamentally external concepts forces Ainslie to resort to internal "thought experiments" like Newcomb's problem (p. 134), rather than real experiments, as evidence for his theory.