lier payoff or raising the later one could resolve the preference reversal in favor of the later behavior.

A second, perhaps complementary, mechanism might rely on the self-blame, process-oriented component of regret. We have shown in recent work (Reb 2005) that making regret salient to experimental participants can lead them to use more careful decision processes, acquire more decision-relevant information, and deliberate longer before deciding. Perhaps the participants in Richard et al.'s study responded to the regret-salience manipulation by searching more diligently for alternative choices, weighing the costs and benefits of the unsafe behavior more carefully, or considering one of the familiar self-control strategies discussed by Ainslie (p. 73ff).

The hypothesis, then, is that regret can be more than a mere symptom of failed decision making. Regret may, in some circumstances, play a role in improving decisions: the experience of regret can drive learning in repeated decisions; its anticipation can shape single decisions. Outcome regret affects decisions by modifying the relative attractiveness of different payoffs. Process-related regret does so by motivating the search for trap-evading strategies such as decision bundling, precommitment, and the like. In both cases, the interweaving of thought and feeling that characterize regret provide the bridge between System 1 and System 2 processes, between the thoughtful appraisal of the distant goal and the visceral appeal of the immediate indulgence. Without venturing into evolutionary speculation that regret may have developed to serve such a system-bridging purpose, it is not difficult to see that some level of regret can be highly functional for control of the self-defeating processes that temporary preference problems represent. The hypothesis seems to us worthy of serious consideration.

The will: Interpersonal bargaining versus intrapersonal prediction

Luca Ferrero

Department of Philosophy, University of Wisconsin–Milwaukee, Milwaukee, WI 53201-0413. **ferrero@uwm.edu** **http://www.uwm.edu/~ferrero**

Abstract: Ainslie is correct in arguing that the force of commitments partly depends on the predictive role of present action, but this claim can be supported independently of the analogy with interpersonal bargaining. No matter whether we conceive of the parties involved in the bargaining as interests or transient selves, the picture of the will as a competitive interaction among these parties is unconvincing.

I am unpersuaded by Ainslie's central claim that the will is the product of transtemporal bargaining among successively dominant, transient interests analogous to the emergence of cooperation in a repeated Prisoner's Dilemma (Ainslie 2001, pp. 90–93). It is questionable that we could make sense of the parties involved in this bargaining as truly separate sources of agency. And even if we could, it is hard to see how cooperation could emerge out of the interactions between these parties. It is doubtful, therefore, that the will can be understood as a genuine *inter*personal phenomenon.

Consider the alleged competition among separate interests. Ainslie presents the interests as independent agencies that strive for selection (Ainslie 2001, pp. 39–41, 61, 73) as if they were replicators in a process of natural selection. But he gives no reason to believe that there are heritability and differential fitness in the competition among interests. What the selection amounts to is just that the strongest interest is satisfied at the expense of the conflicting, weaker interests. This satisfaction does not alter the chances of any interest to reappear with equal strength in the future. Nor does it promote the development of any adaptive strategy by the interests themselves. Understanding of the effects of hyperbolic discounting does not depend on the unwarranted reification of the agent's preferences into independent sources of agency that compete strategically in a genuinely selective process. Talk of selection among the preferences and the development of strategies to deal with the conflicts of preferences is more appropriate at the level of agents, even if the agents operate in response to the varying strengths of their preferences. In any event, interaction among interests could not explain the emergence of commitments. A short-range interest has no incentive to submit to a commitment, because commitments preclude the interest's present and future satisfaction. Interests seek nothing other than their satisfaction, hence nothing can be offered to them in exchange for their frustration.

What if the parties are not interests, but successive temporal selves? Ainslie occasionally shifts from talk of transient interests to talk of temporal selves (Ainslie 2001, pp. 40, 93, 161). The two notions are not identical, however. Contrary to transient interests, temporal selves are sources of agency and can have multiple interests. It seems that temporal selves might agree to be under commitments that frustrate their dominant short-range preferences in exchange for the satisfaction of other preferences. However, temporal selves are transient, hence they have no incentive to settle for less than the satisfaction of their short-range dominant interest.

The problem would not arise for parties that are transient in the sense that they act just once, but have stakes in the long-term outcomes of their actions (see Ainslie 2001, p. 93). These parties have no problem seeing the long-term benefits of a commitment. Nevertheless, they are tempted to make an *exception* now, thereby satisfying their dominant short-range interest while still reaping the long-term benefits of future compliance. However, if the present action counts as a precedent, a single exception to the commitment is self-defeating, given that present defections invite future ones. For Ainslie, transient interests/selves happen to be related so that their actions count as precedent for future ones, whence the stability of commitment. However, the fact that transient parties with long-term stakes can strategically agree to cooperate does not explain the will. First, there is no need to look at interpersonal scenarios to appreciate that actions can work as *intra*personal precedents. Given that the *same* agent is going to face the *same* choice at the future time with the *same* set of preferences, it is not surprising that her present action is a precedent for her future ones, thereby defeating temptations to make exceptions to her commitments. This is not really *strategic* thinking, but just reflection on the import of one's present action in the context of one's continued existence as one and the same agent who is going to face exactly similar choices in the future. Appeal to transient selves adds nothing to this straightforward *intra*personal explanation. Moreover, in order to make the repeated Prisoner's Dilemma scenario envisaged by Ainslie (2001, p. 93) truly explanatory, special interpersonal conditions must be assumed: The parties must face exactly the same choice over time and share the same long-term preferences. But these conditions are not special from the intrapersonal point of view. They are just distinctive features of the agent's temporal identity.

Second, the fact that we are subject to hyperbolic discounting and thus prone to inconsistent shifts in short-term preferences is no reason to think that we are made up of competing transient selves with long-term stakes. What makes this false picture attractive is the misleading focus on scenarios like "Ulysses and the sirens" as if they were paradigmatic of diachronic agency. Ulysses' situation, however, is unusual. When Ulysses listens to the sirens, he does not just reverse his short-range preferences, rather, he is also insensitive to long-term considerations. Hence, he does not care that his action could be a precedent. But this makes him impervious to commitments. He can be controlled only by physical restraints or short-range disincentives. If hyperbolic discounting were to make us always like Ulysses, our lives would indeed be best described in interpersonal terms. But then there would be no will, just crude transtemporal manipulation. On the other hand, if temporal selves are depicted as having not just shifting short-term

https://doi.org/10.1017/S0140525X05260119 Published online by Cambridge University Press

preferences, but also stable long-term ones, then they have no explanatory force. They are just like ordinary extended agents with stable long-term preferences, except that they act just once. Speaking of an agent as made up of a succession of these selves is just a convoluted way of saying that, as a result of hyperbolic discounting, we are going to have dynamically inconsistent short-term preferences.

Ainslie is right in claiming that we are subject to hyperbolic discounting and that the will is partly a matter of predicting future conduct from one's present action. But this conclusion is independent of the misleading picture of a bargaining among competing interests/selves. The will is not an organ, but it is not a bargaining situation either. It is rather the product of the agent's reflection on the long-term effects of her actions, including their role as precedents, under the assumption of her own identity over time.

Hyperbola-like discounting, impulsivity, and the analysis of will

Leonard Green and Joel Myerson

Department of Psychology, Washington University, St. Louis, MO 63130.
LGreen@wustl.edu JMyerson@wustl.edu

Abstract: Ainslie's insightful treatment of dynamically inconsistent choice stands in contrast to traditional views in psychology, economics, and philosophy. We comment on the form of the discounting function and on new findings regarding choice between delayed rewards. Finally, we argue that the positive correlation between temporal and probability discounting is inconsistent with the view that impulsivity represents a unitary trait.

In *Breakdown of Will*, Ainslie (2001) proposes a radical departure from the traditional approach to the concept of will. He argues against the notion of will as a faculty that exerts top-down control (an executive function, as it were). A major problem with this traditional, top-down approach, as Ainslie points out, is that it cannot easily account for dynamically inconsistent choice. For example, an individual may agree to give a talk at an upcoming meeting, yet regret having made that commitment as the date gets closer. From the traditional perspective, it seems odd that the same person who is endowed with a sufficient degree of will at an earlier point in time would not possess it at a later point. Even odder from the traditional perspective is that the same person often will agree to give yet another talk if asked well in advance of a subsequent meeting.

Rather than blaming such apparently inconsistent behavior on inadequate willpower or selective stupidity, Ainslie suggests that such dynamic inconsistencies, as well as the behavioral phenomena traditionally explained in terms of will, can be better accounted for by conceptualizing will as a strategy that involves thinking of individual choices as precedents. Thinking of choice in terms of precedents that establish choice policies has the effect of "bundling" the sequence of rewards that would be obtained by following a specific policy. Giving a talk, for example, may be thought of as part of a larger, ongoing plan for career advancement associated with larger rewards than might result from a single talk (see also Rachlin 1995; 2000). According to Ainslie, the utility of this bundling strategy derives from the hyperbolic nature of the discounting of future consequences.

Ainslie was among the first to argue that hyperbolic discounting predicts that preferences may reverse with the passage of time, whereas the standard economic model, which assumes exponential discounting, predicts stable preferences. That is, although a larger, later reward may be chosen over a smaller, sooner reward when both are sufficiently delayed, as time passes and the delay until both rewards is reduced equally, preference may reverse and the smaller, sooner reward becomes the more attractive alternative. In the absence of some form of overt commitment, well-intended resolutions frequently are broken as preference reverses. Such inconsistency emerges naturally from hyperbolic discounting and, according to Ainslie and others, provides compelling evidence against exponential discounting.

Preference reversals, however, do not, in and of themselves, rule out exponential discounting. If larger delayed rewards are discounted less steeply than smaller delayed rewards, then exponential and hyperbolic discounting both lead to preference reversals (see Fig. 1 in Green & Myerson 1993). This is important because a number of studies have shown that discounting rate is, in fact, inversely related to amount of reward (Chapman & Elstein 1995; Green et al. 1997; Kirby 1997; Raineri & Rachlin 1993), and thus an alternative method for distinguishing between exponential and hyperbolic discounting models is needed. Curve fitting provides such a method, and when exponential and hyperbolic functions are fit to individual discounting data, the hyperbolic consistently accounts for a larger proportion of the variance (Kirby & Marakovic 1995; Myerson & Green 1995; Simpson & Vuchinich 2000).

As it turns out, curve fitting reveals that a hyperbola-like discounting function in which the denominator is raised to a power provides an even better fit to individual discounting data than either an exponential or a simple hyperbola (Myerson & Green 1995; Simpson & Vuchinich 2000). More specifically, the power to which the denominator is raised is often significantly less than 1.0, and never significantly greater than 1.0. Moreover, often when the data from an individual cannot be described by a simple hyperbola, it can be described by a hyperbola-like discounting function. It is important to note that the hyperbola-like function preserves the essential characteristic required by Ainslie's framework – steeper discounting at short delays and less steep discounting at longer delays, rather than the constant (stationary) discounting rate predicted by the exponential. Thus, although we disagree with Ainslie's argument against exponential discounting based on preference reversals, nevertheless we agree with Ainslie's bottom-line conclusion that the discounting function is hyperbolic (or at least hyperbola-like).

Studies in which curve fitting has been used to evaluate the form of the discounting function almost always have examined choice between an immediate and a delayed reward. An important issue arises when choice involves two delayed rewards. In such situations, people do not simply compare the present (hyperbolically discounted) values of the two rewards, as most behavioral-economic analyses of preference reversals, including Ainslie's, would imply. Neither does behavior conform to the predictions of an *elimination-by-aspects* decision process (Tversky 1972). For example, in a situation where one has to choose between a smaller reward available in one year and another, larger reward available in two years, both alternatives share the common aspect of a one-year wait, yet people do not ignore the common one-year wait. Instead, recent work in our laboratory (Green et al., in press) suggests that discounting functions are still hyperbola-like, but people give less-than-full weight to the common aspect of the delays. By not taking the actual delays fully into account, individuals increase the likelihood that they will choose the sooner, smaller reward, thereby exacerbating their problems with self-control and furthering the "breakdown of will."

Which brings us to the question of what is the best way to talk about the important issues under consideration in Ainslie's insightful book. It should be clear to his readers that when Ainslie refers to *self-control*, he means control *of* the self (in the sense of control of one's own thoughts and behavior), rather than control *by* the self, and we strongly endorse his use of the term. A related issue concerns the use of the term *impulsivity*, a term that does not occur in the index to *Breakdown of Will* but which is increasingly used by researchers to refer to the tendency to choose smaller, sooner over larger, later rewards. Ainslie's arguments against a unitary faculty of willpower also apply to the notion of a unitary trait of impulsivity, which appears to mean simply a lack of willpower. According to the *Diagnostic and Statistical Manual of*

https://doi.org/10.1017/S0140525X05260119 Published online by Cambridge University Press