

Commentary: Huang & Bargh for BBS 37(2) 2014

Selfish goals must compete for the common currency of reward¹

George Ainslie^{a,b}

^a*School of Economics, University of Cape Town, Rondebosch 7701, South Africa;*

^b*Department of Veterans Affairs, Coatesville, PA, 19320.*

George.Ainslie@va.gov

<http://www.picoeconomics.org>

Abstract:

Selfish Goal Theory is compatible with a behaviorally based theory that recognizes mental processes as behaviors. Both envision choices as made by the competition of purposive processes, which are autonomous in that they are not coordinated by an agentic “self.” However, the survival of mental processes – termed “goals” or “interests,” respectively – depends on a well-documented active mechanism: reward.

Text:

Selfish Goal Theory is a step toward bridging the conspicuous gap between cognitive and behavioral psychology that has existed since the “cognitive revolution” of the 1970s (Baars 1986). A key tenet of the cognitive approach has been to avoid positing a directly selective mechanism for choice, such as reward, reinforcement, or utility – a reaction against the behaviorists’ total substitution of such mechanisms for explanatory mental constructs (as in Rachlin 1985). Between avoiding selective mechanisms on one hand and avoiding mental constructs on the other, theorists have not developed the obvious possibility that mental processes compete with one another on the basis of prospective reward. Interestingly, Bargh’s earlier writings proposed that much of our “thinking, feeling, and doing ... is driven by current features of the environment” (1997, p. 2), leading to the suggestion that he was essentially a behaviorist who recognized internal behaviors (Mischel 1997). Huang’s and his current proposal interposes a market-like mediating process between environmental features and behavioral output, based on the competition of goals. This is clearly an advance, but it suffers from ambiguity about the basis on which the goals compete.

Huang & Bargh (H&B) analogize the internal selection of goals to the natural selection of organisms. This is an apt comparison, but whereas genes are selected for differential survival by adaptiveness, the selective process for goals is not specified. H&B are clear (and bold, for cognitivists) in stating that an overarching faculty of choice is not necessary – that the conscious self serves mainly a public relations function (sect. 5.2) – but they are vague about their alternative. H&B have it in mind that goals are “incentivized,” and they name B. F. Skinner’s “operant” as a synonym of “goal” (sect. 4). They note that “representation of a goal can become associated with positive affect” (sect. 4.2.1, para. 5), but this only “signals” goal desirability, rather than creating it. The most important determinant of selection is said to be “the passive activation or *priming* of higher-order concepts by contextual features” (sect. 2, para. 3; authors’ italics), but there is no mention of how, in a conflict, some contextual features get more priming power than others.

All H&B’s examples of conflicts involve long-term versus short-term payoffs: behaviors that are “contrary to his or her general self-interests” (sect. 4.5, para. 1) such as “dangerously

unhealthy” sexual attractants, putting off work and eating tempting foods (sect. 4, para. 3), and “self-control failures” (sect. 5.2, para. 2). Such actions are said not to “map onto” the person’s own well-being, “even to the point of feeling helpless to prevent the goal pursuit” (sect. 4.5, para. 2). The lack of a clear hypothesis about goal selection leads to trouble in explaining why the mapping fails, that is, why contradictory goal pursuits endure over time rather than resulting in simple preference for one over the other. Given the authors’ extensive documentation of unconscious processes in all phases of choice, we might have thought that they were building an explanation for such conflicts using Freudian repression. But no. H&B’s evidence is that unconscious processes are capable of the same operations as conscious ones, including such self-control procedures as “activation of motoric avoidance goals” against addictive stimuli, and “strategically automatic implementation intentions” to promote healthy behaviors (sect. 5.2, para. 2), so the dimension of consciousness is unimportant. H&B suggest that there is an underlying coherence to these contradictory behaviors, which may only appear inconsistent “to an outsider” (sect. 4.5, para. 1), but they go no further.

On the other side of the behavioral-cognitive gap, a behaviorism that recognizes internal behaviors has no problem with selfish goals. It calls them interests, the set of behaviors (and thoughts) that are based on a particular source of reward (Ainslie 1992, pp. 88–94). Reward is the process that selects for the mental activity that led to it. Far from being a hypothetical construct, reward performs in controlled experiments with mathematical precision (e.g., Daw & Doya 2006) and can be directly visualized with magnetic resonance imagery in specific coordinated brain centers (e.g., Kable & Glimcher 2007). Like the selection of selfish goals, selection by reward has been compared to natural selection (e.g., Gilbert 1972; Vaughan & Herrnstein 1987). The reward process shapes interests – or goals – as a proxy for the evolutionary adaptiveness that must have led to the selection of its mechanism in turn (Ainslie 1992, pp. 179–84). Although the cognitive revolution rightly faulted classical behaviorism for requiring reward to come from external events, cognitivists might not object to endogenous (self-generated) reward. Expected reward is devalued according to a hyperbolic or hyperboloid function of delay (Berns et al. 2007), which, among other effects, could prevent endogenous reward from short-circuiting the behavioral selection process (Ainslie, 2013), and might govern “higher-order goals” (sect. 4.2.3). The finding of hyperbolic discounting is especially relevant to the recurrent dominance of “temporarily active goals” such as unhealthy behaviors (sect. 4.1), because it predicts a disproportionately increased forcefulness of goals when they can be imminently fulfilled. As in Selfish Goal Theory, the property of consciousness is not important.

Hyperbolic discounting also predicts a self-control phenomenon not envisioned here: bargaining among separately motivated interests (or goals) that are expected to be dominant at different times, a variant of repeated prisoner’s dilemma (Ainslie 1992; 2012). In such intertemporal bargaining, the perception of a current choice as a test case for similar choices in the future recruits incentive for a long-term goal against a short-term one. Like the interaction of autonomous goals, intertemporal bargaining takes place mostly at an intuitive (or unconscious) level, as demonstrated indirectly by thought experiments (Ainslie 2007), and generates familiar ego functions from the bottom up, “obviating the agentic ‘self’” (sect. 2, para. 3). When such bargaining is taking place, rationalization serves to maintain not only public consistency (sect. 5.2), but also intertemporal trust, avoiding the agent’s perception of a defection that would undermine intertemporal cooperation.

Thus, two lines of development, one from behaviorism, the other from social-cognitive psychology, lead to a description of the apparent self as a population of autonomously competing processes. However, the two models operate differently. The main determinant of choice in intertemporal bargaining theory is contingent reward, a process that used to be thought of as strictly external and thus too narrow for human behavior. Selfish Goal Theory seems to view the main determinant of choice as priming, a natural variant of hypnotic suggestion (Bargh 2006, p. 155), which has been unrecognized for the very reason that it is largely unconscious. Certainly the great contribution of priming research has been to show how much choice-making occurs without awareness; but this does not mean that the resulting choices escape the influence of prevailing incentives in any major way. Unhealthy behaviors persist because they are attractive in the short run, and, perhaps tellingly, they respond poorly to hypnotic therapies (Barnes et al. 2010). Conversely, hypnotists cannot impose repugnant behaviors on subjects. To account for disruptive impulses and strong defenses, Selfish Goal Theory needs to take account of how goals bargain with each other in terms of a common currency, best characterized as reward.

<c-text ends>

ACKNOWLEDGMENT

I thank John Monterosso for comments and suggestions. The cited works under the author's control can be downloaded from www.picoeconomics.org. This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA.

Note

1. This is a work of the U.S. Government and is not subject to copyright protection in the United States. The opinions expressed are not those of the Department of Veterans Affairs or of the U.S. government.

References:

Ainslie, G. (1992) *Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge University Press.

Ainslie, G. (2007) Can thought experiments prove anything about the will? In: ed. D. Spurrett, D. Ross, H. Kincaid & L. Stephens, *Distributed Cognition and the Will: Individual Volition and Social Context*, pp. 169–96. MIT.

Ainslie, G. (2012) Pure hyperbolic discount curves predict “eyes open” self-control. *Theory and Decision* 73:3–34. doi: 10.1007/s11238-011-9272-5

Ainslie, G. (2013) Grasping the impalpable: The role of endogenous reward in process addictions. *Inquiry* 56: 446-469. DOI: 10.1080/0020174X.2013.806129.

Baars, B. J. (1986) *The cognitive revolution in psychology*. Guilford.

Bargh, J. A. (1997) The automaticity of everyday life. In: *Advances in social cognition, vol. 10*, ed. R. S. Wyer, Jr., pp. 1–61. Erlbaum.

- Bargh, J. A. (2006) What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology* 36:147–68.
- Barnes, J., Dong, C. Y., McRobbie, H., Walker, N., Mehta, M. & Stead, L. F. (2010) Hypnotherapy for smoking cessation. *Cochrane Database System Review* 2010; (10):CD001008.
- Berns, G. S., Laibson, D. & Loewenstein, G. (2007). Intertemporal choice –Toward an integrative framework. *Neuroeconomics* 11:482–88.
- Daw, N. D. & Doya, K. (2006) The computational neurobiology of learning and reward. *Current Opinion in Neurobiology* 16:199–204.
- Gilbert, R. M. (1972) Variation and selection of behavior. In: *Reinforcement: Behavioral analyses*, ed. R. M. Gilbert & J. R. Millenson, pp. 263–76. Academic Press.
- Kable, J. W. & Glimcher, P. W. (2007) The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience* 10:1625–33.
- Mischel, W. (1997) Was the cognitive revolution just a detour on the road to behaviorism? On the need to reconcile situational control and personal control. In: *Advances in social cognition, vol. 10*, ed. R. S. Wyer, Jr., pp. 181–86. Erlbaum.
- Rachlin, H. (1985) Pain and behavior. *Behavior and Brain Sciences* 8:43–83.
- Vaughan, W., Jr. & Herrnstein, R. J. (1987) Stability, melioration, and natural selection. In: *Advances in behavioral economics, vol. 1*, eds. L. Green & J. H. Kagel, pp. 185–215. Ablex.