

MIT OpenCourseWare
<http://ocw.mit.edu>

24.120 Moral Psychology
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

VIII Strength of Will: Descriptive Issues

AINSLIE'S EIGHT MARKS OF THE WILL

1. A new force in addition to desires etc.
2. Throws its strength on the weaker side. (or perhaps it *can*)
3. Unites actions under a common rule. (?)
4. Strengthened by repetition
5. Very vulnerable to nonrepetition
6. Requires no diversion of attention
7. Doesn't normally depend on each single choice, except where stakes are very high.
8. Tendency of failure in one sphere to spill into others is variable.

WHY DO WE EXHIBIT WEAKNESS OF WILL?

Exponential and Hyperbolic discounting. \$100 now or \$120 in a year's time; \$100 in ten year's time or \$120 in eleven year's time. From pigeons to people (though non-humans are not good at delaying reward). Hyperbolic discounting is, in some sense, irrational.

Hyperbolic discounting is probably also involved in what we might call *one more won't hurt* reasoning:

One more won't hurt

You are a smoker, currently smoking forty cigarettes a day. You believe that this is very bad for you, and that you should give up. But you reason as follows: one more cigarette won't hurt me, and it will give me considerable pleasure now. So I might as well have this one more. And then you use this argument forty times a day.

Without hyperbolic discounting it looks as though the subject would simply be making a mistaken calculation (is that so in all cases?); but with it they need not be. Is hyperbolic discounting an *explanation* or merely a *description*?

AINSLIE'S EXPLANATION OF HOW STRENGTH OF WILL IS POSSIBLE

Competition between egoistic time slices, akin to an iterated prisoners' dilemma. ('Picoeconomics') The reward game. See how this gives us the eight marks of the will.

PROBLEMS WITH AINSLIE'S EXPLANATION

- (i) General problem of time slices: how short should they be? Can we make sense of actions done by time slices?
- (ii) Each time slice is dead by the time of fulfillment, so in what sense is there a long run benefit for which they can trade?
- (iii) What has become of the irrationality?

REVISING AINSLIE'S EXPLANATION

An alternative, solving the second problem: give up on the egoism. Assume that the time slices have desires (i) for others' well being, which (ii) they will never know to have been fulfilled. But:

- (a) we've now lost the sense of competition; this is quite unlike a prisoner's dilemma.
- (b) even these altruistic desires aren't enough, since they don't explain a further phenomenon, that having kept to a resolution provides a reason for continuing to do so (compare the sunk cost phenomenon).
- (c) it still doesn't help with the third problem. Why should we think that it is irrational for different actors to have different desires?

A MORE RADICAL REVISION

Better just to put back the enduring self. What is left from the prisoners' dilemma model is just the idea that in deciding what to do, we are acutely sensitive to what else will be done, but by us at a later time, rather than by others. This still keeps all of the advantages. In particular, we keep (indeed can make even better sense of) this reasoning:

- (i) If I resist temptation now, then I'll resist it in the future.
- (ii) If I don't resist temptation now, then I won't resist it in the future.

OTHER FEATURES OF THE PROPOSAL

The status of rationalizations and exceptions.

Bright lines. Alcoholics Anonymous on the need never to drink.

WORRIES

Does this help explain all cases of strength of will in the face of desire? What about a case in which there is no likelihood of repetition? The highly unusual one night stand.

Does this explain how there can be *effort* involved in resisting temptation?

EGO-DEPLETION

Baumeister *et al.*: it seems that the will does resemble a muscle. Complications: Choice also causes ego depletion. Prior resolutions seem to prevent ego-depletion.

MISCHEL'S FINDINGS

SEPARATE NEURAL SYSTEMS: MCCLURE *ET AL.*