Jamison, Julian; Wegener, Jon

Working Paper

Multiple selves in intertemporal choice

Working paper series // Federal Reserve Bank of Boston, No. 09-17

Provided in Cooperation with:
Federal Reserve Bank of Boston

Working Papers

FEDERAL RESERVE BANK OF BOSTON™

No. 09-17

Multiple Selves in Intertemporal Choice

Julian Jamison and Jon Wegener

Abstract:

We propose that individuals consider future versions of themselves to truly be separate persons, not simply as a convenient modeling device but in terms of actual brain systems and decision-making processes. Intertemporal choices are thus quite literally strategic interactions between multiple agents. Previous neuroscientific studies have found evidence that systems involved with Theory of Mind (that is, mentalizing other agents) are similar to those involved with prospection (imagining oneself in the future). We provide a conceptual framework for this work and suggest that, instead of prospection, a more analogous future task is one that concerns intertemporal choice and time preferences, since these involve implicit prediction of future actions. Recent functional imaging studies appear to confirm such a link. Additional studies—behavioral, clinical, and neuroimaging—are proposed in order to confirm the specific nature of the correspondence and to elucidate the underlying mechanisms. Finally, given that society may have a vested interest in promoting the welfare of future selves, we discuss possible policy implications of departing from the standard framework in which individuals act in their own best interests as defined over the entire lifetime.

Keywords: intertemporal choice, theory of mind, mentalizing, intrapersonal games, neuroeconomics

JEL Classifications: D90, D87, D62, D03

Julian Jamison is a senior economist in the Research Center for Behavioral Economics and Decisionmaking at the Federal Reserve Bank of Boston and a visiting faculty member in the department of economics at Yale University. His e-mail address is julian.jamison@bos.frb.org. Jon Wegener is a postdoctoral fellow in the Decision Neuroscience Research Group at the Copenhagen Business School. His email address is jonw@drcmr.dk.

This paper, which may be revised, is available on the web site of the Federal Reserve Bank of Boston at http://www.bos.frb.org/economic/wp/index.htm.

We thank participants (especially Lauren Willis) at the USC conference on Law, Economics, and Neuroscience, two anonymous reviewers, and Suzanne Lorant for their helpful comments. We also thank Carl Nadler for his valuable research assistance.

This version: November 30, 2009

RESEARCH CENTER FOR BEHAVIORAL ECONOMICS AND DECISIONMAKING

Introduction

The notion of *self* has a long tradition in both philosophy and psychology, dating back to at least Hume (1739). In economics it has been primarily an implicit focus, prominent in the assumption of maximizing an individual's utility function and in performing welfare analyses by aggregating and comparing across individuals. The concept also has a long legal history, with questions of autonomy rising to the forefront. Of course, *self* has essentially no meaning except in distinction to some *other*, and so the relevant question becomes where to draw the borderline between the two. In social psychology, studies of group identity have suggested that sometimes the notion of self extends beyond that of the individual. Here, we suggest the reverse: in one specific but important way—namely across time—the notion of self is considerably narrower than the individual.

Although it may sometimes be useful to consider a model in which multiple selves compete simultaneously within an individual agent, we do not do so here. Rather, we focus on the temporal aspect: any two decisions made by an individual are not in fact made by the same self of that individual, as something has necessarily changed in the interim (Reed 2008). Heraclitus said, "You could not step into the same river twice; for other waters are ever flowing on to you." But it is just as true that you cannot step into the same river twice because you yourself have moved on, incorporating new information, new memories, and new thoughts. More importantly, whether or not this change occurs, your current self perceives that it may occur and treats the world accordingly.

Our assertion goes beyond the standard economic "as if" modeling trope where (in this instantiation) people behave as if they have multiple selves over time. Instead we propose that the decision-making process involving a tradeoff between our current and future selves is substantially the same as the decision-making process involving a tradeoff between ourselves and other individuals. In the next two sections we describe the neuroscientific data that leads us to this conclusion and suggest additional studies that could substantiate and refine this link. In the final section of the paper we discuss various possible policy implications arising from the idea that the process itself (and not just the outcome of the process) is so similar in these two apparently disparate situations.

To fix some terminology, note first that the terms *intertemporal choice* (used primarily in neuroscience), *delay discounting* (primarily in psychology), and *time preference* (primarily in economics) all refer to roughly the same concept. Namely, these refer to decisions in which an individual agent must trade off outcomes (monetary or otherwise) received at different points in time. This concept is related to *prospection*, in which an agent is asked to imagine him- or herself at some future point in time. Meanwhile, *retrospection* can be considered to be the analogue of prospection for the past. Note in particular that for both prospection and retrospection there is no current action to be taken, nor is there any prediction of behavior at the alternate point in time.[1]

The notion of a *game* in economics refers to strategic interactions between agents, for which the payoff of one player depends on the choices of another. In particular, one key element to playing a game is to form beliefs about the motivations and actions of other players. When economists formally model an individual's choices of consumption vs. savings behavior over the lifecourse, they often write this down as a game between multiple selves—with each self making a single decision at each point in time.[2] This general approach to modeling intrapersonal behavior is by no means limited to economics; typical applications include Ding (2007) in the marketing literature and Frank (1996) in the addiction literature. Much attention has been paid recently to the issue of *time-inconsistent* preferences, which refers to utility functions under which agents may deviate from planned future behavior only because time has passed and not because any new information has arrived. That is, the agent at time 0 and the agent at time $t>0$ desire different actions at time t, meaning that they are (at least to some degree) in conflict with one another. This can naturally be modeled as a game between two different selves with different goals; both Laibson (1997) and O'Donoghue and Rabin (1999) give more general technical expositions of the mechanics.

[1] This is obvious for retrospection, since any such behavior has already been taken and need not be predicted. For prospection, this is because the standard paradigm asks only that a future scenario be imagined.

[2] See Thaler and Shefrin (1981) for an early example of this approach, and Krusell and Smith (2003) for a more recent and comprehensive example.

Notably, there are typically multiple equilibria in all such models (see Vieille and Weibull 2009) —often a 'good' one in which, for example, all the selves exhibit self-control and a 'bad' one in which the selves do not. The good equilibrium is supported by the threat that if any one self breaks down (for example, has a cigarette, or fails to maintain a diet) then all future selves will also lapse. Thus, the consequences of one's actions are magnified via the strategic interaction. Future selves affect current selves not simply because of what might be called altruism (that is, placing positive weight on future welfare) but also because of this interaction. Just as in a typical game (whether simultaneous or sequential), what is key are the beliefs about the behavior of the other players, so there is a direct link from the future selves back to the current self. Intuitively, this is what gives rise to the "micro-commitments" (Reed 2008) that people use every day, maintaining the importance of a 'bright line.' One example is alcoholics who convince themselves that even a single drink will lead to ruin, in some sense artificially exaggerated the importance of every possible drink precisely because of expectations about future behavior.

The general process of putting oneself into the shoes of another, often to better predict her behavior by emulating her thoughts and desires, is the central component of *Theory of Mind* (ToM forthwith; see Premack and Woodruff 1978) or *mentalization* (Frith et al. 1991). This is defined as the ability to perceive one's own mental state and (more importantly) to attribute analogous but distinct mental states to others. The exact constituents of ToM are still debated in the literature, but we focus here on three distinct dimensions: prediction of behavior by others; empathy with others; and affinity with others. While affinity is not generally considered part of the definition of ToM, it seems to correspond to greater success at ToM-related activities, as indicated by a neuroimaging study showing increased activity in parietal and frontal regions when morphed faces contained more of the subject's own face (Uddin et al. 2005; see also Mobbs 2009).

Given the correspondence above between intertemporal choices and strategic games (which heavily utilize the ability to mentalize), one might expect to find activation during such choices in the same regions known to be involved with ToM. Indeed, our central idea is that essentially the same brain system is utilized when individuals attempt to predict or empathize with others and when they attempt to predict or empathize with themselves in the future. This brain system consists primarily of the temporal poles, the medial prefrontal cortex, and areas both within and in proximity to the temporoparietal junction (Frith and Frith 2006). Furthermore, we believe that this ability to predict or empathize with future selves is the prime component of pure time preferences and is therefore activated whenever an intertemporal choice is encountered. Hence, ToM applies not only to other individuals but equally to other selves within an individual. More distal selves, like more distal acquaintances, are more difficult to predict or empathize with, triggering less ToM activity and corresponding to lower levels of affinity. Future selves are not simply treated the same as other people, but indeed for the same reason.

Previous Data

A classic behavioral task used to test ToM in children is the false-belief task: The child sees the experimenter, two dolls named Sally and Anne, a basket and a box. Sally places a marble in a basket while Anne is observing. Then Sally exits and Anne transfers the marble from the basket to the box. Sally reenters and the experimenter asks the child where Sally will look for her marble. If the child answers that Sally will look in the basket, it indicates an understanding of false beliefs in other agents—and hence that the child has developed a theory of mind. The ability to understand and verbally explain this task develops around the age of four in healthy children, while studies on children with autism and Down's syndrome, between the ages of 6 and 17, demonstrated that four out of five failed the Sally-Anne task (Baron-Cohen et al. 1985).

Together with the ability to mentalize, children at the age of four have developed the ability to forgo immediate rewards for larger, delayed rewards. Furthermore, individual differences in the ability to delay rewards has been shown to correlate positively with intelligence, higher achievement striving, and greater social responsibility later in life (Mischel et al. 1989). Additionally, time preferences change over the lifespan according to a U shape, where young adults and elderly discount the future more than the middle aged (Read and Read 2004).

Imaging studies on ToM in healthy subjects have focused on the ability to ''mentalize'' and have used stories, cartoons, and animated geometric shapes that differed in the degree to which they represented the intentions, beliefs, and desires of others (for a review, see Gallagher and Frith 2003). Other studies have involved the neuroimaging of subjects while they played strategic games (McCabe et al. 2001, Gallagher et al. 2002, Rilling et al. 2004) with either a virtual or real partner. All of these studies have repeatedly given evidence for the involvement of three brain areas in mentalizing: the temporal poles, the temporoparietal junction, and the medial prefrontal cortex.

A handful of neuroimaging studies have indicated that mentalization relies on the same discrete network of brain areas as prospection, which involves mentalizing about oneself in the future. Studies have generally focused on episodic prospection whereby subjects are asked to imagine themselves in a future scenario. To our knowledge, five papers have explicitly focused on the neural processing of prospection using fMRI (functional Magnetic Resonance Imaging) (Addis et al. 2007, D'Argembeau et al. in press, Sharot et al. 2007, Botzung et al. 2008, Szpunar et al. 2007). Common to all five studies is that subjects are prompted to imagine themselves in the future, using either events that have been experienced by the subject in the past (Botzung et al. 2008, D'Argembeau et al. in press), common life experience events (for example, a future birthday, Szpunar et al. 2007), events constructed by the experimenter (Sharot et al. 2007), or events that are constructed and elaborated on by the subject (Addis et al. 2007). Although the control tasks used in the study varied from fixation (Sharot et al. 2007) to more elaborate semantic tasks (Addis et al. 2007, Botzung et al. 2008), prospection consistently activated the temporal poles and the temporoparietal junction, while the medial prefrontal cortex was also found to be activated in tasks requiring the ability to mentalize (Spreng et al. 2008).

An important limitation to the above studies explicitly investigating prospection is that they rely on subjects' ability to imagine themselves in future events without having real incentives to do so. There is no possibility to control for any mind wandering to non-task-related issues that might occur if subjects are allowed 20 seconds to elaborate on future scenarios (for example, Addis et al. 2007). Furthermore, as some prospection studies rely on imagining common life experiences in the future, or imagining experienced events as happening in the future, it is difficult to separate brain activation related to episodic memory from brain activation related to prospection. Given these inherent limitations of prospection, perhaps better are the intertemporal choice tasks that provide real incentives for subjects to engage in empathizing about themselves in the future. These require subjects to trade off the provision of more immediate resources (for example, monetary resources) with more distal ones.

Recent theoretical developments have highlighted how memory is queried in valuing options before executing a decision (Johnson et al. 2007). With data replicating a well-known asymmetry in intertemporal choices showing higher time discounting when the more immediate option is shown first (Loewenstein 1988, Loewenstein and Prelec 1992), Weber and colleagues extended this finding, demonstrating that there are a greater number of thoughts associated with the first option (Weber et al. 2007). The authors suggest that more memory retrieval resources are allocated to the first option shown (Weber et al. 2007), consistent with a large body of evidence showing that the retrieval of memories associated with the first item, interferes with subsequent items (review: Levy and Anderson 2002). This is one of the few other papers that attempts to make a link between the processes behind memory, prospection, and choice. On the other hand, there have been a range of neuroimaging studies investigating different components of intertemporal choice (for example, McClure et al. 2004, 2007, Boettiger et al. 2007, Glimcher et al. 2007, Kable and Glimcher 2007, Wegener et al. 2008). Whereas the study by McClure and colleagues focused on comparing intertemporal choices involving an immediate reward or only delayed rewards, with results suggestive of two separate systems as in Laibson (1997), the studies by Glimcher and colleagues attempted to isolate the subjective value component of intertemporal choice and found evidence for a single system at work.

In the intertemporal choice study by Wegener et al. (2008), the time component in intertemporal choices was isolated. Here, an adaptive paradigm was used that adjusted the monetary reward offered at each delayed time point (one week, one month, one year) to approach indifference within each subject between the fixed immediate reward and the delayed rewards. Interestingly, the regressor isolating the time component demonstrated progressively declining activation around temporoparietal junction and anterior temporal areas bilaterally known to be associated with prospection and mentalizing, as illustrated in Figure 1 (Buckner and Carroll 2007, Spreng et al. 2008). In addition, the decline in activation around the temporoparietal junction within individual subjects correlated positively with how much the individual discounted the future, that is, how much larger the reward should be in order for the subject to be indifferent to the immediate reward, as shown in Figure 2. Firstly, this suggests that the subjects as a group mentalize less with more distal selves, as represented by the decline in activation in areas involved in prospection and mentalizing. Secondly, this suggests that intertemporal choice tasks may be used to study prospection by providing incentives to empathize with future selves. Since the temporoparietal junction is not exclusively involved in mentalizing processes (for example, Mitchell 2008), we face (as always with fMRI) a "reverse inference" problem (Poldrack 2006). Some of the future studies proposed in the next section partially address this issue by further investigating the dynamics between time preferences and mentalizing.

Future Work

It is difficult to determine, using only observed behavioral data, whether a similar mechanism is being used for decisions relating to others and to future selves. For instance, consider a modified dictator game in which individuals must choose between two options: option A yields $10 for oneself, whereas option B yields $8 for oneself and $4 for a randomly chosen anonymous partner. In this case, some people prefer option A, some prefer option B, and some are indifferent. It would be relatively straightforward to find a delay time such that similar proportions of people chose analogously between the following options: option A yields $10 immediately, while option B yields $8 immediately and $4 at the specified future time. But we would learn almost nothing about actual pathways from such a study, since these two similar outcomes could very easily have arisen from entirely separate brain processes, evolved to address what appear to be separate situations.

On the other hand, it is known that some subjects are better at ToM than others, and it would be possible to compare this trait with a related version regarding future selves. In particular, focusing on the prediction dimension, one could test whether individuals who are proficient at predicting the behavior of others are also [relatively speaking] proficient at predicting their own future actions, controlling for age and other relevant variables. Such a correlation would hardly be conclusive, but would be suggestive that the links pointed out in the previous section apply in terms of purely behavioral outcomes, and that they apply along the sub-dimension of predicting choices. For the other dimension that we have identified, it is conceivable that standard survey assessments of empathy could be modified to address one's future self. Then the responses could be compared, although this is a less compelling test than the one regarding prediction.

Studies on non-human primates have shown that conspecific calls, recruiting processes thought to involve theory of mind, result in higher activity than non-conspecific calls, in areas homologous to the human temporoparietal junction (Gil-da-Costa et al. 2004).

Patients with a specific lesion to the temporoparietal junction are known to have theory of mind impairments (Samson et al. 2004). According to recent reviews (Buckner and Carroll 2007, Spreng et al. 2008) these patients should also demonstrate impairments in prospection (and autobiographical memory, not covered in the present paper), but this remains to be seen. Similarly, it would be interesting to test whether these subjects discount future outcomes more heavily as predicted by Wegener et al. (2008). Additionally, patients with autism, who are known to demonstrate impairments in mentalizing (Frith and Frith 2006), would also be predicted to discount the future more than healthy subjects in an intertemporal choice task.

Using fMRI, a study could be conducted that would include two experimental paradigms within the same group of healthy subjects: an intertemporal choice task and a mentalizing task. The intertemporal choice task would be designed similar to the task in Wegener et al. (2008) in order to estimate how the activity around the temporoparietal junction decreases with delay time, while the mentalizing task would be designed similar to the task in Castelli et al. (2000), Schultz et al. (2003), and Martin and Weisberg (2003)—where moving geometric figures are used to elicit activity in areas associated with mentalizing, including the temporoparietal junction. The degree of inherent intentionality in moving shapes decreases as the complexity goes up, so we would expect there to be decreasing activation in the temporoparietal junction as the level of complexity increases. If there were additionally a correlation across individuals between the steepness of this parametric response and the analogous steepness in the same area corresponding to time discounting preferences, then this result would provide more confirmation for the connection between mentalizing about others (represented here by the shapes) and mentalizing about our future selves. Such a study also has the potential to reveal not only neural similarities, but also possible dissociations between mentalizing about others and about future selves.

Finally, recall the modified dictator game described at the beginning of this section. As pointed out there, merely observing choices is insufficient to draw any conclusions about the underlying processes at work. Here is where neuroimaging techniques such as fMRI, as in the previous section, truly come to the fore. Such a task could be carried out while subjects were being scanned, allowing direct comparison of brain activity in various regions during decision-making in each case. Of course, finding overlapping areas does not prove that precisely the same system is at work, but it is highly suggestive that similar cognitive processes are utilized. Similarly, the prediction task described above could be performed while subjects undergo scanning, allowing an analogous comparison to be made in addition to the purely behavioral outcome comparison.

Policy Implications

The research agenda—past, present, and future—that has been described here fits within the recent field of neuroeconomics. This new discipline has been defined as experimental, empirical, and theoretical analyses of the decisionmaking process that take into account the physical (and especially the neurological) embodiment of the decisionmaker.[3] It combines neuroscience, economics, and psychology, but also touches on the concerns of philosophy, medicine, and public policy. It is relatively straightforward to see how such an intersection can

[3] See Jamison (2008), as well as the other papers in that same journal issue, for further discussion of the definition and potential benefits of neuroeconomics.

benefit cognitive neuroscientists, since economists and psychologists have a wealth of tasks and theories about human behavior, each of which can now give rise to a search for corresponding brain activation and other neurophysiological responses.

It is less obvious, but still eventually evident, how to reverse this flow: what are the benefits to social scientists from observing embodied behavior? That is, if the ultimate goal is to study, understand, and perhaps predict choices themselves, then what is the value added of information about the choice process? In some ways, the question answers itself: to understand and predict choices, knowing something about the process that leads to those choices, is likely to prove extremely useful. In particular, suppose that one wishes to determine whether it will be empirically more efficient to model two putatively similar types of decisions (for example, risk and ambiguity; or in the present case, time preferences and theory of mind) together or separately. With sufficient choice data, there may be no need to look at the actual mechanisms—but using the biological data may be a far faster route to the same conclusion. Similarly, there are many types of choices (for example, involving house purchases, severe medical interventions, or the environment) that are simply infeasible to study via controlled non-hypothetical laboratory decisions. Given the fact that survey or self-reported responses are viewed as less inherently trustworthy, there is a clear rationale for augmenting them with concurrent neurological data in order to determine at least whether the decisionmaking process is proceeding in a manner known to be valid and consistent in other circumstances.

Finally, we turn to possible implications of neuroeconomics, particularly this link between other agents and future selves, for policy at both the social level, whether governmental or institutional, and at the individual level. We shall briefly discuss each in turn, but first a reminder of our main conclusion: humans seem to use the same brain systems to think about themselves in the future as they do to think about other conscious agents. By "think about," we refer to empathy (not just affinity), the mentalization of intentionality, and the prediction of behavior.

At the social level, there has recently been much discussion in both the scientific and popular press about the appropriate role for behavioral economics in government policy.[4] The central debate revolves around whether recent evidence suggesting that people may choose sub-optimally (even with respect to their own preferences) is enough to imply a role for government regulation. Such evidence provides a tension between the conservative or libertarian view (that individuals should generally be left alone to act in their own best interest) and the liberal view (that government is necessary to help some individuals, especially when exogenous factors impact either their material circumstances or their perceived set of available options). One intermediate proposal has been to use *nudges* or strategically chosen defaults—leaving the final choice up to the individual but making it more likely that he or she will choose the 'preferred' option.[5] This is moderately compelling when it applies, but there is no default position for most personal choices. For example, should an individual quit smoking now, next week, or not at all? In those cases, such an approach provides no guidance because it does not address the thornier issue of whether or not there is ever sufficient reason to overrule current individual preferences.

One might be tempted to say that these sorts of passive approaches are more natural, and that an active intervention (in this case, effectively placing a higher value on future members of society than individuals themselves seem to) would be unprecedented. And yet there are many situations where society explicitly chooses not to allow individuals to freely determine their own course of action: children, drug addicts, psychotic patients, and so on. Sometimes this entails a parent or other relative having decisionmaking authority, but sometimes it entails a court-appointed guardian. Even when it is, say, a parent and his or her young children, the state often decrees that certain behaviors (such as primary education) are mandatory—for the sake of the child (or rather the adult that the child will grow into and who will reap the benefits of the education). In the present context, the idea would be that the state

[4] For example, see Leonhardt (2008) in the *New York Times* on behavioral economists in the new Obama administration.
[5] Thaler and Sunstein (2003) refer to this as *libertarian paternalism*, while Loewenstein et al. (2007) refer to a variant of it as *asymmetric paternalism*.

might be obligated to 'protect' future selves from the uncaring (relatively speaking) choices made by current selves.

Society does not allow individuals to excessively harm their neighbors, and the empirical findings discussed here (that in the brain, one's future self is the same as a neighbor) suggest that perhaps it ought to likewise protect the welfare of future selves. As with relations between neighbors, this does not imply that the government or a panel of experts would (or should) tell anybody what choices to make or exactly how to behave. Rather, the government might make certain behaviors harder (or more expensive) in order to counterbalance the underlying tendencies. Neuroscience can provide a scientific foundation for why we as a society might want to do this, and it can inform the debate as to when, to what extent, and how we should collectively engage in trading off present freedom of choice against benefits to future selves.

Of course, this would require an explicit social decision that future members of society have rights that are commensurate (or nearly so) with current members. This may or may not be the social consensus, but if it were agreed to explicitly, then policies that take the multiple self model into account could be considered. We do not propose a comprehensive or even representative list of those policies, but a few examples may illustrate the possibilities:

- Prohibiting suicide—which is already the case in many countries, but for different reasons;
- Prohibiting drugs or narcotics only if they are highly addictive;
- Requiring a minimal financial reserve for large expenditures or donations (that is, preventing individuals from spending or giving away all of their assets);
- Subsidizing preventive health above and beyond the financial rationale;
- Subsidizing saving, especially long-term holdings;
- Waiting periods for everything from marriage to career choices.

Note that these possibilities do not uniformly restrict choices: prohibiting suicide increases future choices (which is the point); prohibiting only highly addictive drugs would be more permissive than current policies. Rather they suggest an overall approach. To many people, the final three may be the most attractive, since they involve a 'soft' paternalism[6] that guides but does not limit. Note also that we are in no way suggesting that neuroscience can be used to predict specific future choices of individuals, but merely that it tells us something about how the brain models future selves. In particular, the brain treats them as 'other' instead of as 'I,' and we therefore cannot straightforwardly use the traditional rationales for non-intervention.

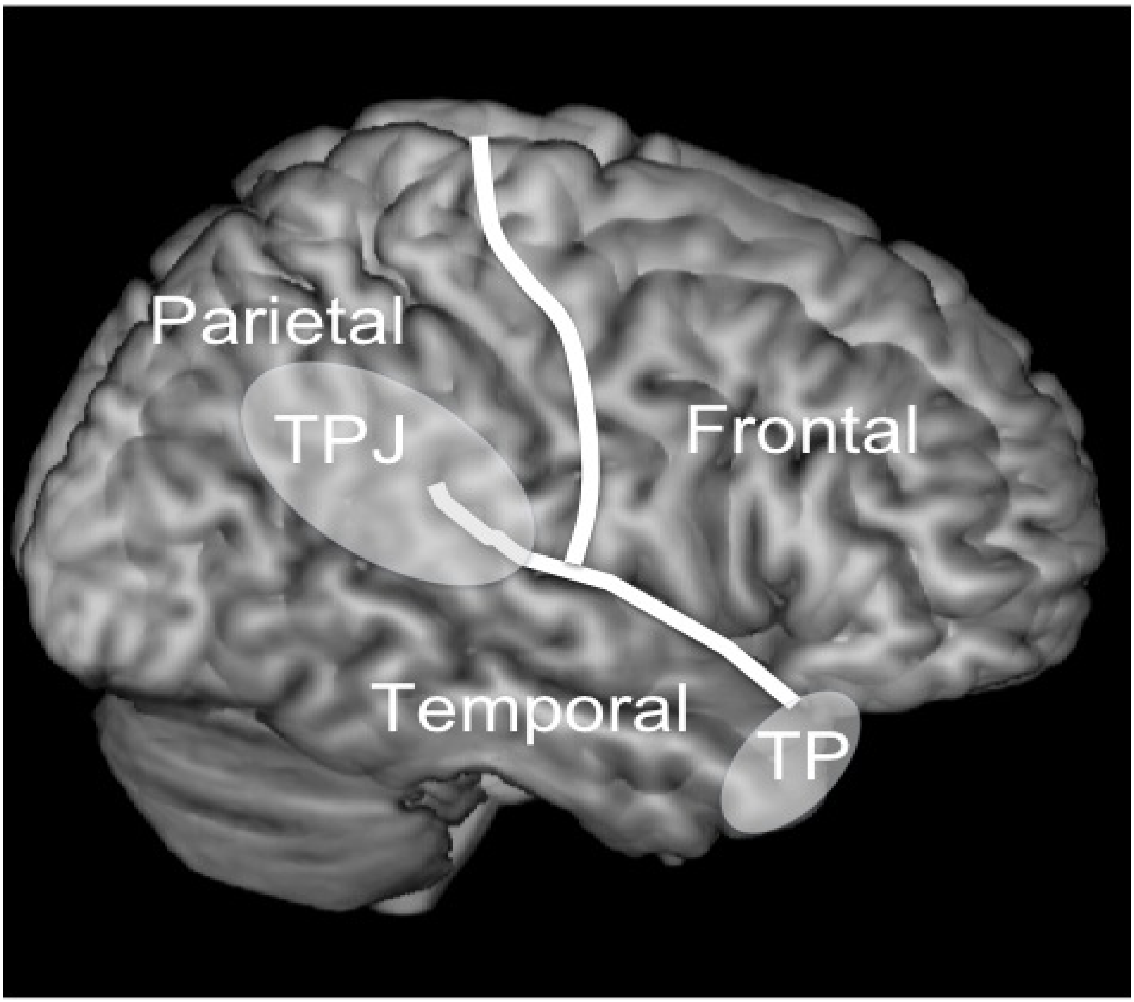
We reiterate that we are not explicitly suggesting such policies; they would depend both on further scientific work, as described previously, and on broader social decisions regarding the relative rights of future selves (not future generations, as is more commonly debated). The goal of this approach is to make these sorts of tradeoffs explicit and to provide the scientific input, which is necessary but not sufficient for sound policymaking. Since future selves cannot speak for themselves, this raises the question of who gets to speak for them, and hence any such policies face unusual constraints and need to be weighed especially carefully.

Nevertheless, we believe that this view of future selves is fundamentally different from the prevailing one, that it is based on sound data from multiple sources, and that it has deep implications for policy that should be openly discussed. See Farah (2002) for an excellent discussion of broader ethical issues in neuroscience.

It is also clear that if one takes seriously the idea of multiple selves over time, then there are also individual responses that do not require government intervention of any sort. These can range from simply being more attuned to discrepancies between past and present (leading to better predictions of one's future actions or future welfare); to playing the equivalent parental role not only with children but also with friends and relatives; to voluntarily joining or creating institutions that mimic some of the proposed government policies listed above. These ideas will surely be expanded and refined as the evidence base grows and the debate continues.

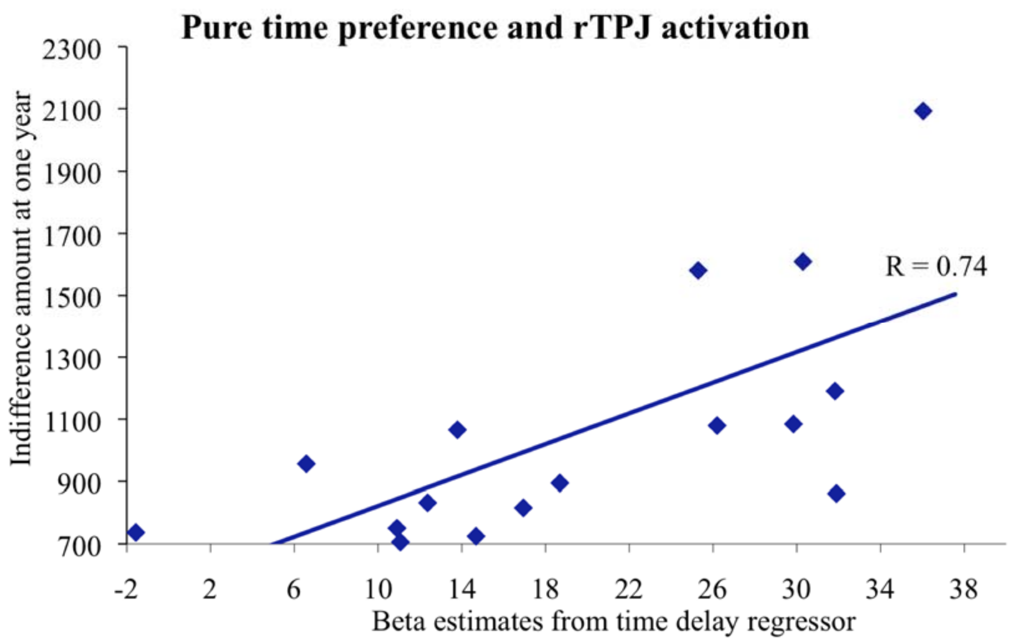
[6] We hesitate to use the word "paternalism" at all, because the entire idea of the paper is that the intervention is on behalf of a different agent (the future self), but this is admittedly indistinguishable in practice from what others may call by that name.

Figure 1



Temporoparietal junction and temporal pole illustrated on right hemisphere on a rendered brain template. Both Temporoparietal junction (TPJ) and temporal poles bilaterally have been found increasingly activated in mentalizing about others (Theory of Mind)(review: Frith and Frith 2006), mentalizing about oneself in the future (prospection) (review: Spreng et al. 2008), and decreasingly activated as a function of time delay in an intertemporal choice task as seen in Figure 2 (Wegener et al. 2008)

Figure 2



Significant correlation between inverse slope of BOLD activation at one year in right temporo-parietal junction (rTRPJ) [66 -45 33] and estimated amount at indifference point at one year within individual subjects (n=16)] (r=0.74; p=0.001). Horizontal axis shows beta estimates from the time-delay regressor in the fMRI analysis, and vertical axis shows indifference amount in DKK (Wegener et al. 2008).

References

Addis, D. R., A. T. Wong, and D. L. Schacter (2007) "Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration." *Neuropsychologia* **45**: 1363–1377.

Bernheim, B. D., and A. Rangel (2004) "Addiction and cue-triggered addiction processes." *American Economic Review* **94(5**): 1558–1590.

Boettiger, C. A., J. M. Mitchell, V. C. Tavares, M. Robertson, G. Joslyn, M. D'Esposito, and H. L. Fields (2007) "Immediate reward bias in humans: Fronto-parietal networks and a role for the catechol-O-methyltransferase 158(Val/Val) genotype." *JournaL of Neuroscience* **27**: 14383–14391.

Botzung, A., e. Denkova, and l. Manning (2008) "Experiencing past and future personal events: Functional neuroimaging evidence on the neural bases of mental time travel." *Brain & Cognition* **66**: 202–212.

Boyer, P. (2009) "Evolutionary economics of mental time travel?" *Trends in Cognitive Sciences*, in press.

Buckner, R. L. and D. C. Carroll (2007) "Self-projection and the brain." *Trends in Cognitive Sciences* **11**: 49–57.

Castelli, F., F. Happe´, U. Frith, and C. D. Frith. (2000) "Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns." *NeuroImage* **12**: 314–325.

D'Argembeau, A., G. Xue, Z.-L. Lu, M. Van der Linden, and A. Bechara, A. (in press) "Neural correlates of envisioning emotional events in the near and far future." *Neuroimage*.

Decety, J., and J. Grezes (2006) "The power of simulation: Imagining one's own and other's behavior." *Brain Research* **1079**: 4–14.

Ding, M. (2007) "A theory of intraperson games." *Journal of Marketing* **71:** 1–11 (April).

Farah, M. J. (2002) "Emerging ethical issues in neuroscience." *Nature Neuroscience* **5(11)**: 1123–9.

Frank, Bjorn (1996) "The use of internal games: The case of addiction." *Journal of Economic Psychology* **17:5**: 651–-660.

Frith, U., and C. D. Frith (2006) "The neural basis of mentalizing." *Neuron* **50(4)**: 531–534.

Frith, U., A. Leslie, and J. Morton, J. (1991) "The cognitive basis of a biological disorder: Autism." *Trends in Neurosciences* **14**: 433–438.

Fudenberg, D., and D. K. Levine (2006) "A dual self model of impulse control." Harvard Institute of Economic Research Discussion Paper #2112.

Gallagher, H.L. and C. D. Frith (2003) "Functional imaging of 'theory of mind'." *Trends in Cognitive Sciences* **7**: 77–83.

Gil-da-Costa R., A. Braun, M. Lopes, M. D. Hauser, R. E. Carson, P. Herscovitch, and A. Martin (2004) "Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque." *Proceedings of the National Academy of Science U S A.* **101(50)**: 17516–-2.

Glimcher, P. W., J. W. Kable, and K. Louie (2007) "Neuroeconomic studies of impulsivity: Now or just as soon as possible." *American Economic Review* **97(1**): 142–147.

Hume, D. (1739; republished 2000) *A treatise of human nature*. Oxford, United Kingdom: Oxford University Press.

Jamison, J. C. (2008) "Well-being and neuroeconomics." *Economics and Philosophy* **24**: 407–418.

Johnson, E. J., G. Haubl, and A. Keinan (2007) "Aspects of endowment: a query theory of value construction." *Journal of Experimental Psychology: Learning Memory and Cognition* **33**: 461–474.

Kable, J. W., and P. W. Glimcher (2007) "The neural correlates of subjective value during intertemporal choice." *Nature Neuroscience* **10**: 1625–1633.

Krusell, P., and A. A. Smith, Jr (2003) "Consumption-savings decisions with quasi-geometric discounting." *Econometrica* **71(1):** 365–375 (January).

Laibson, D. (1997) "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics* **112(2)**: 443–477.

Levy, B. J., and M. C. Anderson (2002) "Inhibitory processes and the control of memory retrieval." *Trends in Cognitive Sciences* **6**: 299–305.

Leonhardt, D. (2008) "Budgets behaving badly." *New York Times*, Economic Scene. (Dec. 3) B1.

Loewenstein, G. (1988) "Frames of mind in intertemporal choice." *Management Science* 200–214.

Loewenstein, G. T. B., and K. G. Volpp (2007) "Asymmetric paternalism to improve health behaviors" *Journal of the American Medical Association* **298(20)**: 2415–2417.

Loewenstein, G. and D. Prelec (1992) "Anomalies in intertemporal choice: Evidence and an interpretation." *Quarterly Journal of Economics* **107**: 573–597.

Martin, A. and J. Weisberg (2003) "Neural foundations for understanding social and mechanical concepts. "*Cognitive Neuropsychology* **20(3–6)**: 575–87.

McClure, S. M., D. L. Laibson, G. Loewenstein, and J. D. Cohen (2004) "Separate neural systems value immediate and delayed rewards." *Science* **306**: 503–507.

McClure, S. M., K. M. Ericson, D. L. Laibson, G. Loewenstein, and J. D. Cohen (2007) "Time discounting for primary rewards." *Journal of Neuroscience* **27**: 5796–5804.

Mischel, W., Y. Shoda, and M. L. Rodriguez (1989) "Delay of gratification in children." *Science* **244***:* 933–938.

Mitchell, J. P. (2008) "Activity in right temporo-parietal junction is not selective for theory-of-mind." *Cerebral Cortex* **18**: 262–271.

Mobbs, D., R. Yu, M. Meyer, L. Passamonti, B. Seymour, A. J. Calder, S. Schweizer, C. D. Frith, and T. Dalgleish (2009) "A key role for similarity in vicarious reward." *Science* **324**: 900.

O'Donoghue, T. and M. Rabin (1999) "Doing it now or later." *American Economic Review* **89(1**): 103–124.

Poldrack, R. A. (2006). "Can cognitive processes be inferred from neuroimaging data?" *Trends in Cognitive Sciences* **10**: 59–63.

Premack, D. and Gary W. (1978) "Does the chimpanzee have a theory of mind?" *Behavioral and Brain Sciences* **4**: 515–526.

Read, D., and N. L. Read (2004) "Time discounting over the lifespan." Working Paper No: LSEOR 04.63

Samson, D., I. A. Apperly, C. Chiavarino, and G. W. Humphreys (2004) "The left tempero-parietal junction is necessary for representing someone else's belief." *Nature Neuroscience* **7**: 449–500.

Sharot, T., A. M. Riccardi, C. M. Raio, and E. A. Phelps (2007) "Neural mechanisms mediating optimism bias." *Nature* **450**: 102–105.

Schultz, R. T., D. J. Grelotti, A. Klin, J. Kleinman, C. Van der Gaag, R. Marois, and P. Skudlarski (2003) "The role of the fusiform face area in social cognition: implications for the pathobiology of autism." *Philosophical Transactions of the Royal Society London* B **358**: 415–424.

Singer, T. (2006) "The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research." *Neuroscience and Biobehavioral Reviews* **30:** 855–863.

Spreng, R.N., R.A. Mar R.A., and A. S. N. Kim. (2008) "The common neural basis of autobiographical memory, prospection, navigation, theory of mind and the default mode: A quantitative meta-analysis." *Journal of Cognitive Neuroscience* [Epub ahead of print].

Szpunar, K. K., J. M. Watson, and K. B. McDermott (2007) "Neural substrates of envisioning the future." *Proceedings of the National Academy of Sciences of the United States of America* **104**: 642–647.

Thaler, R., and S. Hersh (1981) "An economic theory of self-control." *Journal of Political Economy* **89(2)**: 392–406.

Thaler, R., and C. Sunstein (2003) "Libertarian paternalism." *American Economic Review* **93(2**): 175–179.

Tomlin, D., M. A. Kayali, B. King-Casas, C. Anen, C. F. Camerer, S. R. Quartz, and P. R. Montague (2006) "Agent-specific responses in the cingulate cortex during economic exchanges." *Science* **312**: 1047–1050.

Uddin, L.Q., J. T. Kaplan, I. Molnar-Szakacs, E. Zaidel, and M. Iacoboni (2005) "Self-face recognition activates a frontoparietal 'mirror' network in the right hemisphere: an event-related fMRI study" *NeuroImage* **25**(*3*): 926–935.

Vieille, N. and J. W. Weibull (2009)"Multiple solutions under quasi-exponential discounting." *Economic Theory* **39**: 513–526.

Weber E.U., E. J. Johnson, K. F. Milch, H. Chang,J. C. Brodscholl, and D. G. Goldstein (2007) "Asymmetric discounting in intertemporal choice—a query-theory account." *Psychological Science* **18**: 516–23.

Wegener, J. S., K. Madsen, M. Christensen, and J. Jamison (2008) "Parieto-temporal activation modulates pure time preference." presented at the *Conference of Neuroeconomics 2008* (May), Copenhagen, Denmark; currently under review.