To appear in L. Nadel & W. Sinnott-Armstrong (eds.) *Benjamin Libet and Agency*. Oxford: Oxford University Press.

Draft, do not quote without permission.

What are Intentions?

Elisabeth Pacherie & Patrick Haggard

Elisabeth Pacherie
Institut Jean Nicod, EHESS, DEC-ENS, CNRS, Paris, France

Patrick Haggard
Institute of Cognitive Neuroscience, University College London, London UK

Address for correspondence :
Elisabeth Pacherie
Institut Jean Nicod
Département d'Etudes Cognitives
Ecole Normale Supérieure
29, rue d'Ulm
75005 Paris – France
Tel: +33 1 44 32 26 89
Fax: +33 1 44 32 26 99
Email: pacherie@ens.fr

Acknowledgements:

EP was supported by a project grant from Agence Nationale de la Recherche (ANR-07-NEURO-039-03.

PH was supported by a Royal Society Research Fellowship, a Leverhulme Trust Research Fellowship and a project grant from ESRC.

Introduction

Neuroscientific work on intentions and voluntary actions has tended to focus on very short time scales, immediately before movement onset. As a prime example, the intentions investigated by Benjamin Libet are states that are first consciously experienced on average 200ms before action onset. Libet's experiments showed that these conscious intentions were reliably preceded by a few hundred milliseconds by a negative brain potential, the so-called ‘readiness potential’. The existence of this antecedent unconscious brain activity indicated that the action was initiated unconsciously rather than by the conscious intention. This led Libet to the conclusion that we do not have full-blown “free will”. However, he attempted to salvage a limited form of free will by suggesting that although we cannot consciously initiate actions, we can still consciously veto them in the 200ms interval between conscious intention and action onset. Libet's conception of free will and his interpretation of his results have been widely discussed and criticized.

Here, we take as our starting point one of these lines of criticism, voiced notably by Shaun Gallagher (2006). Gallagher argues that it is misguided to attempt to frame the question of free will at the time scale and in terms of the very short term motor intentions and control processes Libet considers. Rather, free will involves temporally extended deliberative processes and applies to intentional actions considered at levels of description typically higher and more abstract than descriptions in terms of motor processes and bodily movements. In earlier work, one of us (Pacherie, 2008) proposed a three-tiered hierarchical model of intentions, the DPM model, distinguishing distal or prospective intentions, proximal or immediate intentions, and motor intentions; the other (Haggard, 2008) offered a naturalized model of human volition involving a set of decision-making processes concerned with whether to act, what to do (and how) and when to act. If Gallagher is right about the temporal and intentional framework relevant for the exercise of free will, a discussion of free will must at least include not only the contribution of intentions to the final process of action initiation itself, but also the anterior decision processes that take place at the level of prospective intentions.

1. Immediate Intention and Action Initiation

Providing a satisfactory definition of intention is notoriously difficult. In this chapter, we assume that intention is a mental state, which may be associated with particular brain states. But what *kind* of mental state is an intention? We suggest that intentions have two distinguishing features. First, they are accessible to consciousness. Second, they bear some relation to subsequent action. This relation could be distinctive for two reasons: a causal reason or a content reason. Let us take a physical movement of the body (I raise my arm) as a paradigm of action. The causal reason suggests that the intention (I intend to raise my arm) is simply the mental state that causes the action of lifting my arm (Wittgenstein, 1953). Intentions thus explain why actions occur, and serve as the guarantors of volition. This view is clearly vulnerable to sceptical attack: folk psychology may find it convenient to have some appropriate explanation of a person's actions, and the concept of intention could be designed to fulfil this purpose. The fact that intentions do a good job of explaining actions does not therefore constitute evidence that they are a bona fide mental state.

The content argument suggests the content of the intention ("I will raise my arm") is somehow linked to the specific details of the arm-raising action. This view makes clearer predictions about what might constitute an intention. For example, if I perform two different actions, raising my left arm on one occasion and my right arm on another, the intentions for each action should have different contents, capable of explaining which arm is used for the action in each case. The content of intention should be discriminative, in the sense that it should predict specific details of action. The content argument emphasises the continuity between decision and intention: when someone decides to do A rather than B, they may develop an intention whose specific content will relate to A or to B. A number of neuroscientific studies have attempted to decode the brain processes predicting the specific content of a subsequent action (Soon et al, 2008; Haggard and Eimer, 1999). This level of motor content would typically be generated once the specific situation and context of action are established, and only immediately before action initiation. Because intention, viewed in this way, is very close to the details of motor execution, we use the term "immediate intention" to refer to it.

Interestingly, although Libet's work (Libet et al., 1983) occupies a central role in modern scientific work on intention, he himself appeared to avoid the word. On the one hand, he speaks of the "unconscious initiation" of action. This refers to the set of brain processes that ultimately give rise to muscular movement. The readiness potential generated by the frontal motor areas of the brain is a convenient marker that these processes have begun, but Libet avoids making the simplistic claim that the onset of the readiness potential simply constitutes initiation. On the other hand, the conscious experience of immediate intention (W judgement) occurs several hundred milliseconds after the readiness potential onset, and only slightly before movement itself. If the W judgement is taken as the marker of conscious intention, then, our conscious intentions cannot be the cause or explanation of our actions, because intention follows neural initiation of action, rather than precedes it.

But is the W judgement really a marker of immediate conscious intention? Libet himself speaks of an "urge to act". Participants are supposed to report the moment when this urge begins. This is clearly one of the weaker points of the experimental method. How do participants know what they are supposed to report? Could the instruction to report urges somehow suggest to the subject that they should have a specific experience of immediate intention that would otherwise remain unconscious? Could the instruction suggest to subjects the need to report a moment slightly before action, even if they have no distinctive conscious experience at that moment? Participants might interpret the instructions in such experiments as “Behave as if you had free will, and make your reports of intention consistent with this concept of free will”. If this were the case, then such experiments could not separate the influence of folk psychology from any genuine mental state of intention, making them vulnerable to sceptical attack, or even scientifically worthless.

Evidence from direct cortical stimulation

Clearly, experimental manipulations of intention which do not depend on instructions, and therefore avoid the worst problems of suggestion, are highly desirable. Perhaps the most informative data come from reports of direct cortical stimulation prior to neurosurgery for epilepsy. Methodologically, these data clearly differ from psychological experiments relying on participants’ understanding of instructions. In fact, no instruction is given at all: the patient’s behaviour during stimulation is observed, and they are invited to report anything that they feel. Little detail is generally given about *how* the reporting is done. Few neurosurgical studies seem to address the problems of experimenter-led suggestion and response bias, for example, by including catch trials without stimulation. Nevertheless, these data have particular significance for the psychology of intention, and are therefore worth examining in some detail.

Direct stimulation data broadly support a distinction between initiation of action and conscious immediate intention. In particular, we shall argue that direct stimulation of the pre-Supplementary Motor Area is accompanied by an anticipatory conscious experience of immediate intention. In contrast, direct stimulation of the deeper Cingulate Motor Area produces a strong *motivation* to perform a specific action, and can trigger action initiation, but without any particular *specific* conscious experience prior to action. In the neurosurgical literature, and in Libet’s work also, the word ‘urge’ is widely used. We argue that the same word is used with two quite different meanings, which have been unnecessarily confounded. On the one hand, an urge involves a conscious experience of being about to act. On the other hand, an urge involves a feeling of compulsion, or having to. We suggest these two components are localised to the pre-SMA and the CMA respectively. Rather than the general term ‘urge’ we suggest that the terms immediate intention and motivation to act might be more appropriate.

Pre-SMA stimulation can evoke a state resembling immediate intention

The awake patient reports a subjective experience or ‘urge to move’ during stimulation of characteristic cortical regions, notably the supplementary motor area. The study closest to our interest is that by Fried et al (1991). The paper reports responses to stimulation through intracranial grids over the mesial frontal cortex. In one patient, several reports of 'urge' were obtained following low-amplitude stimulation over the supplementary motor area. The responses typically referred to a specific contralateral body part, as in 'urge to move the right elbow'. In some trials, different verbal formulas appear: 'need to move…', 'feeling as if movement were about to occur'. At higher stimulation intensities, actual movements were often evoked. The authors comment that the actual movement evoked was not necessarily commensurate with the urge. However, urge and movement at least referred to the same limb in the majority of trials reported for this patient.

The ability to evoke by external intervention a mental state that appears close to conscious intention is intriguing. However, several important methodological questions remain. How general are these sensations: they receive prominent attention in the report of one case, but it is unclear whether they were investigated and found to be absent, or merely not investigated, in the remaining cases? What phenomenal experience does the stimulation cause? Beyond the frequent use of the word 'urge' there is little information on phenomenology. One particular concern would be whether the experience reported as 'urge' is truly an anticipatory experience of central origin, and occurring in advance of movement. Could 'urge' actually reflect subtle muscle contractions caused by low-intensity stimulation, which lacked the strength required to produce observable movement? Alternatively, could 'urge' reflect a sensory experience, like the 'tingling' sensation frequently reported following stimulation at sites close to those provoking 'urge' (Fried, 1991)? The preSMA is known to receive sensory afferent input, probably after initial processing in somatosensory cortical areas (Mima et al, 1999). In conditions such as Tourette's Syndrome and Restless Legs Syndrome, the urge to move is strongly associated with, or is simply described as, a *sensory* quality localised in specific body parts, and relieved by movement of those body parts. If urges were essentially sensory in nature, they clearly would not be a good model for conscious intention. Interestingly, however, a recent review of a series of 52 patients who underwent electrical stimulation suggests sensory experiences are not a normal feature of preSMA stimulation, being recorded in only a single instance (Chassagnon, 2008). In fact, they were much more common following stimulation of the posterior portion of the CMA. It seems likely that preSMA stimulation produces a specific conscious experience, distinct from both stimulation-evoked sensation and from peripheral sensation. This experience, like immediate intention, is motorically-specific, and linked to an impending action.

CMA stimulation produces motivated but automatised actions

In fact, the stimulation of the CMA, and particularly of the region of the cingulate sulcus immediately below the pre-SMA, seems to correspond more closely to Libet's 'unconscious initiation of a … voluntary act'. Chassagnon et al. (2008) report 4 instances where CMA stimulation elicited reaching and grasping behaviours, 'as if the patients were groping around and handling a small object in the dark'. There is no specific report or evidence of urge *prior* to actual movement. In an extended report of one patient in this series (Kremer et al., 2001) shows that these behaviours had a compulsive, irresistible quality. This patient had a strong drive to perform the movement once stimulation began, making scanning eye movements and exploratory arm movements to identify a potential target for grasping. The patient is described as having an "urge to grasp something". However, it remains hard to locate this feeling of urge within the chain of events linked to the action. In particular, no quantitative data is given on two details which are of primary importance for the psychology of intention: the delay between stimulation onset and movement onset, and the delay between stimulation onset and any sense of 'urge'. We suggest that this patient showed 'urge' in the motivational sense during CMA stimulation, but they did not experience the kind of anticipatory conscious awareness characteristic of immediate intentions.

A more extensive study of actions evoked by CMA stimulation in 83 epileptic patients was reported by Bancaud et al. (1976). Stimulation generally produced an increased state of arousal and attentiveness, often at low stimulation intensities. This was interpreted as a non-specific form of attention to action. At higher stimulation intensities, a range of coordinated manual, buccal and oculomotor actions were produced. Interestingly, if an object were given to the participant during stimulation, it would evoke complex series of object-appropriate movements. For example, when one patient was given a cigarette, they lit and smoked it in a compulsive manner, stopping smoking when stimulation ceased, and restarting when stimulation restarted. In other cases, patients compulsively ate food they were offered, or brought objects to the mouth and sucked them. Again, ceasing stimulation caused the action to end. When the experimenters physically restrained the patient's arms, the patient often strove to continue the action, especially at greater stimulation intensities. This sustained drive to achieve the action is not merely a matter of maintaining motor output in the face of perturbation, since in one case the patient transferred an object repeatedly between the hands to overcome the experimenter's interference.

What did the patients experience? While Bancaud et al do not address this point systematically, the general attitude of the patients towards their own evoked actions appeared indifferent. Patients acknowledged the action they had performed immediately afterwards, but did not generally give specific reasons why they performed it, nor did they appear surprised by actions which might prima facie seem strange. On questioning the next day, the patients did not find their actions under stimulation in any way surprising or unusual. One way of interpreting this unusual phenomenology of action would suggest that the CMA drives actions, without any reference to conscious intentions, desires, or reasons for action. For example, a patient presented with a fruit in the absence of stimulation would merely hold it. But once stimulated, the patient would grasp and eat the fruit for as long as the stimulation lasted. This compulsive eating was not part of a normal desire for food, since it ceased with the end of stimulation.

In summary, CMA stimulation transiently induced a syndrome similar to utilisation behaviour (Lhermitte, 1983). The overall impression is of a CMA role in motivating and driving behaviour, but not in anticipating, or monitoring or adjusting it to circumstances, nor in providing a conscious experience of an impending action. The state evoked by CMA stimulation therefore appears to be closer to a motivational drive than to an intentional decision. The evoked actions appear to happen **to** the patient, but are quite decoupled from their conscious mental life, and play no role in it. This explains why the patient does not produce convincing or detailed reasons to explain why they occurred.

A model of frontal contributions to intentional action

--- INSERT FIGURE 1 ABOUT HERE ---

One simple model, which could encompass Libet's (1983) concept of conscious intention, is shown in figure 1. Selection between competing alternative actions that are currently available might occur in dorsolateral prefrontal cortex (Rowe et al., 2000). This process may involve conscious thought about the range of action alternatives, but only at the level of abstract action possibilities. The DLPFC selects the appropriate action, and forwards the decision to two separate cortical motor areas to implement it. On the one hand the decision is sent to the CMA, which provides a motivational drive to initiate the action. On the other hand the decision is sent to the pre-SMA, which provides a stage of flexible, contextual modulation of internally-generated action, weaving the selected action into the ongoing flow of behaviour and experience. This flexibility is required since a behaviour may be appropriate in one context but not in another: even a strongly motivated action can and should sometimes be stopped or delayed. Pre-SMA therefore provides contextual arbitration, according to which a drive may be developed into an impending action plan, or alternatively inhibited. This contextualising role of preSMA can explain three specific findings from the neurophysiological literature which may otherwise be hard to explain (see Haggard, 2008 for a detailed review). First, cells in the preSMA appear to play a key role in integrating single actions into coordinated superordinate sequences of behaviour. Second, lesions in this area produce compulsive action tendencies, reminiscent of the automatised reaching and grasping evoked by CMA stimulation. Third, the preSMA plays a key role in arbitrating involving conflict between the various alternative actions that could be consistent with a given situation. The pre-SMA is therefore involved not in the raw drive to action, but in reconciling action drives to current contexts.

Interestingly, the conscious experience of immediate intention seems to involve the same circuits that contextually constrain action drives. The conscious 'urge' evoked by preSMA stimulation, and perhaps underlying W-judgements of intention in Libet-type experiments, would correspond to the moment of opening the gate between drive and motor action. The pre-SMA would then pass the contextualised action plan to the SMA-proper, M1 and CMA for execution. On this model, Libet is absolutely right that our actions are initiated unconsciously, by the normal functioning of the sensory and motor network of the cortex. The conscious experience of intention-in-action occurs when the prefrontal executive opens the gates between this network and motor executive areas, such as M1, so that the drive built up within this network can now proceed to appropriate action execution.

What are immediate intentions?

The discussion above allows us to to revisit our question "what are immediate intentions"? From a neural point of view, immediate intentions are conscious experiences of impending action, generated by the motor systems of the medial frontal cortex. From a psychological point of view, two important aspects of immediate intention are worth emphasising. First, immediate intentions are predictive, in the sense that they precede actions. Second, immediate intentions have an episodic, time-locked quality, rather than being abstract and semantic. Thus, the content of an immediate intention prefigures at least some of the specific motor details of the action itself. Immediate intentions are not linked to actions in a vague and general way, but in a motor-specific way (Haggard and Eimer, 1999), even in artificial cases such as pre-SMA stimulation (Fried et al., 1991). Put another way, immediate intentions incorporate the specific contextual detail, corresponding at least to the P-level and often to the M-level in the DPM hierarchy. An interesting conscious correlate of this episodic quality is the very integrated experience we have of our own voluntary action. Intention, action and goal are not experienced as separate disconnected events, but as a tight and integrated flow. In particular, intentional actions, but not involuntary movements, display an effect called 'intentional binding', whereby the experiences of action and effect are perceived as temporally compressed and bound together (Haggard et al., 2002; Haggard and Cole, 2007), as if part of a single episode.

2. *Prospective Intentions*

We share with other animals the capacity to act purposefully, but we also regularly make more or less complex plans for the future and our later conduct is guided by these plans. We are, in Michael Bratman's words, planning agents and this planning ability appears to be distinctively human. People can, and frequently do, form intentions focussed on actions that may occur years or even decades later. Intentions to choose particular careers, to become prime minister, or to choose a destination for next year’s holiday all offer examples. The length of time-scale associated with intentions is effectively unlimited. These long-range intentions seem effectively connected with short-range intentions, and therefore with action itself. General intentions formed at one time point cascade into much more detailed intentions prior to action execution

However, almost nothing is known about how these long-range, prospective intentions connect to immediate, short-term intentions. Indeed, experimental studies of voluntary action deal hardly at all with the concept of prospective intention. On one view, the prospective intention in such studies consists in the participant’s decision to participate in the experiment in the first place, and thus lies beyond what can be measured in the experimental setting itself.

We start this section with a brief review of Bratman's influential account of prospective intentions (or as he calls them future-directed intentions), what their main characteristics are and what makes it useful to have them. We then turn to the issue what kind of cognitive processes are involved in the formation of prospective intentions and how these relate to the processes involved in immediate intentions.

Bratman on intentions

Bratman's account of future-directed intentions (Bratman, 1987) stresses the commitment to action that is a distinctive characteristic of intentions. When I intend today to go Christmas shopping tomorrow, I do not simply want or desire today that I go Christmas shopping tomorrow. Rather I am committed now to going shopping tomorrow. What exactly does this commitment involve? Bratman distinguishes two dimensions of a commitment to action: a volitional dimension and a reasoning-centered dimension. The volitional dimension concerns the relation of intention to action and can be characterized by saying that: "Intentions are, whereas ordinary desires are not, conduct-*controlling* pro-attitudes. Ordinary desires, in contrast, are merely *potential influencers* of action" (1987: 16). In other words, unless something unexpected arrives that forces me to revise my intention, my intention today to go shopping tomorrow will control my conduct tomorrow. The reasoning-centered dimension of commitment is most directly linked to planning. What is at stake here are the roles played by intentions in the period between their initial formation and their eventual execution. First, intentions have what Bratman calls a characteristic *stability* or inertia: once we have formed an intention to *A*, we will not normally continue to deliberate whether to *A* or not. In the absence of relevant new information, the intention will resist reconsideration, we will see the matter as settled and continue to so intend until the time of action. Intentions are thus *terminators of practical reasoning* about ends or goals. Second, during this period between the formation of an intention and action, we will frequently reason from such an intention to further intentions, reasoning from instance from intended ends to intended means or preliminary steps. When we first form an intention, our plans are typically only partial, but if they are to eventuate into action, they will need to be filled in. Thus, intentions are also *prompters of practical reasoning* about means. Finally, the volitional and reasoning-centered dimensions of intentions together account for another important function of prospective intentions, namely their role in supporting both *intrapersonal and interpersonal coordination*. Because intentions have stability, are conduct-controlling and prompt reasoning about means, they support the expectation that I will do tomorrow what I intend today to do tomorrow. Such expectations facilitate coordination. My intention to go Christmas shopping tomorrow supports my sister's expectation that I will, and she can go ahead and plan to join me in this shopping expedition. Similarly, I can go ahead and plan my activities for the day after tomorrow, on the assumption that by tomorrow evening I will be done with Christmas shopping.

As noted by Bratman himself, future-directed intentions have an air of paradox. They are typically stable but they are not irrevocable. Such irrevocability would be irrational, since things can change and our anticipation of the future is not infallible. This suggests that, having formed today an intention to do something tomorrow, I should persist in that intention tomorrow only if it would then be rational for me to form such an intention from scratch. But then, asks Bratman, why I should I bother deciding today what to do tomorrow? Isn't that just a waste of time?

Bratman offers several complementary answers to that challenge. They stem from the fact that we are epistemically limited creatures, with limited cognitive and time resources for use in attending to problems, gathering information, deliberating about options, determining likely consequences, and so on. There are several reasons our epistemic limitations make it useful for us to form prospective intentions. First, if our actions were influenced by deliberation only at the time of action, this influence would be minimal as time pressure isn't conducive to careful deliberation. Advance planning frees us from that time pressure and allows us to deploy the cognitive resources needed for successful deliberation. Second, intentions once formed have characteristic stability. They resist reconsideration. This doesn't mean we never reconsider. Intentions may be revoked. But as Bratman points out, revocability does not entail actual reconsideration. Unless new facts come to light, we will normally simply retain our intentions. Furthermore, in settling on a course of action, we have already rehearsed and weighted the considerations for and against that course of action. This prior rehearsal puts us in a better position to assess whether a new piece of information is actually relevant or not to our plans. If nonreconsideration is the default option, once an intention is formed the precious cognitive resources that were engaged in deliberation about ends are free be used for other tasks, including planning about means and ensuring both intra- and interpersonal coordination. To achieve complex goals, I must coordinate my present and future activities and coordinate with activities of other agents. If I now intend to the concert tomorrow night, I first need to procure a ticket and make sure I have a baby-sitter for the evening. If I were to leave it to the last minute to decide whether I go to concert tonight or not, I may well be frustrated to find out that tickets are sold out or that the baby-sitter is not available. Thus, the success of many of our actions depends on our ability to coordinate our own activities over time and to coordinate them with the activities of other agents. This coordination is best achieved if we plan ahead of time.

So-called Buridan cases provide a third reason for forming intentions. We may be forced to choose between options that we find equally desirable. I may have an equal desire to go to a concert or to go see a play tomorrow evening. But if I am to do either, I had better decide now among these options. For one thing, it may not be worth my while looking for further information in the hope of finding new reasons to decide between them, as the effort and time needed to gather further information may well exceed the potential benefits, say, enjoying the concert slightly more than I would have the play. Moreover, once again, intrapersonal and interpersonal coordination require that I reach a decision. I need to know whether to buy a ticket for the play or for the concert, and if I wish friends to join me, I need to let them know whether I intend to go to the concert or to go see the play.

Future-oriented cognition and mental time travel

Prima facie, it would seem that the reasons that make it useful for us to form prospective intentions also apply to other species. Limited cognitive resources and a need for coordination are not unique to humans. So why is it that we alone appear to exhibit such distinctive planning abilities? One obvious answer is that other species are even more limited then we are in their cognitive resources; a complementary answer is that how much need and use we have for planning also depends on the kind of environment we live in. There wouldn't be much use for planning in an environment that were completely unpredictable, for planning exploits regularities and in such an environment there would be none to exploit. On the other hand, in an environment both simple and reasonably predictable, there may be cheaper ways of coping than those involving advance planning. Suddendorf and Corballis (2007) describe several ways in which behavior may be future-oriented without involving a capacity to think about the future as such. First, future-directed behavior may be instinctual, as when, through natural selection, a species has evolved behavioral predispositions to exploit significant long term regularities. For instance, an animal can gather food for hibernation, although it has yet to experience a winter. Second, future-directed behavior may be driven by procedural learning, allowing an individual to track short-term regularities. For instance, through association, a conditioned stimulus can predict the future arrival of an unconditioned response and trigger a future-directed response. Third, future-directed behavior may exploit semantic memory about regularities, which provides the basis for inferential and analogical reasoning and allows learning in one context to be voluntary transferred to another. Procedural learning allows for greater flexibility than instinctual patterns of behavior, allowing behavior to be modulated by individual experience; semantic memories provides even greater behavioral flexibility as they can be triggered endogenously rather than being stimulus bound. Yet, the environment in which humans live is unique in both its ecological and its social complexity. Humans also have an extraordinary range of desires and motivations, going far beyond the basic drives and simpler desires present in other species. Dealing with this spectacular environmental, social and motivational complexity may require in turn forms of future-oriented cognition that exhibit unique flexibility and versatility.

A prime candidate for this more flexible form of future-oriented cognition is *mental time travel*, the faculty that allows a person to mentally project herself backward in time to relive past events or forwards to pre-live events (Suddendorf & Corballis, 1997, 2007; Suddendorf & Busby, 2003, 2005; Wheeler *et al.*, 1997). Mental travel in the past, known as episodic memory, has been intensively studied (e.g., Tulving, 1983, 2005). Mental travel into the future, in contrast, has only recently begun to draw attention. Recent work indicates that mental travel into the past and into the future are closely related, involving similar cognitive processes – a combination of episodic memory and imagination under executive control – and recruiting strongly overlapping neural systems (D'Argembeau & Van der Linden, 2006; Hassabis, Vann, & Maguire, 2007; Klein, 2002, Gerrans, 2007). Several researchers have argued that mental time travel into the future is a crucial cognitive adaptation, enhancing planning and deliberation by allowing a subject to mentally simulate and evaluate contingencies, and thus enhancing fitness, and that mental time travel into the past is subsidiary to our ability to imagine future scenarios (Dudai & Carruthers, 2005; Suddendorf & Corballis, 2007).

Mental time travel, whether into the past or into the future, involves episodic memory and inherits its two main characteristics. First, it is not about regularities but about constructing or reconstructing the *particularities of specific events*. Second, mental time travel involves *autonoesis*, i.e., awareness of a self as the subject of actual, recalled or imagined experience. But what are exactly the benefits that accrue from using mental time travel rather than simply reasoning from general knowledge stored in semantic memory in planning future actions? As we have seen, prospective intentions involve making a number of decisions. The intention is first formed when one reaches a decision about what to do. Once the intention is formed, one must still typically make a number of decisions about how to implement the chosen goal. Another important decision, not explicitly considered by Bratman, concerns when to act. What can mental time travel contribute to these what-decisions, how-decisions and when-decisions?

What-decisions

Not all what-decisions involve explicit conscious deliberation. Some decisions are pretty straightforward. If my closest friend invites me to her wedding, of course I'll accept the invitation and form the intention to attend the wedding. If, however, being on the job market, I am offered academic positions in two different universities, I might spend quite a while weighing the pros and contras of each option before reaching a decision. Yet, it may be that performing a logical cost-benefit analysis of the two options does not suffice to motivate me to choose one over the other even if this analysis yields a clear advantage for one of the options. Rather, I might have to imaginatively rehearse future experiences occupying one or the other position as part of the process of deliberation.

Patients with damage to the ventromedial prefrontal cortex (VMPFC) are often described as having impaired ability for planning and decision-making despite retaining intact capacities for explicit reasoning. Philip Gerrans (2007) argues that this impairment is best explained by a deficit in mental time travel. In his view, Damasio's somatic marker hypothesis (Damasio et al., 1991; Bechara et al., 1999), according to which the deficits of VMPFC patients result from a failure to link an implicit emotional response – a somatic marker – with an explicit representation of a situation, is deficient in two ways. First, it uses an account of emotions which explains salience and motivation in terms of valence and within this framework interprets somatic markers as valencing systems whose activation is required to produce suitable motivation. However, recent research shows that the mechanisms which make objects salient and motivate behavior are independent neurally and cognitively from those which determine valence. The mesolimbic dopamine system plays a central role in salience/motivation by predicting reward (rather than valence), while valencing appears to be realized by a number of other systems, including the opioid and benzodiazepine systems (Berridge & Robinson, 2003; Berridge, 2007, Robinson & Berridge, 2003). Second, the somatic marker hypothesis underspecifies the nature of the explicit representations involved in decision-making. These representations can either be declarative, as when one performs cost benefit analysis by manipulating probabilities, or episodic, as when one uses past experiences to imagine future ones. According to Gerrans then, the planning and decision making deficits of VMPFC patients result not so much from their inability to associate semantic markers to their explicit declarative representations than from their inability to perform mental time travel, that is imagining themselves living out future scenarios and thus activating the motivationally relevant contingencies salient in these imagined experiences.

If this conception of the link between mental time travel and motivation is on the right track, mental time travel could also help explain one unique characteristic of human planning. According to the Bischof-Köhler hypothesis (Bishof-Köhler, 1985, Suddendorf & Busby, 2005), non human animals cannot anticipate future needs or drive states. Humans, in contrast, can plan for the future not just on the basis on their current motivational states but also on the basis of what they anticipate their future motivational states to be. The ability to project oneself forward in time and imagine future scenarios may be an important key to motivation regulation.

How-decisions

The construction of plans for future actions depends in part on semantic memory since it is crucial to their success that the plans we come up with be consistent with our general knowledge about the world. Yet, filling in the details of a plan may depend on our ability to imagine future episodes, since they provide the particularities that will help us fine-tune the plan to the particular occasion. However, trade-offs need to be considered since mental time travel is effortful and cognitively costly. When I form the prospective intention to go to my office tomorrow rather than to work from home, there is no need for me to mentally rehearse the route to my office. The route is familiar enough that I can trust myself to do the right thing when the time comes. Suppose, however, that I have an appointment tomorrow in some other part of the city I am less familiar with. In that case, it may be worthwhile rehearsing possible ways of getting there and using memories of past episodes to decide between options. For instance, I may remember that changing lines at this station takes forever and involves walking along endless, badly lit, corridors or I may remember getting stuck in heavy traffic on a given bus line. Or imagine again, I am about to visit Beijing for the first time and have no clue what the public transportation is like there. In such a case it may be a waste of time and energy imagining potential future scenarios for how to get around in Beijing. The scenarios I come up with may be far off the mark and completely useless in the end; better just way and see.

More generally, whether we make how-decisions early or late and the extent to which we use mental time travel to make those decisions depends on a number of factors, among them: how predictable we think the future situation is; how knowledgeable we are; whether our knowledge is mostly declarative or based on prior personal experience; how motivated we are (as rehearsing a future scenario may help reinforce motivation); how novel or difficult the prospective action is; how neurotic our personality is. In addition, there appear to be important individual differences in the ability to project oneself into possible future events. A recent study (D'Argembeau & Van der Linden, 2005) provides evidence that the individual differences in dimensions known to affect memory for past events similarly influence the experience of projecting oneself into the future. People less adept at recalling in vivid detail past episodes of their life, are also less able to simulate specific future events. Note that these results also provide support for the view that view that mental time travel into the past and mental time travel into the future rely on similar mechanisms.

When-decisions

A prospective intention is an intention to perform an action at some future time. But if the intention is to eventuate into action, it is important that the time of action be specified. An initial when-decision can take at least two forms. The time of action can be specified in explicit temporal fashion, say as "next Tuesday" or "on the 1st of November" or it can be specified in relation to some specific future event, say "when I next meet Charles" or "as soon as the bell rings". Work in the field of prospective memory sheds light on interesting differences between the time-based and the event-based strategies.

Prospective memory is a field of cognitive psychology dealing with remembering to perform an action in the future (e.g., I must remember to stop to buy fruit on my way home from work). The starting point for prospective memory is clearly an intention to perform an action at a future time. Most experimental studies deal with event-based prospective memory, in which a specific event that will occur in the future is used as a cue for an action. Translating a long-range intention into action then becomes a matter of identifying that the cue has occurred, and retrieving the appropriate action in response to it. Several studies of 'implementation intentions' in Applied Psychology (Gollwitzer, 1999), suggest this strategy is effective: intended actions such as taking medication are more likely to occur if people link their implementation to a specific external event. According to Gollwitzer (1999), what explains the efficacy of implementation intentions is the fact that their formation triggers two sets of processes. First, when an implementation intention is formed, mental representations of the relevant situational cues become highly activated, leading to heightened accessibility and thus a better detection, of these cues when they are encountered (Aarts, Dijksterhuis, & Midden, 1999; Gollwitzer, 1999; Webb and Sheeran, 2007). Second, implementation intention formation not only enhances the accessibility of the specified situational cue, but also forges an association between that cue and a response that is instrumental for obtaining one's goal, thus making action initiation more immediate and efficient.

Such 'implementation intentions' may take advantage of the fact that externally-cued intentions are normally more strongly held, in the sense of being harder to overturn, than internally-generated intentions (Fleming et al., in press).

Prospective memory can also be time-based, rather than event-based. In time-based prospective memory, an intended action is performed at a designated future time, without any particular cue event occurring at that time. Thus, time-based prospective memory seems to be purely endogenous, while event-based prospective memory effectively reduces endogenous actions to cue-triggered reactions. The distinction between the two forms is supported by the dissociation between different rostral prefrontal activations in time-based and event-based prospective memory tasks (Okuda et al., 2007).

Recent studies of time-based prospective memory suggest an interesting role for unconsciously-initiated processes, similar to Libet's action initiation, in linking long-range intentions to eventual action. Kvavilashvili and Fisher (2007) asked participants to call an experimenter at a self-chosen time one week after an initial briefing session. In the intervening week, they noted the circumstances in which they remembered this intention, using a diary. Although the authors refer to these memory events as 'rehearsals' they were primarily automatic and uncued events, in which the intention to make the phone call simply 'popped into' the participant's mind, without obvious cue or antecedent. The frequency of these recall events increased dramatically in the day before the phone call was due, but this increase was less dramatic in those participants who in fact failed to return the phone call on time.

3. Linking prospective intentions to immediate intentions

Actions are not always the product of prospective intentions, they may often simply be the outcome of immediate intentions, formed on the spot so to speak. But let us focus on cases where actions are preceded and brought about by prospective intentions. What is the additional contribution, if any, of immediate intentions to such actions?

Recall that in the previous section we characterized the content of immediate intentions as involving episodic representations. Forming an immediate intention involves fitting one's endogenous goal to the current situation, using contextual information to generate a representation of a specific episode of acting. When one has a prospective intention to perform an action, how much work there is left for an immediate intention to do at the moment of action itself will depend on how episodic the content of the prospective intention already is. This will in turn depend on the extent to which the agent made use of mental time travel in forming and shaping his prospective intentions. For example, a person forming a prospective intention may become fully involved in mental time travel and may simulate the full details of how and when the action will occur. Conversely, one can have a genuine prospective intention while knowingly leaving it for later to decide on the means. At one extreme of a continuum is the "neurotic planner", at the other end is the "optimistic improviser".

The neurotic planner makes extensive use of mental time travel, imaginatively combining and recombining elements from prior stored episodes to generate early on precise scenarios concerning the action to be performed and the situation in which it is to be performed. His strategy is to generate as much episodic information as he can as early as he can. When mental time travel serves well, this front-loading strategy leaves little left for immediate intentions to do.

Using Gollwitzer's terminology, we can say that neurotic planners tend to make early detailed how- and when-decisions, thus forming implementation intentions. A key feature of this strategy of early planning is that it allows for later automatization. As Gollwitzer points out, implementation intentions automatize action initiation: "The goal-directed behavior specified in an implementation intention is triggered without conscious intent once the critical situational context is encountered" (Gollwitzer, 1999: 498). Thus, the use of external cues to trigger action seems partly to shift the action from an endogenous or voluntary one to a stimulus-driven or reactive one.

In contrast, the optimistic improviser generates little episodic information early on. She makes a what-decision, possibly a time-based when-decision, but keeps her options open as to how and in what specific situation the action is to be performed. She is committed to generating relevant episodic information in real-time, at the moment of the action itself. The prospective intentions of agents following this strategy contain as yet too little episodic information to yield action. To fill this informational gap between her prospective intention and action initiation, the agent will have to form an immediate intention specifying the missing information. This means that the agent must retain some endogenous control over action initiation and cannot delegate it to automatic responses to environmental triggers.

Episodic information must be generated in order to produce a specific action. It can be generated either early (neurotic planner) or later (optimistic improviser). These are in some sense alternative reciprocal responses to the common challenge of deciding exactly what one will do. Despite the personality-based labels we used, early vs. late planning isn't just a matter of temperament. Each strategy may be better suited to some situations than to others. Early planning has its dangers. If the agent's anticipations were not correct, the external cues on which action initiation depends may fail to materialize. Or, worse perhaps, the cues may be present and automatically trigger the action when other unanticipated and unattended aspects of the situation make it unadvisable to pursue as planned. The late planner may be more flexible, but she risks unpreparedness when the time of acting comes. Having left it to the last moment to deliberate about means, when she finally does so she also risks re-opening the Pandora's box of deliberation about ends. What-decisions and how-decisions aren't strictly compartimentalized. The costs and efforts involved in deliberating about how to *A* under time pressure, may lead one to reconsider whether to *A* in the first place, when giving up *A*-ing may well tempt us as the less costly option.

Often, and perhaps most of the time, our planning strategies will be mixed strategies, taking into account various factors beyond mere temperament; among them, the expected predictability of relevant future situations, one's store of relevant semantic and episodic information, one's degree of motivation, the degree of novelty or difficulty of the planned action, and how strong one thinks time constraints will be at the time of acting. The generation of episodic information about future actions will thus be distributed over time in various ways according to our assessment of these factors. One example of these differing distributions comes from the contrast between an event-based and time-based prospective memory. In event-based prospective memory, specific details of the action episode are already present in the prospective intention itself. In contrast, time-based prospective memory lacks any concrete details about the specific context in which the action will occur. Most people can and do use both forms of planning. This flexibility in the temporal distribution of episodic information is a fundamental dimension of the psychology of intention. The skilled planner is the one who knows how best to take advantage of this flexibility.

4. Conclusion

The concept of intention can do useful work in psychological theory. We have made a distinction between prospective and immediate intentions. Many authors have insisted on a qualitative difference between these two regarding the type of content, with prospective intentions generally being more abstract than immediate intentions (e.g., Searle, 1983; Pacherie, 2008). However, we suggest that the main basis of this distinction is temporal: prospective intentions necessarily occur before immediate intention and before action itself, and often long before them. In contrast, immediate intentions occur in the specific context of the action itself. Yet both types of intention share a common purpose, namely that of generating the specific information required to transform an abstract representation of a goal-state into a concrete episode of instrumental action directed towards that goal. To this extent, the content of a prospective and of an immediate intention can actually be quite similar. The main distinction between prospective and immediate intentions becomes one of when, i.e., how early on, the episodic details of an action are planned.

In our view, the conscious experience associated with intentional action comes from this process of fleshing out intentions with episodic details. In the field of episodic memory, representations of episodes are thought to include an autonoetic type of consciousness (Tulving, 1983). We suggest that intentional actions reach conscious awareness at the point where they become specific action episodes. However, the time when this occurs can vary. We have argued that episodic detail can be generated either as part of advance planning, in the form of prospective intentions, or as part of an immediate intention in real time. In the former case, one might have a conscious mental image of what one will do, but the doing itself may be automatised and only marginally conscious. In the latter case, one may have a specific conscious experience linked to the initiation of action, along the lines studied by Libet.

References

Aarts, H., Dijksterhuis, A., & Midden, C. (1999). To plan or not to plan? Goal achievement or interrupting the performance of mundane behaviours. *European Journal of Social Psychology*, 29, 971–979.

Bancaud J, Talairach J, Geier S, Bonis A, Trottier S, & Manrique M. (1976). Behavioral manifestations induced by electric stimulation of the anterior cingulate gyrus in man. *Revue Neurologique (Paris), 132*, 705-724.

Bechara, A., Damasio, H., Damasio, A. and Lee, G.P. (1999). "Differential Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Human Decision-Making." *The Journal of Neuroscience*, 19: 5473-5481.

Berridge, K. C. (2007). The Debate over Dopamine's Role in Reward: the Case for Incentive Salience. *Psychopharmacology*, 191, pp. 391–431.

Berridge, K. C. and T. E. Robinson (2003). "Parsing reward." *Trends in Neurosciences*, 26, (9): 507-513.

Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Cambridge University Press.

Chassagnon S, Minotti L, Kremer S, Hoffmann D & Kahane P. (2008). Somatosensory, motor, and reaching/grasping responses to direct electrical stimulation of the human cingulate motor areas. *Journal of Neurosurgery, 109*, 593-604.

Damasio, A.R., Tranel, D., Damasio, H., (1991). Somatic markers and the guidance of behaviour: theory and preliminary testing. In: Levin, H.S., Eisenberg, H.M., Benton, A.L. (Eds.), *Frontal Lobe Function and Dysfunction*. Oxford University Press, New York, pp. 217–229.

Dudai, Y. & Carruthers, M. (2005) The Janus face of Mnemosyne. Memory: Some systems in the brain may be better equipped to handle the future than the past. *Nature*, 434:567.

Fleming SM, Mars RJ, Gladwin TE & Haggard P (in press). When the brain changes its mind: flexibility of action selection in instructed and free choices. *Cerebral Cortex*, in press.

Fried, I., Katz, A. McCarthy, G., Sass, K.J., Williamson, P., Spencer, S.S., & Spencer, D.D. (1991). Functional organisation of human supplementary motor cortex studies by electrical stimulation. *Journal of Neuroscience*, 11, 3656 – 3666.

Gerrans, P. (2007). Mental time travel, somatic markers and "myopia for the future". *Synthese*, 159, 3: 459-474.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503.

Haggard P & Eimer M. (1999) On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research, 126*, 128-133.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Science*, 9(6), 290-295.

Haggard, P., & Cole, J. (2007). Intention, attention and the temporal experience of action. *Consciousness and Cognition*, *16*(2), 211-220.

Haggard, P., (2008). Human volition: towards a neuroscience of will. *Nature Reviews: Neuroscience, 9, 934-946.*

Hassabis, D. K. D., Vann, S.D. and Maguire, Eleanor A. (2007).Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Acadeedmy of Sciences*, 104, 1726-1731.

Klein, S. B. (2002). Memory and temporal experience: the effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition*, 20, 353-379.

Kremer S, Chassagnon S, Hoffmann D, Benabid AL & Kahane P. (2001). The cingulate hidden hand. *Journal of Neurology, Neurosurgery and Psychiatry, 70*, 264-265.

Kvavilashvili L & Fisher L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Role of incidental cues, ongoing activity, age, and motivation. *Journal of Experimental Psychology: General, 136*, 112-132.

Lhermitte F. (1983). 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain, 106*, 237-255.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential).

Mima T, Ikeda A, Yazawa S, Kunieda T, Nagamine T, Taki W & Shibasaki H. (1999). Somesthetic function of supplementary motor area during voluntary movements, *Neuroreport, 10*, 1859-1862.

Okuda J, Fujii T, Ohtake H, Tsukiura T, Yamadori A, Frith CD & Burgess PW. (2007). Differential involvement of regions of rostral prefrontal cortex (Brodmann area 10) in time- and event-based prospective memory. *International Journal of Psychophysiology, 64*, 233-246.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107, 1:179-217.

Robinson, T. & Berridge, K. C. (2003). Addiction. *Annual Review of Psychology*, 54: 25–53.

Rowe JB, Toni I, Josephs O, Frackowiak RS & Passingham RE. (2000). The prefrontal cortex: response selection or maintenance within working memory? *Science, 288*, 1656-1660.

Searle, J. (1983). I*ntentionality*. Cambridge: Cambridge University Press.

Soon, C., Brass M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience, 11*, 543-545.

Suddendorf, T., & Busby, J. (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, 7, 391–396.

Suddendorf, T., & Busby, J. (2005). Making decisions with the future in mind: Developmental and comparative identification of mental time travel. *Learning and Motivation*, 36, 110–125.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, 123, 133–167.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to human? *Behavioral and Brain Sciences*, 30: 299-351.

Tulving, , E. (1983) *Elements of episodic memory*. Clarendon Press.

Tulving, , E. (2005) Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds),: *The missing link in cognition: Origins of self-reflective consciousness*. Oxford:, Oxford University Press, ed. H. S. Terrace & J. Metcalfe, pp. 3–56. Oxford University Press.

Webb, T. L. Sheeran, P. (2007). How do implementation intentions promote goal attainment? A test of component processes. *Journal of Experimental Social Psychology*, 43: 295-302.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, 121:, 331–354.

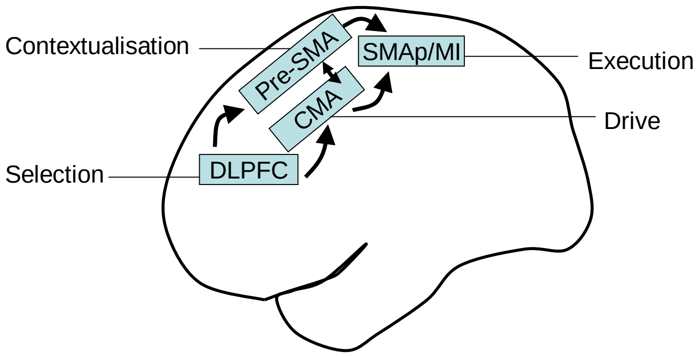
Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

Figure 1

Figure Caption

Figure 1. A simple model of the division of labour between frontal cortical areas in the initiation of intentional action. Selection between alternative action plans occurs in the dorsolateral prefrontal cortex (DLPFC). The signal corresponding to the selected action is forwarded along two major neural pathways: to the Cingulate Motor Area (CMA) to provide a motivated drive to perform the action, and to the Pre-Supplementary Motor Area (Pre-SMA) to modulate the action according to current context, competing action representations etc. Hypothesised interactions provide an arbitration between the push from drive and the constraints provided by context. Both areas have access to the main motor execution pathway through their connections to the Supplementary Motor Area proper (SMAp) and the primary motor Cortex (MI).

Figure 1


Contextualisation
Pre-SMA
SMAp/MI
Execution
CMA
Drive
Selection
DLPFC