Mental Time Travel (MTT) is, roughly, an individual’s capacity to project herself into the past or future by remembering or imagining first-personal experiences respectively. MTT is further presumed to have a distinct, concrete though dispersed neural correlate, and hence describes a neuro-cognitive phenomenon.

Opening with a brief sketch of the development and current state of the art, the essay pursues three central aims: Firstly, it constitutes a plea for more conceptual rigour on the cognitive side of the fence, so as to ensure that meaningful lessons can be drawn from neurological enquiry about it. Secondly, a partial conceptual qualification of the necessary requirements of MTT as traditionally conceived is proposed, as they seem vague, uninformative and arbitrary. Finally, a revision of MTT is attempted, which aspires to include a variety of mental states so far not associated with MTT. MTT, as it is currently defined and investigated, I will argue, stands too heavily in the genealogical debt of research into episodic memory, and suffers from an astonishing neglect of considerations pertaining to imagination.

1. What is Mental Time Travel?

The fact that certain types of imagination activate the same brain zones as episodic memory provoked the hypothesis that there is a single neuro-cognitive system which enables human beings to engage in mental time travel (MTT). Mental time travel is, roughly, an individual’s capacity to project herself into the past or future by remembering or imagining first-personal experiences respectively. As a neuro-cognitive phenomenon, MTT is presumed to have a distinct, though dispersed, neural correlate. Evidence from various disciplines is consistent with the MTT hypothesis. Studies in ontogenetics have confirmed that episodic memory and prospection (mental time travel into the future) emerge in parallel in children aged around three to four. Furthermore, episode specific details decrease with age both for generated past and future events.1

Lesion studies show that ventromedial frontal damage leads to loss of episodic memory and prospection, while leaving large parts of the cognitive apparatus in tact.2 Moreover, patients with hippocampal amnesia are unable to generate everyday imaginary experiences.3

Neuroimaging draws a similar picture. According to a variety of studies, episodic states about future and past have a common underlying cerebral base: When talking freely about past or future events, PET and fMRI scans revealed shared activity in regions including the ventromedial prefrontal cortex, and parts of the medial temporal lobe.4

I’ll open with a brief sketch of the leading account of mental time travel.5 I will then proceed to argue that this account is misconceived for two fundamental reasons: (1) Its genealogical debt to episodic memory and autonoetic consciousness as well as the shallow conception of imagination in play give rise to an ad hoc and unnecessarily constrained account of episodic states. (2) The necessary capacities for MTT as traditionally conceived are unfounded, their formulation is conceptually vague and uninformative. This will severely obstruct empirical research into the ontogenetic and neurological foundations of the phenomenon.

2. Foundations of the Traditional Account

MTT has developed out of the psychological study of memory, and in particular Tulving’s (1972) landmark distinction between semantic and episodic memory. The former takes propositional form since it is factual. I

---

3 Hassabis et al. (2007).

---

5 The core references are Wheeler, Stuss & Tulving (1997) and Tulving (2002). An alternative, but similar account is proposed by Suddendorf & Corballis (1997; 2007).
can e.g. recall that Paris is the capital of France. Episodic memory, by contrast, refers to an individual's engaging in an episode of past personal experience, e.g. when I remember what my first arrival in Paris was like. It is characteristically accompanied by a particular feeling of “warmth and intimacy” (W. James, 1890), in other words it is phenomenologically rich. Later on, autonoetic – “self-knowing” – consciousness became the distinguishing mark of episodic memory. Tulving, who coined the term “autonoetic consciousness”, defines it as “the kind of consciousness that mediates an individual’s awareness of his or her existence and identity in subjective time extending from the personal past through the present to the personal future” (1985: 1). In response to neuropsychological findings, the restriction of episodic states soon dropped away; episodic states were henceforth considered to encompass mental time travel both into the past and future.

My first criticism concerns the scope of the leading account: MTT is restricted to episodic states exclusively concerned with an individual’s personal past, present and future. Employing a detailed typology of imaginative states I will draw up a rival account which construes the phenomenon in question more broadly. I will argue from the following two hypotheses:

**Common Kind Hypothesis:**
There exists a basic mental state, called “episodic state”, in which we undergo phenomenologically rich experiences from a first person perspective. Such states are “quasi-perceptual” in so far as they resemble perceptions, and draw heavily on past perceptual and proprioceptive intake, but are not direct representations of reality. Episodic states comprise of episodic memory – quasi-perceptions of the past, and participatory imagination – quasi-perceptions of hypothetical and potentially future episodes.

**Common Capacity Hypothesis:**
Episodic states supervene on a single type of brain state; they can be characterized by a differentiated neuronal correlate and are the product of a particular neurosystem of the brain.

### 3. Problems of Scope and a Rival Account

MTT, as it is traditionally conceived, is construed unnecessarily narrow, since it insists on (i) a clearly defined temporal component, which (ii) involves an explicit awareness of a narrative self and concerns (iii) episodes which are explicitly personal/autobiographical – in the sense that the subject involved must be the thinker’s empirical self, and the scenarios must be true past or probable future experiences. Both constraints can be directly derived from Tulving’s characterization of MTT taking place in “subjective time extending from the personal past through the present to the personal future” and countless other passages. They are equally present in Suddendorf & Corballis (1997, 2007).

The main reason for the narrow construction of MTT seems to lie in the genealogy of the term, coming from episodic memory and hence focusing heavily both on a temporal component and some sort of autobiographical element. Autonoetic awareness as the central property of episodic states has further helped to foster such a questionable conception. The problem extends into the experimental paradigms: Not only is rather few research done on future MTT (i.e. the imagination component as narrowly conceived), but furthermore hardly any experiments included imagination not structured autobiographically or temporally in the relevant way.

Mental voyage, as I propose it, explicitly includes three types of episodic imagination which are not encompassed in MTT: (i) Episodes about what would have happened to myself if I had acted differently in the past. (ii) Imaginations involving my empirical self in scenarios which do not have a temporal specification whatsoever. (iii) Imaginations not involving my empirical self, but another self, or a general self – i.e. episodes concerning what it is like for Jack to win an Oscar, or for someone to win the lottery respectively.6

---

6 The proposal to define MTT wider than is commonly done is not revolutionary. A similar point of view is shared by D’Argembeau & Van der Linden (2005), Hassabis (2007) and Buckner and Caroll (2007).
4. Central Capacities for MTT

Methodological Considerations

Mental time travel is generally conceived as a neuro-cognitive or brain/mind system\(^7\), that is, as a mental capacity which has a real, singular, though dispersed neural correlate. Crucially, the description of its functions and properties take place in two conceptual spaces: an abstract vocabulary pertaining to mind as elaborated in psychology and philosophy and the vocabulary referring to concrete phenomena of the brain as employed by neuroscience. The mental vocabulary is dominant due to our still very limited understanding of the brain. Hypotheses concerning the functioning of the brain as well as experiential paradigms are largely formulated in mental terms, which gives rise to a variety of complications: (i) The conceptions of the mind and its capacities are manifold, so a choice has to be taken which (ii) is likely to leave its mark on the formulation of the respective hypothesis, experiments, and hence the empirical “findings”. (iii) A neuro-cognitive hybrid vocabulary facilitates conceptual confusion if (for instance) a brain phenomenon is “associated” with a mental phenomenon which gives rise to different interpretations in distinct conceptual frameworks of the mind.

Given the complications arising from two distinct types of interacting vocabularies, empirical underdetermination, and the sensitivity of the subtraction method in hemodynamic techniques (PET and fMRI), two things should be clear: Adequate enquiry into such neuro-cognitive phenomena can only succeed if its constituents are (i) conceived of in minimal, rather than complex units and (ii) defined as rigorously as possible on the cognitive side so as not to obstruct and confuse enquiry on the neurological side. The literature on MTT does not adhere to these criteria. The central capacities of MTT are both theoretically ad hoc and so ill defined that while scholars take themselves to be in conceptual agreement they frequently operate with – and do empirical research based on – radically different concepts. We will now turn both to the misspecification of the central capacities of MTT and the latent conceptual anarchy.

Autonoetic Consciousness

The central criterion of episodic memory and states of mental time travel more generally, is autonoetic awareness (Wheeler, Stuss & Tulving, 1997; Tulving 2002; Suddendorf & Corballis 1997, 2007). Though a relatively recent, and entirely technical concept, astonishingly there are four conceptions of autonoetic consciousness which stand in rivalry.

Autonoetic consciousness understood as the distinguishing mark of mental states which have a particular “feel” to them is roughly equivalent with Block’s (1995) “phenomenal consciousness”. Autonoetic awareness in this sense is usually cashed out by reference to the feeling of “warmth and intimacy” (W. James, 1890), or the “subjectivity” of such states in comparison to, for example, “objective” semantic memory. A second account puts the stress on (narrative) self-awareness, reasonably enough, since autonoesis has frequently been characterized as “self-knowing” consciousness (Tulving, 1985), and it is due to autonoesis that an individual is supposedly able to project his self into the past and future. A third account focuses on the kind of (phenomenal) feature which allows us to distinguish, for instance, an episodic memory from an imagination or a daydream. Episodic states seem to come pack and parcel with a certain phenomenological feature pertaining to time, which allows us not to confound past, present and future episodic states; in the case of memory we witness, for instance, a “feeling of pastness”. Finally, there are various hybrid accounts which include some or all of the mentioned features. Unsurprisingly, with autonoesis - “the hallmark of episodic memory” (Tulving) – being such a promiscuous concept, the episodic/semantic memory distinction is also drawn in all sorts of ways.

The three mentioned aspects, phenomenality, narrative self-awareness, and subjective temporal indexing are neither inconsistent nor necessarily unrelated, but nonetheless distinct. Problematically, however, different authors seem to work with different

definitions. There is, furthermore, no apparent reason for lumping them together - especially inexplicitly, and in varying constellations. In fact, it seems favourable to keep them separate so as to curb the spreading conceptual confusion, and in particular so as to control – as far as possible - separately for each in experiments.

**Consciousness**

How best to reorganize what is left of autonoetic consciousness? Block (1995) distinguishes four concepts of consciousness: **Self-consciousness** constitutes the possession and competent mastery of the concept of the self; **monitoring consciousness** is the metacognitive process of one’s realizing to be in a certain state (e.g. to know that one believes X). **Access-consciousness** is the property of a representation which “is broadcast for free use in reasoning and for direct “rational” control of action (including reporting)”. Finally, and for our purposes probably most importantly, there is **phenomenal consciousness**, which is notoriously hard to define. Under **phenomenal consciousness** we understand the experiential properties of a conscious state, or what it “is like” to be in that state. Block highlights that phenomenal and access consciousness are conceptually distinct, though admits that they might contingently always appear in parallel in human subjects.

My proposal is to abolish the concept of **autonoetic awareness** since it is vague and unnecessarily lumps together all sorts of phenomena which are best kept separate. Phenomenal consciousness (or phenomenality) is most salient in perceptual experiences, however it also characterizes episodes of remembering and sensory imagination, though the phenomenal properties are not as pronounced. It is probably phenomenality, intimately related to the first-person perspective, which determines the distinctions between episodic and semantic memory and between sensory and propositional imagination in the first place.

**Subjective Time**

As mentioned before, **time** (like colour or smell, for instance) also seems phenomenologically salient, in particular, **duration** (“an instant seeming an eternity”) and the **subjective temporal location** (“it being early”; “something being a long time ago”). It is a feature of episodic states that they frequently present themselves as a particular episodic state (e.g. memories through an attached feeling of “pastness”). However, as pointed out above, various types of participatory imagination do not have an explicit temporal component, and are hence not located in what Tulving calls “subjective time” (Tulving, 2002, 2005; a feature reproduced in virtually all articles on MTT). Furthermore, though there frequently seems a phenomenally salient indicator present in episodic states, it can easily be wrong (as for example in implanted or false memory, possibly in states of déjà-vu experiences if one wants to count them amongst episodic states).

In short, the temporal component should be abolished, as there is reason to include episodic states without them into the common kind; the fact that temporal features do not form part of memory traces (Friedman, 1993) leaves us with the perfectly viable option of their being “attributed” by an extra capacity, if at all.

**Involvement of the Self**

Self-awareness to a rather high extent is considered a fundamental prerequisite of MTT by all leading accounts. For Tulving...

---

8 The first conception of autonoetic consciousness is prevalent e.g. in D’Argembeau & Van der Linden’s (2004, 2006), Gardiner’s (2001) and some of Tulving’s work. Fink et al (1996), Gerrans and Tulving (2002) seem to opt rather for the second, whereas Wheeler, Stuss and Tulving (1997) seems to work strongly with the third conception of autonoetic consciousness.

9 Cf. also Nagel (1975).

10 This observation seems to be well in line with Suddendorf & Corballis’ account of MTT: Drawing on a survey of episodic memory by Friedman (1993), they highlight that neither chronology nor a “sense of ‘pastness’” are basic properties of memory traces. Nevertheless, and in contrast to our account, Suddendorf & Corballis are convinced that experiences of episodic states always have such phenomenal properties concerning time.

11 Gerrans & Sander (MS) propose a radically different account in which MTT is conceived as a non-conscious, implicit phenomenon not involving explicit (overt) representation – and hence does not require (narrative) self-awareness.
and colleagues, the relevant degree of self-awareness required for MTT corresponds to Stuss & Benson’s (1986) third functional level of the frontal lobes. The latter provides “the ability to introspect on one’s own thoughts and to realize the relation of self to one’s social environment” (1997:334), and is further deemed “intimately related” to autonoetic consciousness.

Mitchell’s (1994) threefold account of the self is employed by Suddendorf & Corballis with questionable success to illuminate the role of the self in MTT. In Mitchell’s framework, the second type of self – a “self as built on kinaesthetic-visual matching” allows us to engage in pretense, planning, imaginative experience and fantasy. Suddendorf & Corballis deem this insufficient for MTT, which apparently calls for a self “built on symbols, language and artefacts” so as to allow the individual to understand social norms, and to dissociate self and other. It rests obscure, however, how pretense, imagination and fantasy (i.e. second-level activities) are possible without the capacity to dissociate from one’s present states (only available on the third level) and hence whether Mitchell’s account is not rather ad hoc.

Ignoring the specificities of the two proposed accounts of the self, let us note that they both share two essential features: Firstly, they require the mastery and competent application of an explicit concept of the self, and secondly, they pertain to what is frequently called a “narrative” account of the self.

The narrative self contrasts with the “minimal” self, i.e. immediate awareness of oneself as the subject of experience. A minimal self is conceived to be little more than “a bare locus of consciousness, void of personality” (G. Strawson, 1999: 493), aspects of continuity over time are neither included in the definition nor deemed necessary. By contrast, an adequate account of the latter – i.e. of how to explain an awareness of oneself from past to present and future, both in terms of one’s individual experience and testimony of others’ – is the centrepiece of “narrative” conceptions of the self (cf. Gallagher, 2000 for a recent review). Both minimal and narrative self demand the competent mastery of a self-concept. However, even simpler conceptions are thinkable, amounting to no more than a point of consciousness to which action and experience is relativized.

### The Self Reconsidered

Though self-awareness in the general sense is undoubtedly an important aspect of MTT, it is difficult to say something insightful about it. Coherent ad hoc stories about the involvement of a narrative self in MTT can be told to abundance within the framework of our mental vocabulary; however, it rests entirely in the dark of what should constitute (and how to test for) the involvement of narrative self-consciousness on the level of the brain.

Rather than working with the proposed, extremely complex notions of self-awareness as necessary capacities for MTT, it seems preferable to focus on minimal, separable ingredients. Episodic thought is not fundamentally characterised by its relation to the subject’s self, but rather by being a particular mode of thought – it is essentially first-personal, but not essentially autobiographical as often presumed. Furthermore, episodic thought is susceptible to a certain measure of control: I can bring out more vividly, or shift attention to, different aspects of a past event; in imagination the freedom of control is even more pronounced.

On top of narrativity, we might be inclined to question mastery of a self-concept as a necessary requirement of MTT. Perry (1998) for instance, contrasts agent-relative knowledge with self-attached knowledge. The former takes place “from the perspective of a particular agent who does not need to have an idea of self, or a notion of himself” (BBB) but who is nonetheless capable of placing

---

12 The first two (and hence more basic) levels deal roughly with the ordering or representation of information and executive functions respectively.

13 Furthermore, Suddendorf & Corballis themselves seem to be skeptical about the requirement of linguistic abilities for MTT, so it is surprising that they opt for Mitchell’s third level, rather than the second.

14 Dennett (1991), one of the most prominent proponents of the narrative self, is pessimistic about the inquiry into a self which is more than just an abstract, theoretical postulate in our models of the mind.
himself within his environment. One way to make this process explicit are utterance involving demonstratives, e.g. when he thinks/says “There is an apple”. In self-attached knowledge, by contrast, the agent disposes of a self-notion and expresses his knowledge by means of the (pure) indexical “I”, e.g. “I see an apple”. Both involve the first-person perspective, essential for episodic states, but only the latter type of subject can make this explicit to herself.

Interestingly, the first kind of knowledge, and the very basic type of self involved would satisfy a variety of episodic states: If I engage in a phenomenally rich first-person episode with the content “There was an apple”, or follow the instruction “Imagine an apple”, I do not necessarily need the first-person indexical; an appropriately imagined content of the kind “There is a juicy red apple” fulfills the demand no less than “I see a juicy red apple in front of the eye of my mind”. However, whenever the person who remembers or imagines is to take himself as the object of his episodic state, he or she is required to have an explicit notion of himself, otherwise he couldn’t attribute any experiences or properties to himself. What this confirms, once again, is that the capacity to engage in episodic states is — at least in principle — independent of the capacity to engage in autobiographical episodic states (though due to the contingent set-up of the brain this might not actually be so).

A variety of insights follow from this: Firstly, we have another argument why autonoesis — i.e. “self-knowing consciousness” — might not be a fundamental requirement of MTT in so far as there exists a kind of basic episodic state which does not depend on the mental time traveller having even a very basic notion of himself. Secondly, and relatedly, the first-person perspective might be entirely severed form any type of complex narrative self in so far as there is a variety of possibilities to “fill” it: Apart from one’s own self, it could be another’s or a general self. Thirdly, it might be hypothesized that MTT is intimately related to the mastery of demonstratives (and as concerns autobiographical episodic states, also indexicals) and temporal concepts, or their respective non-conceptual counterparts. So an important open question is to what extend the non-conceptual and pre-linguistic resources (e.g. of small children) suffice for MTT. 

In short, there is no need to presume MTT in the need of an awareness of a complex narrative self, whose involvement will most likely prove impossible (or very difficult) to test for. Furthermore, it is not clear why complex introspective abilities, metarepresentation and awareness of one’s social environment etc. should be necessary for basic episodic thought. As has been argued, the only true requirements for basic MTT are the involvement of the first-person perspective and a given measure of control of the episodic content. Non-autobiographical episodic states (There was an elephant or This will be wet and cold etc.) might not even demand an explicit notion of the self. Finally, at least in so far as mental voyage is concerned, it will be extremely misleading to stress the requirement of (narrative) self-awareness, both in so far as certain types of participatory imagination are neither relativized to “subjective time” (i.e. are not locatable within a subject’s personal narrative) or to the subject’s own self at all. They might simply be about what it would be like to win the Nobel Prize or how it would feel to have one’s neighbour’s problems. Rather than operating with various ad-hoc accounts of the narrative self, it seems advisable to take the first-personal mode and control (or sense of agency) as the basic requirements for MTT, and to devise abilities allowing us to judge to what extent an explicit notion of the self, and conceptual abilities are required as well.

Bibliography


Macphail (2000) contents that MTT is dependent on language; Suddendorf & Corballis (1997) disagree. Clayton et al. 2000 hold that it is impossible to investigate MTT in non-human animals, since language is the only agreed-upon medium to report it. Suddendorf & Busby (2003, 2005) challenge this opinion; they hold that MTT is also expressed in action through „pantomimining past events, practice rehearsal of anticipated events and in acts that only make sense in the light of anticipated future events” (2005: 114).


Gerrans, P. S., D. (MS). "Implicit Mental Time Travel."


