

listened to the trauma script. Note that not only did the neutral identity states claim they did not recognise the trauma story as relevant to themselves, but also their brain response backed that up; there was an amnesic barrier between the alters.

The possible problem for the M&C position, then, is that an episodic memory, fully self-referential and auto-noetic when accessed by one alter, behaves totally differently when accessed by another alter. In some cases, the material is treated as knowledge, and in other cases it is treated as though it does not exist. This seems to create problems if the differences between episodic memory, event memory, and beliefs are simply there in the representations, as M&C seem to claim. The alternative way of thinking about this seeming paradox has two components. First, the emphasis would be on meta-cognitive processes: what other people refer to as executive processes. This is distinct from the metarepresentational format referred to by M&C, which would be seen as the product of current processing. The second component is the retrieval process itself. If one thinks in terms of context-sensitive memory (e.g., Godden & Baddeley 1975), it is natural to treat some notion of *self* as a part of the addressable section of a memory representation. This would have a normal use of distinguishing between representations of events that were first- or secondhand. Its interpretation by the executive (meta-cognitive) processes would give rise to auto-noesis.

Within such a framework, the treatment of the DID case is straightforward as outlined above. Executive processes use the *self* marker specific to the alter that is currently active either to restrict whether material can be made conscious or, in other cases, whether the memory representations can be accessed at all. In this way a particular representation can either have an auto-noetic character or not. The apparent paradox is solved here by the use of processing. It is not immediately apparent how a representational view such as that put forward by M&C would deal with it. I leave it in their hands.

## Epistemic authority, episodic memory, and the sense of self

doi:10.1017/S0140525X17001443, e24

Jennifer Nagel

Department of Philosophy, University of Toronto, Toronto, ON, M5R 2M8, Canada.

[jennifer.nagel@utoronto.ca](mailto:jennifer.nagel@utoronto.ca)

[http://individual.utoronto.ca/jnagel/Home\\_Page.html](http://individual.utoronto.ca/jnagel/Home_Page.html)

**Abstract:** The distinctive feature of episodic memory is auto-noesis, the feeling that one's awareness of particular past events is grounded in firsthand experience. Auto-noesis guides us in sharing our experiences of past events, not by telling us when our credibility is at stake, but by telling us what others will find informative; it also supports the sense of an enduring self.

Human knowledge of past particular events is enabled not only by episodic memory, but also by semantic memory and by inference: for example, in detective work from currently perceivable evidence. With various ways of knowing past events, it is an interesting question why we have auto-noesis, the self-reflexive consciousness that is applied to the past in episodic memory. Mahr & Csibra (M&C) propose that its function is communicative: We need to keep track of whether judgments about past events originate in firsthand experience in order to share these judgments most effectively with others. M&C may be right to connect consciousness and communication, but some details of their argument are debatable.

According to M&C, auto-noesis is “the capacity that enables us to distinguish between cases in which we can assert epistemic authority for our own testimony and cases in which we cannot” (sect. 3.1.1, para. 4), where asserting epistemic authority is a

matter of staking one's credibility on the truth of what one is asserting. However, one's credibility is at stake in any assertion, whether it is grounded in present or past firsthand experience, inference, or testimony. M&C are right that one may choose to assert that another person said something, endorsing only the fact that this person made a statement (“Jane said that Bill was fired”), but one may also assert the bare proposition learned through testimony (“Bill was fired”), taking a source's word for it and risking one's own reputation in spreading the gossip. One's reputation for reliability is at stake even if one explicitly marks the assertion with an evidential like “apparently” to indicate a testimonial basis, or an epistemic modal (“Bill must have been fired”) to indicate inference; indeed, claims made with such markers are not inherently weaker than their directly grounded counterparts (Von Fintel & Gillies 2010).

Rather than seeing epistemic authority as something binary, either staked or not, one might see it as a matter of degree, with the greatest confidence vested in what we have seen for ourselves. However, this would leave it unclear just why auto-noesis was needed alongside the simpler quantitative sensibilities that already regulate the reliability of our reports from semantic memory (Goldsmith et al. 2002). More qualitative ways of understanding epistemic authority are perhaps more promising. Perhaps auto-noesis tells us which facts we know will be most informative for others to hear. Semantic memory encodes similar facts and regularities across a community: Indeed, the background beliefs in which one has greatest individual confidence are most likely to be shared already by others (Koriat 2008). By contrast, episodic memory provides a unique record of particular events not widely known to others, together with information on who else witnessed these events. Knowing whether others know about a past event may indeed be useful in arguments, but it could also be valuable to the species for more cooperative purposes in communication, just as the conscious availability of subjective confidence enables pooling of current perceptual judgments to increase accuracy in joint decision making (Bahrami et al. 2010).

Auto-noesis can fail to accompany the capacity to retrieve events, although this is rare: Klein (2015a) reports finding only three such cases in the literature. One describes a recently amnesiac patient who was able, when questioned, to narrate many details of his older brother's death in an automobile accident roughly a year before. When complimented on his recall, the patient looked puzzled and insisted that it was the interviewer who had just told him about what had happened to his brother (Talland 1964). Despite his confusion between another's testimony and his own recall, the patient was not reported as lacking confidence in the facts surrounding the automobile accident: He expressed no uncertainty as to whether these were real or merely imagined events. If this patient had detailed knowledge of events in his personal past and the capacity to narrate these events to others, what was he missing? Talland notes that recall in this case was secured only by a series of prodding questions, rather than flowing spontaneously. Auto-noesis enables recognition of what one's audience will not already know: The intuitive sense of private access to one's personal past enables recognition that description of these events will be informative to others; lacking that sense leaves one unmotivated to share one's knowledge.

M&C's qualitative characterization of the epistemic role of episodic memory focuses on argument: They define episodic memory as “an epistemic attitude taken toward the simulation of a specific past event, which serves to justify a belief about the occurrence of this event” (sect. 1.3, para. 1). However, one may take an epistemic attitude toward a simulation of a past event without remembering the event in question; for example, when visualizing how an animal might have freed itself from the broken trap one is looking at, or when reconstructing what one must have done in the course of last night's drunken stupor. Beliefs about past events may be well justified by imaginative simulation informed by current perceptual input, and these simulations can make us aware of why

we believe what we do, supporting argumentation without constituting moments of episodic memory.

M&C take the importance of having and enforcing human social commitments as a reason for the development of episodic memory. I agree that commitments and auto-noesis are closely linked, but I wonder about the direction of explanation: It seems to me that in order to enter into a social commitment binding my future self, or to feel bound by my past commitments, I must already have the sort of sense of self, enduring in time, that is enabled by auto-noesis. M&C would be on firmer ground if proto-humans with mere event memory could already make social commitments, where the emergence of auto-noesis could then offer a way to strengthen and enforce those commitments. I wonder whether the power to make social commitments might not be just one of the many adaptive consequences of the human sense of self (Metzinger 2004), so that auto-noesis would be better explained by its contribution to that larger construct, with its diverse adaptive advantages.

### False memories, nonbelieved memories, and the unresolved primacy of communication

doi:10.1017/S0140525X17001455, e25

Robert A. Nash

Department of Psychology, Aston University, Birmingham B4 7ET, United Kingdom.

r.nash1@aston.ac.uk www.robert-nash.com

**Abstract:** Mahr & Csibra (M&C) make a compelling case for a communicative function of episodic remembering, but a less compelling case that this is its *primary* function. Questions arise on whether confirming their predictions would support their account sufficiently, on the communicative function of preserving rich, nonbelieved memories, and on the epistemic benefits of developing false memories via the acceptance of misinformation.

The target article contributes appreciably to the established literature exploring the social functions – among other functions – of accurate and inaccurate remembering (Bluck et al. 2005; Newman & Lindsay 2009). Mahr & Csibra (M&C) prompt us to rethink our conception of these social functions; specifically, they propose that remembering is adaptive primarily because communicating our memories can lead others to share our beliefs. The case for this communicative function is compelling, and M&C's account lays the ground for interesting new directions in empirical research, requiring novel methodological paradigms. But the case for the primacy of this function over other functions is currently unresolved.

What kinds of empirical evidence would strongly support or falsify the primacy of communication? M&C make some reasonable predictions, but support for these can arguably only bolster confidence in the existence of a communicative function, not provide evidence of its primacy. For instance, the authors predict that people should engage in more conservative, effortful source monitoring whenever a prospective listener is likely to be skeptical. This prediction seems complementary to the literature demonstrating “audience-tuning” effects on remembering (Echterhoff et al. 2008); more generally, there is broad agreement that metacognition is strategic, goal-driven (Johnson et al. 1993; Koriat & Goldsmith 1996b), and influenced by similar processes as is social persuasion (Blank 2009; Leding 2012; Nash et al. 2015). But does this broad agreement confirm that episodic remembering must primarily serve communication? Not at all. Indeed, although communicative goals undoubtedly can motivate source monitoring, these goals do not necessarily take precedence over other self-serving goals. When a skeptic challenges the authority of our memories, for example, we seem in fact to

systematically prefer cheap-and-easy strategies, not reliable strategies, for verifying the truth (Nash et al. 2017; Wade et al. 2014).

Additional questions arise when we stay on the matter of people disputing their own memories. M&C emphasize that believing in an event's occurrence does not necessarily imply remembering the event; however, they omit to note that the converse is also true. That is, people frequently retain episodic memories of events that they no longer believe truly occurred (Clark et al. 2012; Mazzoni et al. 2010; Otgaar et al. 2013; Scoboria et al. 2017). Importantly, these so-called *nonbelieved memories* often retain the rich, auto-noetic phenomenology that typifies believed memories. Our ability to preserve these memories could be adaptive, given that our reasons for disbelieving any particular memory may themselves transpire to be misguided (Scoboria et al. 2014). But the existence and characteristics of nonbelieved memories must nevertheless tell us that auto-noesis is more than simply “a proposition to the effect of ‘I had these experiences’” (M&C, sect. 1.1.2, para. 1), and that episodic remembering cannot, by necessity, be epistemically generative. M&C must account for the durability of auto-noesis in cases where a remembered event is not believed to have occurred.

Although M&C do not discuss nonbelieved memories, they do give greater attention to the adaptiveness of false memories. Susceptibility to false memories might offer numerous specific benefits to the rememberer (Bernstein & Loftus 2009; Howe 2011; Nash et al. 2016), but M&C propose that this susceptibility is also generally adaptive, because convincing ourselves of self-serving beliefs is an essential first step toward convincing others. They further propose a reciprocal benefit: adopting other people's beliefs into our own recollections can be “communicatively useful” as a means of enhancing our epistemic authority. Both of these proposals warrant scrutiny.

First, is the adaptiveness of (false) remembering really contingent on whether or not we communicate our memories to others? Many examples of self-serving memories give cause for doubt: Remembering plays well-documented roles in identity formation and maintenance, for instance, and so establishing positive self-regard – even if based on false beliefs – can provide important benefits to well-being (Conway & Pleydell-Pearce 2000; Wilson & Ross 2003). Episodic remembering enables us to generate and preserve self-serving beliefs about our own past such as these and also to generate the same self-serving beliefs in other people. But it seems counterintuitive to imply that the adaptive benefit of having others share our self-serving beliefs must be greater than the adaptive benefit of us holding those beliefs ourselves.

Second, does altering our memories to accord with another person's memories really afford greater epistemic authority? Suppose that Doris and Jack observe a theft, and Doris later claims that the thief had red hair, whereas Jack cannot recall the thief's hair. M&C correctly note that people typically treat the richness and detail in others' memory reports as signals of epistemic authority (Bell & Loftus 1989); the authors therefore suggest that Jack could become an ostensibly more authoritative source by integrating detail from Doris's memory into his own. Yet M&C might have equally noted that people are persuaded by good calibration: We trust witnesses who realize what they remember poorly, as well as what they remember well (Tenney et al. 2007). In this sense, even patchy memories – not only detailed memories – can signal epistemic authority. This interpretation makes it more difficult to construe misinformation acceptance as necessarily adaptive: Jack could gain greater authority as a witness precisely *because* rather than accepting the misinformation, he maintains that he cannot remember the thief's hair. Moreover, the benefit of accepting misinformation is even less clear in cases where memories are altered, rather than supplemented. Suppose that Jack initially recalls that the thief's hair was brown, but nevertheless alters his recollection to accord with Doris's (red hair). Here, Jack's testimony neither becomes more detailed as a result of accepting the misinformation nor