Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism*. Cambridge, England: Cambridge University Press.

A NEW TYPE OF MEMORY SYSTEM OR AN ADDITION TO AN OLD MEMORY SYSTEM?

Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past

By Kourken Michaelian. Cambridge, MA: MIT Press, 2016. 312 pp. Hardcover, $43.

Kourken Michaelian describes his book *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past* as an act of applied epistemology. In chapter 1 the author sets out the three core questions addressed by the book. For a psychological audience, the restatement of these core questions found in the final chapter is more relevant because it succinctly sets out the relationship between psychology and the philosophical inquiry pursued in this book:

> This book has had three main goals: first, to provide a general account of episodic remembering, as it occurs in real human beings, consistent with and shaped by the view of remembering as simulational mental time travel that has emerged in psychology in recent years; second, to provide a general account—again, based on the relevant psychology, including research on metamemory—of the factors ensuring the reliability of simulational remembering; and finally, to provide an account of the evolution of episodic memory, including the distinctive forms of consciousness which characterize it. (p. 237)

In the second chapter the author asks whether memory is a natural kind. By *natural kind* he means a category or distinction that exists in nature, although as he notes there is no one way of determining what constitutes a natural kind. Probably the closest he comes to an empirical definition is the idea that if memory, or a type of memory, can be identified as a natural kind then it will be possible to generalize within that natural kind. The approach taken is to evaluate memory systems with respect to their similarities and differences in terms of the information processing task they perform (the computational level), the procedure they use to accomplish the task (the algorithmic level), and the neurological mechanisms used to implement the procedure (the implementational level). Such an approach is familiar to memory researchers, and the author's take on this issue is both informative and properly cautious. We have no quarrel with the author's conclusion that semantic and episodic memory are similar in many aspects and are not that similar to procedural memory. In addition, it also seems likely, as the author concludes, that procedural memory and the other forms of what are commonly considered nondeclarative memory do not constitute a natural kind.

Nevertheless, we do have some misgivings about this approach. First, the computational level theories that can be proposed for memory systems are very different from the computational level theories discussed by Marr (1982). The prototypical task discussed by Marr was deriving shape from shading. With this task the physics of light and the optics of the eye determine the pattern of retinal activation. A mathematical (computational) analysis of how that retinal pattern provides information about shape can provide strong constraints on algorithmic theories. With memory, however, we do not know how the input to the task is represented, so the best we can probably do is to identify specific tasks, specify the inputs and the outputs, and describe in very general terms how the inputs can be transformed into the outputs (Humphreys, Wiles, & Dennis, 1994). When applied to a memory system that performs many different tasks, any computational level theory may be so general that it will be of very little use.

There is also a cost to the attempt to identify natural kinds in psychology. Explaining a difference between a procedural task such as a finger-tapping sequence and an episodic task such as cued recall by reference to different memory systems is a very weak explanation, as it inherits all the uncertainties involved in deriving the idea of there being separate systems. In addition, the very existence of such a simple explanation for the differences may divert attention away from other explanations such as the different ways the two tasks are cued. That is, the finger-tapping task does not use discrete cues, as in the cued-recall task, and the cues that are used for finger tapping are inherent within the task (if you can perform part of the task you are provided with the cues needed to perform the remainder).

In chapter 3 the author discusses the aspects from epistemology that the remainder of the book will take for granted. For the psychologist this provides a fascinating insight into issues that are usually ignored and to the style of argument used. However, there were two aspects to this argumentative style that caused us some concerns. First, in a hypothetical

example on p. 70 a person experiences an episode and forms a memory of it. The person then entirely forgets the episode. Later, the person is reminded of the episode and, using the information in the reminder and knowledge of similar episodes, forms a detailed mental representation of the episode that happens to be similar to the original episode. This and other hypothetical examples where memories disappear and then something like a memory reappears are used to refute some formal ideas about what constitutes a memory. Our concern is that if we ever observed such a situation we would not be able to rule out the possibility that a weak trace of the original episode had remained and had played a critical role in recreating an accurate memory. We are unsure about the extent to which this weakens the conclusions that can be drawn from these examples. Despite these misgivings, we do not disagree with the central thrust of the author's arguments that the output of the memory system is often constructed or, in the author's terminology, simulated.

In section II, the three chapters look at different philosophical conceptions of memory. Chapter 4 looks at the idea that memory is about the preservation of the past. This is referred to as the common sense conception. Chapter 5 looks at the idea that in order for something to be regarded as a memory there must be an appropriate causal connection between the original episode and the representation that is formed at retrieval. Chapter 6 looks at the simulation theory of memory. The simulation theory is the winner because there is now undeniable evidence that the representation that emerges from an attempt to remember can be constructed out of multiple episodes or out of the memory for an episode and semantic memory. However, the author's conclusion about the simulation theory is likely to be controversial among psychologists. That is, he concludes that there is no essential difference between remembering the past and imagining the past, or other forms of episodic imagination such as imagining the future. These chapters will be of interest to psychologists because they provide an introduction to the ways philosophers think about memory.

Section III of the book addresses the reliability of memory. The overall thesis of this section is that simulation does not necessarily imply unreliability. Chapter 7 explores the implications of the automatic incorporation of testimonial information into memory. Chapter 8 looks at whether source monitoring can increase the reliability of memory. Chapter 9 argues that in addition to the need to discriminate source, there is also a need to discriminate process. That is, there is a need not just to distinguish remembering something that happened to you from remembering something you imagined (source or reality monitoring) but to distinguish between remembering the past and imagining the past (process monitoring).

To a psychologist it is somewhat surprising that the author does not review the laboratory work establishing the reliability of memory (cf. Konkle, Brady, Alvarez, & Oliva, 2010). This appears to be due to a desire to examine reliability in ecological settings. However, this largely reduces the argument for reliability to the claim that an unreliable memory would be maladaptive. Nevertheless the argument that the automatic incorporation of testimony into memory does not necessarily reduce reliability is believable, but it could be stronger. In our book (Humphreys & Chalmers, 2016) we examined two situations where people learn from testimony. One was a situation where two or more people jointly rehearse the day's activities. The other involved one person questioning another in order to learn more about the questioner's concerns. These situations would take place largely between people who knew each other and who shared the same goals. Joint rehearsal would also ordinarily take place after a relatively short retention interval where your memory for some of the information being reviewed would be quite good, increasing your ability to evaluate the truthfulness and competence of the other person. Thus some of the main objections to the reliability of memories produced by the automatic incorporation of testimony, such as the difficulty people have in telling whether a stranger is lying, would not apply.

In chapter 8 the author conceptualizes belief-producing systems as involving two processes. The first process produces information intended to be the content of a belief, and the second process chooses whether or not to endorse the contents that have been produced. This distinction (but not the language) is familiar to psychologists, and applying the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) to it is generally straightforward. That is, psychologists would have no problems with the author's assumption that a word that is produced (emerges into consciousness) can be evaluated in order to make a decision about whether the word was presented or imagined, or whether it was seen or heard. A word that was presented is likely to be accompanied by more sensory detail, and a word that was imagined (generated) is likely to be accompanied by more information about the process. However,

one aspect of the source monitoring framework may be problematic for Michaelian's analysis. Johnson et al. (1993) noted that many decisions about source are made quickly, with no conscious experience of retrieving and assessing information. Michaelian refers to this as unconsciousness source monitoring and assumes that it plays the same role in his two-process account as does conscious source monitoring. However, this assumption may turn out to be incorrect. The unconscious side of the source monitoring framework is less well developed than the conscious side, which makes it difficult to evaluate whether they play the same role. In addition, there are alternative frameworks that do not fit as neatly into the author's two processes for belief-producing systems. In fact, Marcia Johnson, the principal developer of the source monitoring framework, has endorsed one of these, the Banks (2000) model, as a way to implement a fast unconscious process (Mitchell & Johnson, 2009). In this model the memory traces reside in a multidimensional space where the subject is free to set a decision axis. The representation of the probe item is then projected onto the decision axis, and the point on the decision axis is compared with a criterion on the decision axis for saying "yes." The meaning of the response depends on the decision axis. With one axis the participant might be deciding that the test item was source A, and with another axis the participant might be deciding that it was presented, not imagined (generated). The Banks model produces results that are similar to the global matching models (Humphreys, Pike, Bain, & Tehan, 1989). These models match a combination of cues against memory and base a decision on the strength of the match or the feeling of familiarity induced by the match. However, the outcomes from these models are interpretable only if the agent or system in some sense knows the decision axis or the cues that were used in the probe. The Banks model and the global matching models appear to have more in common with ideas presented in chapter 9 about the role of intention in memory than they do with the conscious part of the source monitoring framework.

Like us, many memory researchers have probably never thought about a need to distinguish the process of remembering from the process of imagination. If so, chapter 9 becomes a must read. The crux of the argument is that the process that produced something like the image of an event is not given directly to the agent. Instead, the agent who constructed the event must infer the process from a variety of heuristics based on both content and phenomenology. The heuristics are based on average differences between constructive processes and are therefore fallible. The author discusses a wide range of potential heuristics. We would like to focus on one of these and to propose an additional heuristic. The author discusses Urmson's (1967) proposal that the agent knows whether he or she intended to remember or imagine. The author accepts intent as an important component of process monitoring but points out that it does not apply in cases of involuntary remembering. This proposal also seems to imply a direct knowledge of one's intentions, but an indirect route that may or may not lead to a conscious awareness of one's intentions is possible. This indirect route is related to our previous remarks about alternatives to the source monitoring framework. Memory appears to be cue dependent (Humphreys & Chalmers, 2016), so it seems possible that if the agent or the memory system knows something about the cues used, an inference can be made not only about remembering but about the type of remembering (e.g., episodic or semantic, or whether the word was studied auditorily or visually). This suggestion is supported by the observation that instructions to produce a synonym or antonym to a cue word can result in as fast or even faster responding than instructions to produce the first word that comes to mind (Davidson & Cofer, 1968). It does not appear that there is any need to decide whether the word that emerges into consciousness is a synonym or antonym. Likewise, if you are using a contextual cue there may be little need to determine, after access, whether the representation that emerges into consciousness is an episodic not semantic memory. In many cases of involuntary remembering it is also possible to infer the cues. For example, a representation of an event constructed while you are wandering about the campus on your 40th reunion is likely to be attributed to the cues supplied by the current visual scene and hence a memory not imagination. This would be especially compelling if the event had no natural connection to the visual scene.

With simple episodic queries (e.g., the provision of one member of a studied word pair) and semantic queries (e.g., a general knowledge question with a one-word answer), subjects typically either recall the correct answer or recall nothing. For example, instructions to adopt a more lenient criterion for outputting a response typically make very little difference (Koriat & Goldsmith, 1996). This some-or-none characteristic tends to break down with more complex memories and longer retention intervals, but it can still be present. For example, when we asked a

colleague to remember a birthday party, she reported on an episode involving a young girl attending her son's birthday party. The girl emptied a package of chips on the newly cleaned floor and then stamped on them. When questioned further about what the girl was wearing, she reported that she could not remember, but she could readily imagine a young girl's party dress. The probable reason she could not remember such a detail is because it was irrelevant to the reasons she had for repeatedly retelling the story. This suggests that at least some of the time even very old memories of complex events that have been previously recalled and possibly contaminated by some imagination can still be distinguished from pure imagination.

Section IV of the book is concerned with when and why episodic imagination, which includes episodic remembering, evolved. Chapter 10 discusses different forms of consciousness with special attention to autonoesis (a feeling that you participated in the event) and chronesthesia (a feeling that the event occurred in time). We were convinced by the arguments that the hypothesis that autonoesis and chronesthesia play a functional role in episodic remembering and its evolution are worth pursuing. However, we think it may be premature to define episodic memory in phenomenological terms. The memory literature discussed by the author involves mostly memory for events (Rubin & Umanath, 2015). There are also the laboratory paradigms of paired-associate learning, free recall, recognition, and the like. We think an analysis of the computational, algorithmic, and implementational levels would probably conclude that they, along with the memory for events, constitute a natural kind. The phenomenology accompanying paired-associate learning has not been adequately investigated. However, it seems likely that it does not involve the sense of self or time that generally accompanies event memory. The phenomenology of single-item recognition has been more adequately investigated, but it is not clear that the states of consciousness identified in studies of single item recognition play a functional role (see Dunn, 2008, and the review in Humphreys & Chalmers, 2016). It also seems possible to explain this difference in phenomenology. In Humphreys and Chalmers we reviewed evidence showing that an additional cue that subsumes some of the targets in a paired-associate list eliminates interference from other list items and prior memories. Such a cue can also help explain intrusions from different episodes by suppressing the information in those episodes that would eliminate the intrusion from consideration. Thus cues that specifically target information about what is to be recalled may suppress the recall of information about when or where the information was studied and any phenomenology associated with the *when* or *where* information.

Chapter 11 is concerned with why episodic memory, so defined, evolved. The crux of the argument is that autonoesis and chronesthesia would have been helpful, along with source monitoring and with the content criteria discussed in chapter 9 for process monitoring. We generally agree with this argument, but we also think that the previously discussed some-or-none characteristic also has a role. That is, it is possible for a child to learn that, when attempting to remember, if a representation emerges into consciousness it is generally correct. For example, young children listening to an adult read a story will tell the adult that they have skipped a page, and the adult then reads the skipped page. This allows the children to use their short-term memory of having predicted the next page to confirm that their longer-term memory for what was to come next was correct. Another opportunity for confirmation comes when a child remembers where their mother has placed, or generally places, a toy and then proceeds to that spot to find it. Note that in these examples we are ignoring distinctions between episodic and semantic memory, which would be irrelevant to the child, and are focusing on simple memories, well-learned information, and relatively short retention intervals. Memories with these characteristics probably account for most memories about which we receive confirmation or disconfirmation and are probably very common in the everyday life of a young child.

In support of the idea that forward planning evolved before episodic remembering, the author notes that because construction is essential for forward planning, this could explain the constructive nature of remembering. This is possible, but there are other possibilities that do not involve forward planning. For example, Tehan, Humphreys, Tolan, and Pitcher (2004) showed that at times memories were constructed out of semantic information from the cue and phonological information from the study list. That is, if you cue with *tool,* the presence of *trench* in the study list results in the recall of *wrench*. Humphreys and Chalmers (2016) also asked whether the memory traces of sentences, faces, and events are stored in different formats in the hippocampus or whether they are stored in the same format. If the latter, it would be necessary to convert what was

retrieved from the hippocampus into the appropriate format using long-term (semantic or possibly procedural) memory about the regularities found in sentences, faces, events, and so on. Likewise, learning that may result from repeated encounters over different contexts or in different modalities seems to require a degree of construction. In addition, in Humphreys and Chalmers (2016) we provided several examples of how the basic ability to recover a representation of a prior episode in conjunction with the cues provided by a recurrent environment could have been the beginning of learning how to plan for the future. That is, there is the beginning of a theory about how forward planning could develop from a sufficiently robust "what, where, and when" memory, but there is no indication how forward planning could have developed in the absence of such a memory. This does not deny that autonoesis and chronesthesia might not have evolved until there was a need to discriminate a "what, where, and when" memory from forward planning. One could conceptualize this either as the evolution of a new form of memory or the evolution of a new feature added to the existing memory system. The latter is our preference.

The author has chosen to approach the very difficult question of understanding the phenomenology accompanying episodic remembering by providing a detailed description of the phenomenon. Although we think that at times it would have been helpful to have systematic surveys of a variety of respondents (cf. Rubin, Schrauf, & Greenberg, 2003), the author is to be commended for the clarity of the description provided. An alternative approach to this very difficult problem is to look for analogous processes. One way to understand the author's proposal is that, as he notes, we are able to represent time, probably through a spatial analogy. Thus there may be both an embodied and a more abstract representation of time. Likewise, our sense of self may involve both embodied and more abstract representations. The embodied representations may get activated during the course of remembering an event. This might occur because these representations are part of what was stored about the event or because the abstract components were used as retrieval cues that cued both the event memory and the embodied components about time and self. With these ideas it is not surprising that some people report that remembering an episode from their past would be similar to their initial experience of the episode. They would not actually believe they were experiencing the episode because there would be too many other indicators that they were not. When the process is phrased in this fashion, there appear to be parallels with instances of multimodal perception (Hannah & Humphreys, 2017) and synesthesia (Chiou & Rich, 2014). For example, in the McGurk effect (McGurk & MacDonald, 1976) the perception of a spoken phoneme in conjunction with seeing lips speaking a different phoneme can lead to the perception of hearing a third phoneme. Other examples can be found in the many interactions between smell and taste. For example, westerners report that vanilla smells sweet, which probably occurs because of its association with desserts (Auvray & Spence, 2008). These examples can be described as an inferential process that integrates perceptual information and information based on learned associations. The inferential process does not result in awareness of all the information sources nor of any decision-making process. However, it does manifest itself as a somatic sensation, which would not only supply information but could also motivate behavior. It seems possible that this is an evolved mechanism for making some of the otherwise hidden workings of the brain available to awareness. Once such a mechanism evolved, it could be adapted for additional problems such as discriminating remembering from imagination.

As a final comment, as we wrote our book it became apparent that much research was being conducted in silos where the researchers in one silo were not communicating with the researchers in other silos. There were also instances where earlier research and theories had been forgotten. This failure to communicate and remember may be as big a problem for psychology as the replication crisis (Open Science Collaboration, 2015). The opportunity provided by the *American Journal of Psychology* for authors with different backgrounds to comment on each other's books is an important step toward addressing the communication and remembering problems.

Michael S. Humphreys
School of Psychology
The University of Queensland
QLD 4072, Australia
E-mail: mh@psy.uq.edu.au

Kerry A. Chalmers
School of Psychology
The University of Newcastle
Psychology Building, University Drive
Callaghan
NSW 2308, Australia
E-mail: Kerry.chalmers@newcastle.edu.au

REFERENCES

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition, 17,* 1016–1031.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11,* 267–273.

Chiou, R., & Rich, A. N. (2014). The role of conceptual knowledge in understanding synesthesia: Evaluating contemporary findings from a "hub-and-spokes" perspective. *Frontiers in Psychology, 5,* 105. doi:10.3389/fpsyg.2014.00105

Davidson, E. H., & Cofer, C. N. (1968). Some determinants of controlled association times. *Journal of Experimental Psychology, 78,* 200–207.

Dunn, J. C. (2008). The dimensionality of the remember–know task: A state-trace analysis. *Psychological Review, 115,* 426–446.

Hannah, S. D., & Humphreys, M. S. (2017). *Functional amodality: Intersection, memory and task set in the emergence of abstraction*. Unpublished manuscript, the University of Queensland.

Humphreys, M. S., & Chalmers, K. A. (2016). *Thinking about human memory*. Cambridge, England: Cambridge University Press.

Humphreys, M. S., Pike, R., Bain, J. D., & Tehan, G. (1989). Global matching: A comparison of the SAM, Minerva II, Matrix and TODAM models. *Journal of Mathematical Psychology, 33,* 36–67.

Humphreys, M. S., Wiles, J., & Dennis, S. (1994). Toward a theory of human memory: Data structures and access processes: *Behavioral and Brain Sciences, 17,* 655–667.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21,* 1551–1556.

Koriat, A., & Goldsmith. M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19,* 167–188.

Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746–748.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What we have learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin, 135,* 638–677.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi:10.1126/science.aac4716

Rubin, D. C., Schrauf, R. R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition, 31,* 887–901. doi:10.3758/BF03196443

Rubin, D. C., & Umanath, S. (2015). Event memory: A theory for laboratory, autobiographical, and fictional events. *Psychological Review, 122,* 1–23.

Tehan, G., Humphreys, M. S., Tolan, G. A., & Pitcher, C. (2004). The role of context in producing item interactions and false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 107–119.

Urmson, J. O. (1967). Memory and imagination. *Mind, 76*(301), 83–91.

MUCH (MUCH) MORE THAN MONKEY BUSINESS

A Natural History of Human Morality

By Michael Tomasello. Cambridge, MA: Harvard University Press, 2016. 180 pp. Hardcover, $64.

A Natural History of Human Morality is the fourth book in a series that includes *The Cultural Origins of Human Cognition* (1999), *Why We Cooperate* (2009), and *A Natural History of Human Thinking* (2014), in which Michael Tomasello focuses on the critical differences between human and great ape cognition and the evolution of these differences. This, his latest, is the work of a scholar who has been at the top of his game for at least two decades; it is original, lucid, and carefully conceived and constructed. It also includes helpful summaries and diagrams to facilitate access. Readers, both those new to Tomasello's work and those familiar with it, will find much to appreciate in this standalone text (although readers who fall into the latter category might be forgiven for feeling the faintest sense of déjà vu).

In this short book (163 pages), which is jam-packed with empirical detail, Michael Tomasello tells the story of how ultra-social and ultra-cooperative moral human agents evolved from social but uncooperative great apes in the space of some 1.6 million years (*Homo heidelbergensis* was thought to be fully cooperative some 400,000 years ago). The details consist in empirical results from multiple comparative studies of great ape and human behavior, in particular the behavior of 3- to 5-year-old human children, who are considered closest to great apes in terms of cognitive capacity. Many of these studies are Tomasello's own or those of his collaborating partners, and the results inform his "plausible" theory of the evolution of human morality (p. 3). In short, Tomasello identifies a number of moral psychological capacities that are unique to our species and then defends an empirical hypothesis that attempts to explain how and why these capacities evolved. Two of his main