

Sven BERNECKER: *Memory. A Philosophical Study*. New York: Oxford University Press. 2010. viii + 276 pp. ISBN: 9780199577569.

Finding an entire philosophy book dedicated to memory is rare, when clearly memory is not even close to becoming a hot topic in philosophy of mind (which paradoxically contrasts with the enormous interest that it stirs up in other fields, such as cognitive science or sociology). That is a very good reason to welcome and read Bernecker's new book *Memory: A Philosophical Study*, which is a kind of renewed and improved version of his previous book *The Metaphysics of Memory* (2008)¹.

The book is a wide, punctilious and analytic exposition of the causal theory of memory, which had already been advocated in an article full of ideas written in 1966 by Martin and Deutscher called "Remembering"².

The main principles of Bernecker's account are the same as those expounded in "Remembering"; for a present propositional representation to be a genuine memory of a past event, three objective conditions must be met: a continuous causal relationship between the past representation of that event and the present one via a memory trace, a similarity of content between the two, and the truth of the memory representation. Let's see how Bernecker develops and argues in favor of these conditions.

The causal condition, which constitutes the core of the causal theory of memory, is discussed in chapter 5. In order to exclude from the category of memory both cases of epistemic luck (like suggestion) and cases of relearning, the causal condition establishes that the relationship between the past and present representations must be guaranteed by a persisting memory trace or a contiguous series of memory traces that derives from the past representation and that causes the present one. According to Bernecker, memory representations are brought about by mental memory traces, which are dispositional beliefs that preserve the mental content through time, or subdoxastic states if the content stored is non-conceptual, and which supervene on neurobi-

ological memory traces. From the reading of this section, we find out that mental memory traces are not mere epiphenomena without causal power because they give rise to representations and behavior, that they are opaque to the subject, and that they are not structural analogues of the event represented by the past experience (as Martin and Deutscher had advanced). We also find out that memory traces can be removed from the biological body and thus can be intra and interpersonally transplanted, and even replicated, and continue to give rise to authentic memories. This is compatible with Bernecker's defence in chapter 3 of the notion of quasi-memory and therefore of a psychological theory of personal identity, that is, that memory can be used as a criterion to define personal identity because memory does not imply it. Nonetheless, these eight pages about memory traces in a book of 250 pages seem too schematic and synthetic to fully characterize the core of the causal theory of memory. Bernecker is not an exception; most of the philosophers who have written about memory never analyzed memory traces in depth, maybe due to the force of the metaphorical image which is very easy to grasp. However, this kind of transparent analogy should not justify the omission of a proper characterization of memory traces, especially in a book entirely dedicated to the defence of the causal theory.

The last part of chapter 5 sheds light on the other central notion of the causal theory, that is, on the relation between the different elements involved in the causal bond. On this subject, Bernecker explains the causal dependence between the past representation and the present one in terms of counterfactual dependency, whereas the causal dependence of memory states on memory traces vis-à-vis retrieval cues is analyzed in terms of necessary and sufficient conditions.

The second condition of the causal theory, the content condition, is developed in chapters 6, 7 and 8. The first two chapters constitute a defence of pastist externalism about memory content, that is, the idea that the content of both the past and the present

representations is fixed by and so supervenes on the same past environment, and that is why they are similar. Bernecker's argument resorts to different world-switching scenarios, which allows him to rule out internalist accounts and other possible content externalisms, like presentist and futurist versions (as well as the extended mind thesis). Chapter 8 addresses the kind of relationship between the content of the two representations: Bernecker rejects the idea that memory simply duplicates a past experience (what he calls the identity theory but which in fact very few philosophers defend), but he also disapproves strong forms of constructivism about memory, because they lose sight of the first function of memory that according to his view is the "preservation of the content". But this rejection of constructivism is clearly at odds with the social and pragmatic main functions that the current psychology literature attributes to memory: planning for the future, decision-making, construction of personal identity, guiding present and future actions, are now more associated with the concept of memory than the idea of preservation of the past.³ Furthermore, we could ask to what extent the implicitly assumed preservationist view about memory content differentiates his position from the identity theory. Although Bernecker calls himself a "moderate generativist", his generativism only applies to the justification of knowledge: memory is a generative source of justification because memory can remove defeaters. But memory cannot generate anything new: it cannot bring about new elements of justification or generate new content. In fact, in regard to memory content, his position is incontrovertibly preservationist: the content of the memory representation must be entailed (in the relevance sense of "entailment") by the content of the past representation without the need of additional premises, which certainly forbids any enrichment or increase of content. So the "changes" allowed are minimal: substitutions of synonyms, adjustments of verb tenses seem all that can be done. Even if Bernecker did not mention these "principles of semantic similarity" developed in his previous book (158–163), they seem nonetheless to be implicitly assumed here. And unfortunately, together

with the omission of these principles, Bernecker omitted a remarkable conception that he should definitely have continued to develop: the context-dependence or pragmatic sensitivity of these principles, that is, the idea that what counts as an accurate or veridical memory representation is defined to some degree relative to the interests and intentions of the remembered and the attributors of the conversational context (chapter 10 of his previous book).

Finally, the truth condition is mentioned in chapter 8 and chapter 3. The requirement is that the memory representation must be true in a correspondence sense. However, the past representation need not be true at the moment of forming it. Probably, one reason that motivates Bernecker to defend this conception is his intention to distinguish his theory from the epistemic theory of memory, which considers that memory representations are states of knowledge. To this idea, Bernecker responds in chapter 3: memory does not imply past or present justification, memory does not imply past or present belief, and memory does not imply a true past representation. Nonetheless, it implies a true memory representation, which seems to lead to the anti-Fregean consequence that a change in the world can convert a false memory into a true one, or a true memory into a false one, even if the causal and the content conditions are met. One example: I formed at t_1 the justified but false belief that John has borrowed a blue book from the library; if I remember this at t_2 my memory is false, but if at t_3 without my knowledge John effectively borrowed a blue book from the library, when I remember at t_4 the representation I formed at t_1 my memory is true. Once again, this use of the sense of "false memory" moves Bernecker away from the conception of falsity common among psychologists which mainly refers to memory distortion and intrusion, false recognition, delusional memory, confabulation, etc.⁴ In psychology, there are three cases in which a memory representation can be false: a. the past representation was never formed; b. the past representation was false due for example to a perceptual error; c. the past representation was true but during the retrieval process there were some errors on the "editing"

functions or on the temporal context or other source monitoring. Therefore, in the example given above, on standard use in psychology my memory would be false not only at t_1 but also at t_4 because my past representation, even if justified at t_1 , was false. Unlike Bernecker, for psychologists the false/true distinction of memory representations is not reduced to the truth value of the propositions implied on those memory representations.

The truth condition is required for factive attitudes, like knowing, but not for memories that are “introversive”, that is, memories that represent an attitude toward a content, like memories of beliefs, desires, doubts. The only requirement in this last case is the authenticity of the memory representation, which is defined in terms of functional identity (direction of fit and polarity). The distinction between introversive and extroversive memories is outlined in chapter 1, as well as other non-traditional distinctions between different types of memories. In my opinion these distinctions constitute the weakest point of Bernecker’s proposal. Bernecker’s two main typologies are guided by a grammatical criterion: first, extroversive memories are divided according to the grammatical complement of the verb “remember”, which gives rise to four types of memories: memories of objects, memories of properties, memories of events (gerundive construction) and propositional memories (that-clauses). Only these last cases, that is, when the verb “remember” is followed by a that-clause, require the possession of concepts, and are the object of analysis in the book. This distinction however is problematic. Beyond this grammatical taxonomy these four types of memory do not seem to correspond to psychological kinds. If we take these three examples: *a.* remembering visiting the Colosseum; *b.* remembering that I visited the Colosseum; *c.* remembering that Paris is the capital of France, according to the grammatical criterion *a* and *b* would correspond to different memory kinds, i.e. event-memory and propositional memory respectively, even if the content remembered is the same: an event of my life, whereas *b* and *c* would belong to the same kind, even when in one case I remember an event experienced before and in the other case a simple fact of the world. Moreover, I

could remember *a* without having any concept, but it seems that I could not remember *b* without having the concepts of “self”, “visit”, “Colosseum” and “time”. This shows that not only the distinction between gerundive constructions and that-clauses is problematic to establish by itself different kinds of memory, but what is more, this grammatical distinction seems to be implicitly based on other criteria that go beyond grammar, that is the different modalities in which a subject can experience a memory: experiential memory vs detached memory (or episodic memory vs semantic memory according to the most recent distinction made by Tulving,⁵ which is not centred on the type of content: events vs facts, but on the phenomenology). This would be the only way to explain why *a* and *b* differ, while *b* and *c* belong to the same memory type. But even if we leave these problems behind, Bernecker’s grammatical criterion is exclusively based on the analysis of English, leaving aside the fact that some other languages do not have the same kind of constructions, especially in what concerns the gerundive form.

On the other hand, the second main distinction, that is between extroversive and introversive memories, is also grounded in a grammatical difference: while one contains only a simple that-clause, the other one also contains a second order that-clause: *I remember that I (past tense of an attitude verb) that p*. But again there seems to be some difficulties with this criterion: in the case of some verbs, like “believe”, the type of memory depends exclusively on the omission or presence of this verb. So, if I decide to say that “I remember that I believed that p”, the truth or falsity of p is irrelevant to state a genuine memory, but if I instead omit the verb “believe” and say “I remember that p”, p has to be true for my memory to be true; so in this case, the applicability of the truth condition would depend on the arbitrariness of the linguistic choice of the remember.

These problems show that grammar cannot be used by itself as a criterion to establish different psychological kinds of memory; it can only be used as an analytical tool to designate different kinds of memory that differ according to other criteria, such as their

content (events vs facts), phenomenological properties (knowing vs recalling with rich sensory detail and mentally time travelling) and/or source (perception vs thought vs emotion). In fact, surprisingly Bernecker himself says that “memory reports are sometimes not reliable indicators of whether a given state is a propositional or non-propositional memory” (23) or an extroversive or introversive memory, so it is hard to understand why he creates a memory taxonomy based on a criterion that he himself recognized as unreliable, especially when he dismissed the phenomenological criterion for this same reason, for not being “sharp”. He owes us at least a positive argument in favor of the grammatical taxonomy, as well as why that-clauses constitute a single memory kind.

This focus on the objective conditions and not on the phenomenological aspect of the mental representations does not provide a criterion to distinguish between dreams, imaginings, explicit memory, familiarity (“know experience”), etc. Even if Bernecker reduces his analysis to conscious explicit propositional memories, as he announces in chapter 1, if the field of application of the causal account is extended, it can be deduced that provided all these mental representations have the same kind of content (namely propositional content) and fulfill the conditions enumerated before, the subject is in the same mental state, that is, in a genuine memory state, regardless of the different ways in which the subject experiences it. So a person who read Don Quixote in the past might be remembering that he read Don Quixote, or imagining that he read Don Quixote, or he might know that he read it but not remember any details associated with the episode, or he might be doubting whether he read it; and any of these conscious mental states would satisfy the necessary conditions for being a genuine memory state. As we can see, the causal, content and truth conditions cannot account for the specificity of memory states that make them different from other mental states. This clearly constitutes a gap in the causal theory of memory, a gap that seems to be even more outstanding when in return Bernecker proposes a grammatical typology that is blurred. Furthermore, Bernecker reduces his analysis

to conscious explicit propositional memories that are non-inferential, that is, that they are not the result of a reasoning based on additional premises. But the problem is that in everyday life, non-inferential memories occur only rarely, like memories of shopping lists, telephone numbers and “someone told me that” or “I left my glasses on the desk”. Because most of our memories are inferential, as Bernecker himself recognizes (25), an account exclusively focused on non-inferential memory has limited explanatory power.

At this point, we can ask ourselves why this causal account of memory phenomena should be assumed. Chapter 4 tries to give an answer to this question, arguing that the causal theory of memory offers the most viable explanation of memory processes, because it provides a comprehensive account of the kind of process that is responsible for the capacity to retain and represent again something represented in the past. This ability is expressed through a counterfactual relationship between the past and present representation and can be summarized in the intuitive conditional: *If S hadn't represented at t_1 that p he wouldn't represent at t_2 that p^** . However, the problem is not just that Bernecker again does not give any argument in favor of the theory (he dismisses three possible objections but also two possible arguments in favor), but that the only opponent theory that he evaluates and rules out, the retention theory, that is, the idea that once we represent something we only acquire a disposition to represent it again, is an old-fashioned idea that nobody explicitly or implicitly defends, because as Bernecker notices himself, it does not explain the essential point: how we retain this disposition over time. That is why the retention theory is not a real opponent to the causal theory, and that is why finally nothing convinces the reader that the causal explanation is more plausible than other “non-existent” theories. We just have to accept it. And decidedly, it is not hard to do so, because the causal theory is highly intuitive. At least at first glance. But if we take a look to the cognitive science literature, reconstructivism in memory seems to be the new dogma. And reconstructivism not only claims that “retrieving a memory is like reconstructing a dinosaur from fragments

of bone”⁶ but that memories cannot be “simple, countable connections between two well-defined moments of consciousness”⁷ because the content of memories is not determined at encoding alone, due to the fact that the continual processing of new information transforms the old information stored and, what is more, that the overall context of retrieval (present thoughts, interests, expectations) also contributes to the determination of the content of memory.⁸ As we can see, reconstructivism in memory seems to be in tension with the two main principles of the causal theory: the causal condition and the content condition, which at the same time arouses some doubts about the truth of the counterfactual relationship. Maybe it is possible to represent something at t_2 and have a genuine case of memory without having represented something similar at a past time. More in this direction should be explored, because here could lie a potential and respectable opponent of the causal explanation of the memory phenomena.

To round up, I personally would like to have found a battle-hardened defence of the causal theory of memory, including some arguments in favour. Unfortunately, this is absent in this book: we have to assume the causal theory because there does not seem to be any better option. Despite this absence and other weaknesses, *Memory: A Philosophical Study* is worth it because while inviting us to discuss with the author, it reminds us how many ques-

tions about memory deserve a deeper debate among the philosophical community, questions whose answers are indispensable for a better understanding of the mind.⁹

Marina TRAKAS

Macquarie Centre for Cognitive Science,
Macquarie University, Australia
& Institut Jean Nicod, Ecole des Hautes
Etudes en Sciences Sociales, France

1. Bernecker 2008.
2. Martin & Deutscher 1966, 161–196.
3. See for example, Bluck, Alea, Habermas & Rubin 2005, 91–117.
4. See for example, Kopelman, 1999, 197–214.
5. Tulving 2005, 4–56.
6. Schacter 1996, 69.
7. Schechtman 1994, 8.
8. Campbell 2004.
9. *Further reading*: Adams, Fred 2011: “Husker Du?”. *Philosophical Studies* 153 (1), 81–94; Bernecker, Sven 2011: “Further Thoughts on Memory: Replies to Schechtman, Adams, and Goldberg”. *Philosophical Studies* 153 (1):109–121; Goldberg, Sanford 2010: “The Metasemantics of Memory.” *Philosophical Studies* 153 (1), 95–107; Schechtman, Marya 2010: “Memory and Identity”. *Philosophical Studies* 153 (1), 65–79.

REFERENCES

- Bernecker, Sven 2008: *The Metaphysics of Memory*. Dordrecht: Springer.
- Bluck, Susan, Alea, Nicole, Habermas, Tilmann & Rubin, David 2005: “A Tale of Three Functions: the Self-Reported Uses of Autobiographical Memory”. *Social Cognition* 23 (1), 91–117.
- Campbell, Sue 2004: “Models of Mind and Memory Activity”. In: Peggie DesAutels & Margret Urban Walker (eds.), *Moral Psychology: Feminist Ethics and Social Theory*. Maryland: Rowman & Littlefield, 119–137.
- Kopelman, Michael 1999: “Varieties of False Memory”. *Cognitive Neuropsychology* 16 (3/4/5), 197–214.
- Martin, C. B. & Deutscher, Max 1966: “Remembering”. *The Philosophical Review* 75(2), 161–196.
- Schacter, Daniel L. 1996: *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books.
- Schechtman, Marya 1994: “The Truth About Memory.” *Philosophical Psychology* 7(1), 3–18.
- Tulving, Endel 2005: “Episodic Memory and Autonoesis: Uniquely Human?” In: Herbert S. Terrace & Janet Metcalfe (eds.), *The Missing Link in Cognition*. New York: Oxford University Press, 4–56.