

# From Minimal Traces to Scenarios of the Past: A Neuro-Computational Model on Regaining Categoricity and Compositionality in Remembering

Francesca Righetti<sup>1</sup>, Zahra Fayyaz<sup>2</sup>, Laurenz Wiskott<sup>2</sup>, Markus Werning<sup>1</sup>

<sup>1</sup>Department of Philosophy II, Ruhr University Bochum, Bochum, Germany

<sup>2</sup>Institute for Neural Computation (INI), Faculty of Computer Science, Ruhr-University Bochum, Bochum, Germany

\*Corresponding author: [markus.werning@rub.de](mailto:markus.werning@rub.de)

## Abstract

This paper presents a proof of principle for Trace Minimalism (Werning, 2020), a novel philosophical framework for episodic memory. Trace Minimalism claims that remembering does not involve the storage of representational content but rather the reconstruction of past scenarios through the interaction of minimal traces with semantic information. Minimal traces establish a causal link to prior experiences but lack categorical and compositional content. We provide a neuro-computational model using a vector-quantized autoencoder and a transformer-based semantic completion mechanism. Our findings support the hypothesis that remembering is possible without representational memory traces and that minimal traces, in interaction with semantic information, reliably construct past scenarios. The results offer a compelling alternative to classical representational theories of memory while maintaining causal continuity with past experiences.

**Keywords:** episodic memory; trace minimalism; categoricity; compositionality; vector quantized variational autoencoder.

## Introduction

Episodic memory is often conceived as the retrieval of stored representations, a view mainly supported by the Classical Causal Theory (Martin & Deutscher, 1966). This theory provides two claims about memory traces: (i) they act as a causal link between the original experience and the recollection of the event; (ii) memory traces carry representational content. These traces represent the original event accurately. The representational content is structural analogous to the remembered event. This allows the memory trace to serve as an accurate representation of the foregoing experience.

However, the Classical Causal Theory faces significant challenges. Research shows that memory is not an exact replication of past events but undergoes a construction process that integrates information from various sources. Research supports the idea that memory and imagination rely on a shared system, i.e., the constructive simulation system, (Addis, 2018; Schacter & Addis, 2007; Szpunar et al., 2009), emphasizing a move toward a dynamic process involving construction and simulation. Moreover, if a memory trace (almost) completely represented the original content, remembering would not allow to generate new knowledge (see Werning & Cheng, 2018; Werning & Liefke, 2024).

In response, alternative theories have been proposed. Some researchers have proposed constructive causal theories to account for the generative aspects of episodic memory (cf. Michaelian, 2011; see also Bartlett, 1932; Hassabis & Maguire, 2007; Robins, 2016a). Others

propose that remembering needs no memory traces. Simulationism (Michaelian, 2016) holds that memory involves simulations of past events from available information, without requiring a direct causal link to the original experience. Remembering, in this view, is a simulation based on current available information rather than retrieval of stored information. This aligns with evidence that memory supports imagination and scenario construction. Simulationism suggests memory traces are not necessary as long as a reliable simulation system integrates information from different sources. However, its rejection of a causal link raises concerns about memory's reliability. Werning (2020), in fact, argued that memory causal links with past events, implying that some causal connection is essential for reliable remembering.

Werning's Trace Minimalism offers a potential solution by proposing that memory traces are minimal and non-representational. According to this theory, memory traces are sparse neural information of the distributed neural activity pattern produced during the original past experience. Minimal traces do not carry representational content but only sparse information that interacts with semantic information carried in the neocortex to generate a simulated scenario of the past (cf. Cheng, Werning & Suddendorf, 2016). This process allows for the construction of a reliable and accurate representation of the past event, without transmitting representational content over time. The resulting past scenario is representational *again* once the minimal trace is combined with semantic information, enabling them to reliably generate accurate representations of the past. Due to changes in neocortical synaptic connectivity over time, the interaction between semantic information and minimal memory traces may also lead to the generation of a false memory. To ensure truth-approximating reliability, Werning proposes an analogy with the predictive processing framework of perception (Friston, 2005), where minimal traces act as "error signals" in a top-down model. The brain uses these minimal traces to generate memory representations of past events. This process mirrors predictive processing in perception, where sensory input informs predictions about the present. In memory, minimal traces interact with learned regularities to form a predictive model of the past.

Thus, Werning's Trace Minimalism offers a compelling alternative to traditional theories of memory. By focusing on the interaction between minimal traces and semantic information, this framework provides a reliable account of memory without relying on the transmission of

representational content. This paper provides computational support for Trace Minimalism through a neural network model that demonstrates how mnemonic reconstruction can occur without stored representational content. By implementing a vector-quantized autoencoder with a transformer-based semantic completion mechanism, we show that minimal traces, when supplemented with statistical regularities, can reliably regenerate categorical and compositional representations. Our findings suggest that memory functions through a constriction system rather than retrieval of representational items, providing a compelling alternative to contemporary theories.

## Philosophical Assumptions and Framework

Trace Minimalism challenges classical assumptions about memory by rejecting the idea that stored traces must carry representational content. It is distinguished by five key claims:

**(M1)** Remembering a past scenario is causally linked to a foregoing experience of that scenario by a *minimal trace*.

**(M2)** Minimal traces are non-representational: i.e., these do not carry either categorical nor compositional content, but only sparse information of the neural state underlying the foregoing experience.

**(M3)** A memory representation is constructed through the interaction of the minimal trace with statistical regularities acquired over a lifetime.

**(M4)** Categoricity and compositionality are regained in the scenario construction during remembering, thus, the resulting mnemonic state is representational.

**(M5)** The process of scenario construction is reliable: i.e., the resulting mnemonic state represents the past scenario in a sufficiently accurate way with a probability greater than an appropriate, task-dependent reliability threshold.

Thus, Trace Minimalism proposes that remembering occurs without relying on full memory representations. In traditional accounts, "remembering" is often equated with the retrieval of stored representations. However, Trace Minimalism distinguishes between "remembering" as the immediate mental event (episodic representation) and "memory" as a persistent mental state that links past events to the present. The recollection process in this view is less about retrieval and more about construction—the generation of a past event from minimal, sparse information.

To frame these claims philosophically, we turn to the Representational Theory of the Mind (RTM) (Fodor, 1975, 1998; Smolensky, 1995). In this context, a mental state is considered representational if it satisfies three conditions:

**(R1) Categoricity:** The primitive bearers of representational content are categorical. That is, they subsume objects and events in the world under categories (see also Harnad, 2017).

**(R2) Compositionality:** Complex bearers of representational content are compositional. That is, primitive bearers of representational content can be

structurally/syntactically combined to form complex bearers of representational content such that the representational content of the complex state is a structure/syntax-dependent function of the representational contents of its structural/syntactic parts (see Partee, 1984; Recanati, 2012; Stewart & Eliasmith, 2012; Werning, 2001, 2004, 2005a, 2005b, 2012; Werning et al., 2005, 2013).

**(R3) Causal-nomological dependency:** As being a bearer of representational content, the internal state stands in a causal-nomological dependency relation to the objects/events (and the sets and sets thereof) that the internal state represents. In this paper, we assume a reliabilist account according to which the causal process leading from the objects/events (and, respectively, the sets and sets thereof) to the internal state needs to be reliable for the internal state to represent the objects/events (and the sets and sets thereof) in question.

According to Trace Minimalism, minimal traces are not representational, i.e., non-categorical and non-compositional, but, through their interaction with semantic knowledge, the resulting scenario becomes representational. Thus, the reconstruction process in memory satisfies the conditions of categoricity, compositionality, and causal-nomological dependency, as the reconstructed memory reflects the past with a sufficient degree of reliability.

To further support Trace Minimalism, Werning proposes an artificial neural network model that simulates memory recall. The network is trained to perceive and categorize images, learning both categorical patterns (e.g., shapes) and compositional rules (e.g., digit sequences). During the test phase, the network must recall these patterns from minimal traces, which contain only a small amount of information, with the rest being reconstructed using semantic knowledge. The model demonstrates that minimal traces alone are insufficient for full recollection, but when combined with semantic knowledge, the traces can produce accurate memory reconstructions in a reliable way. Before presenting the neuro-computational model, we present Trace Minimalism's predictions.

## Predictions

Trace Minimalism makes key predictions concerning categoricity and compositionality:

1. In perception, the neural vector, derived from a sensory input  $n$  (such as a digit image), can accurately categorize the input  $n$  as its corresponding concept  $N$  with a probability equal to or greater than a defined reliability threshold.
2. The probability for the memory trace to correctly categorize the input  $n$  with its corresponding concept  $N$  is below the threshold. The trace is considered non-categorical, as predicted by Trace Minimalism.
3. In remembering, the neural vector can be accurately categorized as concept  $N$  with a probability equal to or greater than the reliability threshold.

4. Categoricality is general, i.e., the model can reliably categorize new numeral images.
5. In perception, the probability of accurately identifying the numerical concept  $N$  formed by concatenating two inputs  $n_1$  and  $n_2$  is based on the composition of their respective categorizations from the neural vectors.
6. The probability for the memory traces to accurately identify the concatenated numerical concept  $N$  falls below the threshold, indicating a significant drop in compositional reliability.
7. In remembering, when the input is grammatical (i.e., obeys rules), the probability of accurately identifying the concatenated numerical concept, based on the categorizations of the neural vectors, equal to or greater than the reliability threshold. Conversely, when the input is not grammatical, this probability drops below reliability.

### Methodology: Neuro-Computational Model

To empirically test the claims and predictions of Trace Minimalism, we designed a neuro-computational model that simulates memory recall using a vector-quantized autoencoder (VQ-VAE) and a transformer-based semantic completion mechanism (Fayyaz et al., 2022; Reyhanian, Fayyaz, & Wiskott, 2024). This model is used to investigate whether episodic recall can be reliably performed without storing detailed representational traces.

The model consists of three stages (Figure 1):

- **Encoding and quantization:** Perception starts with an image of a two-digit numeral, processed by an encoder into a feature vector. This vector is then quantized by mapping continuous features to discrete codebook vectors, producing perceptual quantized vectors. This models the visual processing pathway from the retina to higher visual areas. Despite compression, quantized vectors remain categorical and compositional.
- **Masking and trace storage:** A random subset of perceptual vectors is selected using a mask, which discards a percentage of features (the masking level) and stores the rest as a trace vector. This simulates attentional filtering and memory storage in the hippocampus. At high masking levels, the trace loses both categorical and compositional properties.
- **Reactivation, completion, and decoding:** During recall, the trace vector is retrieved and completed using a transformer network (Vaswani et al. 2017), restoring it to a mnemonic quantized vector. Though it may differ from the original perceptual vector, it regains categoricality and compositionality. In our model, we decode the trace back into an image to better compare the remembered items with the perceived ones, but in reality, of course this does not happen, only an internal remembered vector of the scenario is generated.

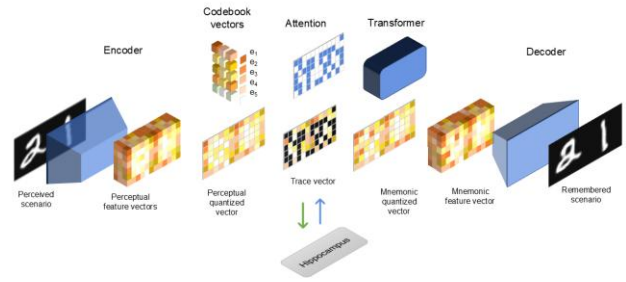


Figure 1. Causal chain from perceptual scenario through trace vector to the remembered scenario in the computational model: The model consists of an encoder module, a trace module, and a decoder module. The encoder transforms an input image, depicting a scenario to be perceived and remembered, into a highly compressed quantized vector. The trace module removes some percentage of the quantized vector by masking and stores the resulting incomplete vector as a trace. The decoder module includes a transformer, which fills in the missing parts in the trace vector to recover a complete quantized vector, and a proper decoder, which transforms the quantized vector back into an image depicting the remembered scenario.

### Experimental Design

#### Training Procedure

The model was trained using a dataset of two-digit numerals created by concatenating MNIST digits (LeCun, 1998), with separate sets for training and testing. A grammatical training set was formed from numerals divisible by three, following specific rules: numerals from  $\{0, 3, 6, 9\}$  pair with the same set, while  $\{1, 4, 7\}$  pair with  $\{2, 5, 8\}$ , and vice versa, which corresponds to the rule that the numeral is divisible by three. Three grammatical numerals (00, 36, 72) were intentionally excluded from training, and no ungrammatical numerals were used for training. To test generalization, some valid numerals were held out, and non-grammatical numerals were included in the test set.

#### Biological significance

The VQ-VAE’s encoder is comparable to the feedforward processing in the visual system, leading to abstract object encoding, while the decoder mirrors feedback connections from higher to lower visual areas, modeling recall processes. However, the model does not claim to replicate exact brain activation patterns. The transformer learns statistical relationships within the VQ-VAE, similar to how the brain learns semantics from experiences, and can fill in missing information like higher cortical areas do.

#### Measuring Categoricality and Compositionality

Categorization performance is assessed using a classifier trained on individual digits. A numeral is deemed categorical if its reconstructed representation can be reliably classified as

its original number concept. The reliability threshold is set at  $r_1 = 80\%$  accuracy.

To evaluate compositionality, we analyze whether the model correctly reconstructs two-digit numerals based on the independent categorization of their constituent digits. Since this is a more complex task, we lowered the reliability threshold from  $r_1 = 80\%$  down to  $r_2 = r_1^2 = 64\%$ .

### Determining Minimality

Minimality is achieved when the trace lacks representational content but still enables reconstruction when integrated with semantic knowledge. This occurs at a masking level where direct classification from the trace is unreliable, but classification from reconstructed outputs meets the reliability threshold.

## Results

Our results provide robust empirical support for Trace Minimalism, demonstrating that minimal traces, when complemented with semantic information, enable reliable episodic recall. The key findings focus on three areas: compositionality in two-digit numerals, prototypicality in memory recall, and the role of semantic information in reconstructive memory processes.

We tested whether, at a critical masking level, the mnemonic quantized vector remains categorial, while the minimal trace does not. The model was fed two-digit numerals, and their perceptual quantized vectors were masked at varying levels to produce trace vectors. These were decoded both without and with semantic completion via a transformer. The focus was on categorizing the second digit (results would be symmetrical for the first digit) using a classifier. Results (Figure 2) show that without semantic completion, reliable categorization ( $\geq 80\%$ ) is only possible up to a 25% masking level. With semantic completion, reliability improves significantly, allowing masking up to 54% when the first digit is fully masked. When both digits are masked equally, compositionality influences results—helping or hindering based on grammaticality. The critical masking level is 62% for grammatical numerals, marking the minimal trace threshold, with a reliability gap of about 33%.

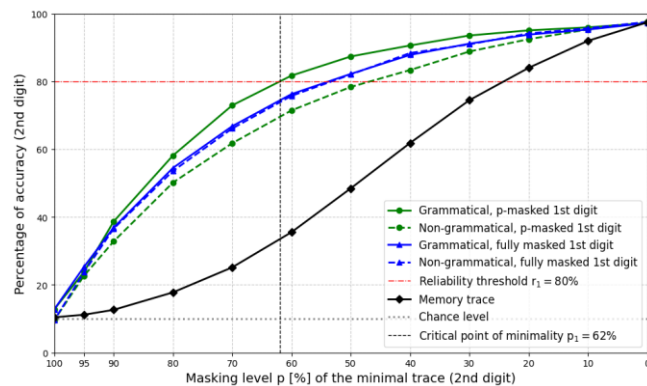


Figure 2. Re-Emergence of Categoricity: A trace vector of a two-digit numeral is decoded either with semantic

completion by the transformer (green and blue lines) or without completion (black line), and then the reliability of categoricity is determined on only the second digit, which had a varying masking level  $p$ . The green lines show the realistic case where the first digit had a masking level of  $p$  as well, while the blue lines show results for a fully masked (100%) first digit as a control. Solid lines indicate grammatical cases, dashed lines non-grammatical cases. Reliability is generally much better with semantic completion and best if the numeral is grammatical and not fully masked (green solid line) and therefore compositionality can help as well. The minimal trace is realized at a critical point of minimality at the masking level of 62% where reliability of categoricity reaches the threshold of 80%.

We examined the information capacity preserved as numeral images are processed from input to minimal trace and back to remembered scenarios. Information capacity was estimated by compressing images using Python’s zip library (Deflate algorithm). For minimal traces, it was approximated by multiplying the quantized vector’s capacity by  $(1 - \text{masking level})$  since masking reduces information, though compressed files may have larger size on disk due to the added random mask complexity. Results (Figure 3) show that the minimal trace has significantly reduced information capacity. Even the quantized vector is highly compressed compared to raw images. Interestingly, remembered images have a higher information capacity than perceptual ones because they include graded gray scales, unlike the nearly binary input images—highlighting that the encoding-decoding process is inherently lossy.

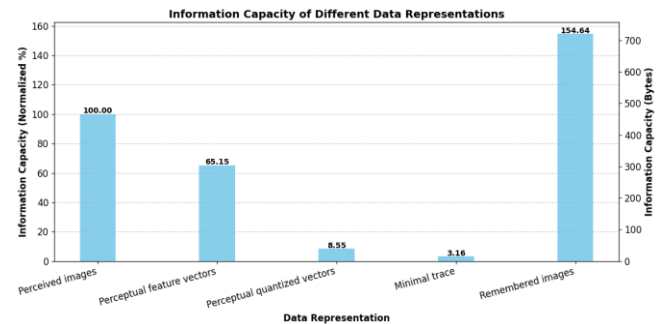


Figure 3. Information capacity across stages: The figure shows the average number of bytes used per image in a zip file in percentage (left ordinate) and in absolute terms (right ordinate). The value of the minimal trace was derived from the value of the quantized vector by multiplying with  $(1 - p)$ .

Expanding on earlier analyses, we assessed compositionality in full two-digit numerals. Using grammatical, non-grammatical, and held-out test sets, we evaluated categorization accuracy with a reduced reliability threshold (64%) to accommodate the increased complexity of two-digit combinations. Results (Figure 4) showed that semantic completion significantly improves recall performance, with higher accuracy for grammatical numerals compared to non-grammatical ones. The model generalized

well to unseen (held-out) numerals, indicating that the transformer effectively learned the underlying grammatical rule (divisibility by three). The critical masking level was identified at 63%, like for the two single-digit classification case, suggesting near-independence in digit categorization but with slight compositional effects.

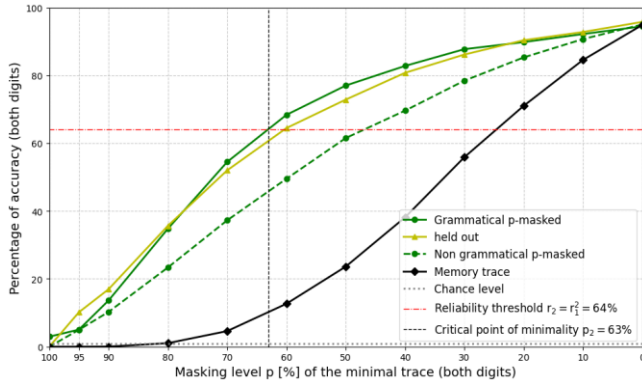


Figure 4. Compositionality: Lines show the reliability of categoricity for trace vectors of two-digit numerals at varying masking level  $p$  decoded either with semantic completion by the transformer (green lines) or without completion (black line). Reliability is generally much better with semantic completion and best if the numeral is grammatical (green solid lines) rather than non-grammatical (green dashed line), which is an indication of compositionality. Interestingly, the transformer generalizes to grammatical cases it has not been trained on (held out set, light-green solid line). The minimal trace is realized at a critical masking level of 63% where reliability reaches the threshold of 64%.

A key finding is that grammatical compositionality is reliably achieved at this critical point, while non-grammatical numerals show an accuracy drop of around 19%. Without semantic completion, accuracy falls to about 10%, confirming that minimal traces alone lack the necessary information for compositional recall.

We observed a clear prototypicality effect in the model’s recall behavior. Memories tend to be biased toward more typical or central representations, consistent with psychological findings across various stimuli (Hemmer & Steyvers, 2009; Bird et al., 2014; Kerrén et al., 2024). Results (Figure 5) reveal that the standard deviation within digit classes decreases as we move from original images to remembered scenarios, indicating that the reconstructed images are more prototypical. This suggests that memory recall is not merely about reproducing past experiences but involves a shift toward statistical norms or “averaged” representations.

Our findings highlight the essential role of semantic information in reconstructive memory. The model relies on statistical regularities learned during training, which function as the foundation for semantic knowledge:

- The VQ-VAE captures image statistics, such as stroke patterns typical of handwritten digits. While it generalizes well to similar data (e.g., handwritten

letters), it struggles with dissimilar inputs like fashion images.

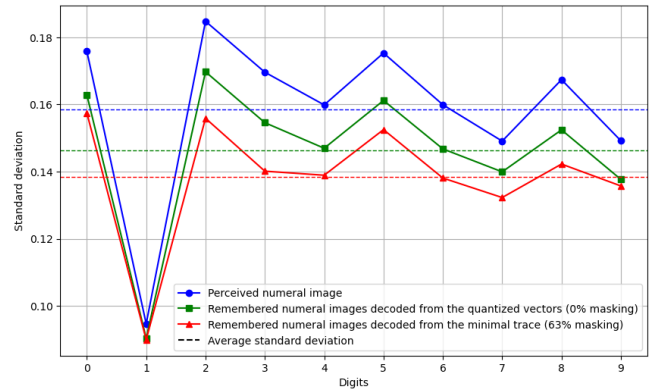


Figure 5. Emerging prototypicality measured: The curves show the standard deviation of gray values per pixel within individual digit categories across different stages of the architecture: perceived numeral images (top blue line), remembered numeral images decoded from the quantized vectors (0% masking, middle green line), and remembered numeral images decoded from the minimal trace with semantic completion (63% masking, bottom red line). Lower standard deviation indicates higher prototypicality, and we see that it is the remembered scenario from the trace vector that is most prototypical. The range of the gray values in images is from 0 to 1. The low value for the digit 1 is probably due to the fact that it consists of the shortest stroke and benefits most from the normalizing effect of centering.

- The transformer learns both within-digit regularities (supporting single-digit completion) and between-digit dependencies (e.g., recognizing patterns related to divisibility by three). This supports compositionality and allows generalization to unseen numeral combinations. Crucially, without the transformer, the VQ-VAE alone cannot maintain reliable remembering. While a supervised model could be trained to categorize minimal traces directly, it would still need to learn the same semantic dependencies, reinforcing the conclusion that minimal traces require semantic completion for remembering.

## Discussion and Conclusions

In this article, we provide a proof of principle for the claims of Trace Minimalism (Werning, 2020) through a biologically inspired neural network model. The model was trained to recognize handwritten two-digit numerals expressing numbers divisible by three, thereby learning statistical regularities and acquiring semantic information about numeral shapes, number concepts, and grammatical rules governing divisibility. The network was then tested in an episodic recall task with various degrees of information loss, demonstrating key aspects of Trace Minimalism at a critical point of minimality, where trace information capacity was only 3.16% of the perceived numeral images.

Our neural network model confirms Trace Minimalism's claims through its ability to reconstruct numerals despite extensive information loss. Minimal traces alone were insufficient for categorization or composition, validating (M1) and (M2). However, interaction with semantic information enabled the successful reconstruction of numerals, confirming (M3). The model demonstrated the reliable regain of categoricity and compositionality, highlighting how semantic information compensates for minimal traces (M4). Moreover, the model generalizes learned patterns to novel numerals and held-out numbers, further confirming (M5).

The model also exhibited sensitivity to grammaticality, reliably applying divisibility rules in recall tasks. Categorization reliability significantly dropped for non-grammatical numerals, demonstrating the model's ability to generalize and apply learned rules rather than simply memorizing specific examples. Empirical parallels include findings that people remember structured information more effectively, such as chess players recalling valid board positions better than random ones (Frey & Adesman, 1976).

Trace Minimalism presents a compelling alternative to existing theories of memory. Trace Minimalism provides a middle ground between representational and non-representational theories of memory. It preserves the causal role of memory traces while rejecting the need for stored content. By denying that experiential content is transmitted in the trace (M2), Trace Minimalism better explains memory reconstruction through interaction with semantic information (M3-M5). This account also provides a reply to Michaelian's claim that remembering needs no memory traces. According to Michaelian, since there is no way to determine how much reconstructed content must stem from information obtained during the foregoing experience, one can propose the slippery-slope argument: if no threshold can be set, then memory traces may not be necessary at all (cf. Michaelian, 2016). Our results challenge this claim by demonstrating that it is, in fact, possible to determine the minimal amount of information from the original experience required for reliable remembering. At this threshold of minimality, memory traces carry just enough information to reliably warrant categorization and compositionality when supplemented with semantic information. This directly calls into question the validity of Michaelian's slippery-slope argument for Simulationism.

Our neurocomputational model strongly supports Trace Minimalism as a viable framework for episodic memory. Remembering occurs not through retrieval of full representations but through reconstruction from minimal traces, aided by semantic knowledge. This approach offers a neurobiologically and philosophically plausible alternative to traditional theories, reinforcing the idea that memory is a constructive process rather than a direct retrieval of past experiences.

## Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) with funding for the research group FOR 2812 Constructing scenarios of the past: A new framework in episodic memory, on the side of L.W by the grant 397530566 (P5) and on the side of M.W. by the grants 419040015 (P6), 419038924 (P3.1) and 397530566 (P3.2, P10). For their helpful comments, we thank Sen Cheng and Jonathan Najenson as well as the audiences of the Bochum workshop Memory and Language and the Grenoble workshop Memory and Causation.

## References

- Addis, D. R. (2018). Are episodic memories special? on the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand*, 48(2-3), 64–88.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.
- Bird, C. M., Berens, S. C., Horner, A. J., & Franklin, A. (2014). Categorical encoding of color in the brain. *Proceedings of the National Academy of Sciences*, 111(12), 4590–4595.
- Cheng, S., Werning, M., & Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neuroscience and Biobehavioral Reviews*, 60, 82–89. doi: 10.1016/j.neubiorev.2015.11.011
- Fayyaz, Z., Altamimi, A., Zoellner, C., Klein, N., Wolf, O. T., Cheng, S., & Wiskott, L. (2022). A model of semantic completion in generative episodic memory. *Neural Computation*, 34(9), 1841–1870. doi: 10.1162/neco.2015.01520
- Fodor, J. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. New York, NY: Oxford University Press.
- Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory & Cognition*, 4(5), 541–547.
- Friston, K. J. (2005). Hallucinations and perceptual inference. *Behavioral and Brain Sciences*, 28(06), 764–766.
- Harnad, S. (2017). To Cognize is to Categorize: Cognition is Categorization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (second edition) (Second Edition ed., pp. 21–54). San Diego: Elsevier. doi: <https://doi.org/10.1016/B978-0-08-101107-2.00002-6>
- Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, 11(7), 299–306. <https://doi.org/10.1016/j.tics.2007.05.001>
- Hemmer, P., & Steyvers, M. (2009). A bayesian account of reconstructive memory. *Topics in cognitive science*, 1(1), 189–202.

- LeCun, Y. (1998). The MNIST database of handwritten digits. Retrieved from <http://yann.lecun.com/exdb/mnist/>
- Martin, C. B., & Deutscher, M. (1966). Remembering. *Philosophical Review*, 75, 161–196. (ISBN: 0031-8108)
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238.
- Michaelian, K. (2011). Generative memory. *Philosophical Psychology*, 24(3), 323–342. <https://doi.org/10.1080/09515089.2011.559623>
- Michaelian, K. (2016). *Mental time travel*. Cambridge, MA: MIT Press.
- Partee, B. H. (1984). Compositionality. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 281–312). Dordrecht, Netherlands: Foris Publications
- Recanati, F. (2012). Compositionality, semantic flexibility, and context-dependence. In M. Werning, W. Hinzen, & E. Machery (Eds.), *Oxford handbook of compositionality* (p. 175–191). Oxford: Oxford University Press.
- Reyhaniyan, S., Fayyaz, Z., & Wiskott, L. (2024, September). Analysis of a generative model of episodic memory based on hierarchical vq-vae and transformer. In *Proceedings of the 33rd international conference on artificial neural networks (icann 2024)*, lugano, switzerland. Springer Nature Switzerland. doi: 10.1007/978-3-031-72341-46
- Robins, S. (2016). Representing the past: memory traces and the causal theory of memory. *Philosophical Studies*, 173(11), 2993–3013. <https://doi.org/10.1007/s11098-016-0647-x>
- Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 773–786. doi: 10.1098/rstb.2007.2087
- Smolensky, P. (1995). Connectionism, constituency and the language of thought. In C. Macdonald & G. Macdonald (Eds.), *Connectionism* (pp. 164–198). Cambridge, MA: Blackwell. (Original work published 1991)
- Stewart, T., & Eliasmith, C. (2012). Compositionality and biologically plausible models. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 596–615).
- Szpunar, K. K., Chan, J. C. K., & McDermott, K. B. (2009). Contextual processing in episodic future thought. *Cerebral Cortex*, 19(7), 1539–1548. doi: 10.1093/cercor/bhn191
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008
- Werning, M. (2001). How to solve the problem of compositionality by oscillatory networks. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the cognitive science society* (pp. 1094–1099). London: Lawrence Erlbaum Associates.
- Werning, M. (2004). Compositionality, context, categories and the indeterminacy of translation. *Erkenntnis*, 60(2), 145–78. doi: 10.1023/B:ERKE.0000012876.85940.b3
- Werning, M. (2005a). Right and wrong reasons for compositionality. In M. Werning, E. Machery, & G. Schurz (Eds.), *The compositionality of meaning and content* (Vol. I: Foundational Issues, pp. 285–309). Frankfurt: De Gruyter.
- Werning, M. (2005b). The temporal dimension of thought: Cortical foundations of predicative representation. *Synthese*, 146(1/2), 203–24. doi: 10.1007/s11229-005-9089-2
- Werning, M. (2012). Non-symbolic compositional representation and its neuronal foundation: Towards an emulative semantics. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 633–654). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199541072.013.0031
- Werning, M. (2020). Predicting the Past from Minimal Traces: Episodic Memory and its Distinction from Imagination and Preservation. *Review of Philosophy and Psychology*, 11, 301–333. doi: 10.1007/s13164-020-00471-z
- Werning, M., & Cheng, S. (2018). Doing without metarepresentation: Scenario construction explains the epistemic generativity and privileged status of episodic memory. *Behavioral and Brain Sciences*, 41, e34. doi: 10.1017/S0140525X17001534
- Werning, M., Hinzen, W., & Machery, E. (Eds.). (2012). *The Oxford handbook of compositionality*. Oxford: Oxford University Press.
- Werning, M., & Liefke, K. (2024). Remembering dreams: parasitic reference by minimal traces in memories from non-veridical experiences. In D. Gregory & K. Michaelian (Eds.), *Dreaming and memory: Philosophical issues*. Springer (in press).
- Werning, M., Machery, E., & Schurz, G. (Eds.). (2005). *The compositionality of meaning and content* (Vol. I: Foundational Issues). Frankfurt: Ontos Verlag.
- Werning, M., Tacca, M., & Mroczko-Wasowicz, A. (2013). High- vs low-level cognition and the neuro-emulative theory of mental representation. In *Models, simulations, and the reduction of complexity* (pp. 141–152). Berlin: De Gruyter.