

# Are Situation Models Embodied?

Beige Jin<sup>1</sup>, Yucheng Wu<sup>2</sup>, Cheok Kei Lo<sup>3</sup>

<sup>1</sup>Peking University, Beijing, China, 1900014901@pku.edu.cn,

<sup>2</sup>Charles Wright Academy, Seattle, Tacoma, 1185362838@qq.com

<sup>3</sup>Fujian Normal University, Fujian, China, locheokkei914@gmail.com

## ABSTRACT

Researchers have proposed that readers construct situation models when reading texts. According to Zwaan's event-indexing model, a situation model can be analyzed in five dimensions: space, time, entity, causation, and motivation. There is robust behavioral evidence showing the existence of situation models from these five dimensions. However, the representational format of situational models, i.e. whether they are embodied or disembodied, is still controversial. Behavioral evidence is unable to answer this question, and only neural evidence testing both situation model construction and modality-specific activation can solve this problem. Three promising paradigms are sorted out, but not all of them give out expected results. We suggest that future experiments could focus on situated conceptualization, application of TMS, and depth of understanding.

**Keywords:** *situation model, embodied, modality-specific*

## 1. INTRODUCTION

Consider the following text: "Eddy and Mary plan to spend the weekend night in the house. Mary prepares dinner in the kitchen, and Eddy plays games in the room, and when Mary finishes the cooking, Mary asks Eddy to help put the dinner on the table. After a while, Mary comes back from outside and has dinner with Eddy."

When reading the paragraph above, you might imagine Mary as an elderly lady, although the text doesn't offer any information about age and gender. You might also feel confused about the last sentence, although there is no logical or grammar contradiction. It seems that people have the ability to read beyond the text. How could that happen?

Researchers have proposed that readers construct mental representations of situations based on the description of the texts and their knowledge of the world. These representations are termed *situation models* [1]. When people read the text, they do not merely comprehend the surface structure and explicit information of the text, but also make sophisticated inferences and create a more detailed representation of the text [2], [3]. One of the most prominent theories about situation models is *the event-indexing model*. It claims that event representations are the basic units of situation models, which are composed of five dimensions: space (where),

time (when), entity (what/who), causation (how), and motivation (why) [3], [4].

In this essay, we firstly sort out some behavioral evidence demonstrating the existence of situation models from these five dimensions. Then we discuss the representational format of situation models, i.e. whether they are embodied or disembodied, and assess the existing evidence. At last, we analyze the limitations of current researches and give some suggestions about future directions.

## 2. BEHAVIORAL EVIDENCE FOR SITUATION MODELS

There are plenty of behavioral experiments on situational dimensions supporting the existence of situation models. The general prediction here is that if the readers construct situation models while reading, then the situational information implied in the text would influence the readers' comprehension and memory, in ways that are consistent with our experience of situations in real life, which can't be predicted merely by the form of the text.

### 2.1.Space

Spatially closer objects are usually more accessible to us, and so it is more predictable in a situation model. Researchers have found that readers are faster in item recognition and reading speed on objects that are spatially

associated with the protagonist than objects that are spatially disassociated [5].

Also, implied spatial information changes people's interpretation of the same word. For example, in "The tractor is just approaching the fence", participants interpret "approaching" as meaning a larger distance than in "The mouse is just approaching the fence" [6].

Researchers have reported that it takes people longer to read when the spatial information is inconsistent with what is implied in the previous text. For example, participants spend more time reading "She approached the mummies quietly" if it follows "Carmen went out the museum" than if it follows "Carmen went into the museum" [7]. You have probably experienced the same effect when you're reading the last sentence of the first paragraph in the introduction.

## **2.2. Time**

Similar to space, temporal proximity in the described events also influences the accessibility of information. For example, participants spend more time recognizing "enter" after reading "An hour ago, John entered the building" than "A moment ago, John entered the building" [8]. Similarly, participants' responses to probe words denoting the protagonist's present occupation are faster than those denoting the protagonist's past occupation [9].

Besides, a mismatch between the narrated and chronological order of events (e.g. "Before he jumped the gate, he patted the dog.") undermines readers' comprehension performance [10]. This phenomenon could be explained by our default that events happen in chronological order in situation models, because this is how we experience events in daily life.

## **2.3. Entity**

The readers infer the existence of different entities in different situations. For example, after reading "Paige went to the restaurant and ordered a steak; after her meal, Paige chatted a bit with the waiter and then she left.", participants would not expect to hear more about the waiter when Paige later entered a movie theater [11]. Similarly, recognition time for "hammer" is shorter when the participants are presented "Jason began pounding the nails into the board" than "Jason pounded the nails into the board" [12].

The readers also make inferences about the properties of entities, such as information about a protagonist's stereotypical gender and an object's physical characteristics. For example, after reading "The babysitter settled down to watch a video", participants would slow down their reading speed if the following sentence is "Then he heard the baby crying", compared with "Then she heard the baby crying" [13]. Another example is that readers' recognition of an object picture would be faster if

it matches the object's implied orientation or shape in the previous sentence [14], [15].

## **2.4. Causation**

People naturally infer causal relationships between real-world events, as well as events described in the text. In one study, participants who read the sentence pair "Mark poured the bucket of water on the bonfire. The bonfire went out." answer faster to "Does water extinguish fire?" than those who read the sentence pair "Mark placed the bucket of water by the bonfire. The bonfire went out." [16] That's probably because the causal knowledge that water extinguishes fire is activated and integrated into the situation model in the former sentence pair but not the latter. This inclination for causation is so strong that readers would form causal relationships between events even if they are not described in adjacent sentences [17].

## **2.5. Motivation**

Readers infer the protagonist's motivation from his/her conditions and actions, which enables them to predict his/her future actions. In an experiment [18], participants are asked to read "After standing through the 3-hour debate, the tired speaker walked over to his chair (and sat down)". The presence of "and sat down" makes no difference to participants' recognition time to "sat" later, indicating the priming effect of "sat" even when the surface information is omitted.

Besides, if a goal hasn't been fulfilled, it may haunt us all the time. The same is predicted by a situation model. Researchers have found that an unsatisfied goal is kept highly available in readers' mental representation [19] and is more accessible than a goal that is just accomplished by the protagonist [20].

In short, there is strong behavioral evidence in all the five dimensions that people construct situation models when reading narrative texts.

## **3. THE EMBODIMENT OF SITUATION MODELS**

### **3.1. Endorse an Embodied Situation Model**

Given the existence of situation models, one may naturally ask what the representation format of situation models in the brain. The traditional framework of language comprehension suggests *the disembodied view of situation models*. According to this view, readers generate amodal propositional representations which connect text content to knowledge about the world. Situation model construction is based on symbolic and abstract representations, having nothing to do with the perceptual systems [21].

However, it is difficult to see how the disembodied view could account for the richness of information that people infer beyond the texts. To construct a fine-grained situation model, thousands of amodal propositions are probably required. However, we don't have good reasons to agree on this assumption, considering the limitations of working memory [22].

The embodied theory of situation models states that situation models are instantiated by the same sensorimotor neural representations formed while experiencing real events [21]. By "embodied" we refer to the framework of embodied cognition proposed by Barsalou [23]. People interact with real events in the world via multiple modalities, and corresponding modality-specific neuron systems are activated. These activation patterns are then integrated and stored in long-term memory as "simulators". Later, when readers comprehend the events in the text, those simulators are triggered to produce a "simulation", which is a re-enactment of previous experience. Crucially, this simulation involves activation in the same modality-specific areas when we interact with real events in the first place [23]–[25]. A situation model is constructed on the foundation of such simulation.

The embodied theory could explain the richness of information in a situation model. Sharing the same neural network with perception, perceptual simulation entails rich information about space, time, etc., and such information is easily available and at little processing cost, compared with generating amodal representations [22].

### ***3.2. Behavioral (non-) Evidence for Embodiment of Situation Models***

How can we support the embodiment of situation models? Behavioral evidence can only support the existence of situation models, but fail to distinguish the disembodied view from the embodied theory. Consider the evidence that readers interpret the implied distance of "approach" differently in "The tractor is just approaching the fence" and "The mouse is just approaching the fence" [6]. The embodied theory hypothesizes that people produce different simulations for these two sentences, thus inferring different spatial information. Nevertheless, the disembodied view could also insist that people elicit different propositional knowledge about distance when comprehending these two sentences. In principle, all the behavioral results about situation models are compatible with the disembodied view, at least post hoc [23]. Thus, they are not good evidence for the embodiment of situation models.

### ***3.3. Neural (non-) Evidence for Embodiment of Situation Models***

In our opinion, only certain kinds of neuroimaging evidence can address the problem of whether situation models are embodied or disembodied, because the

disembodied view and the embodied theory make contradicting predictions here. If situation models are based on amodal propositional representations, people would not activate modality-specific areas when constructing situation models. In contrast, if situation models are indeed rooted in simulations, then activities in modality-specific areas should be observed.

Thus, we propose that any successful experiment that supports the embodiment of situation models should give positive answers to these two questions:

Q1. Do participants construct situation models in the task?

Q2. Are activities in modality-specific areas observed?

Following this criterion, three promising experimental paradigms are sorted out from the literature. They all adopt discourse comprehension to answer Q1, as numerous previous studies have shown that participants construct situation models when reading stories [26]. In response to Q2, fMRI is implemented throughout the experiments.

The first paradigm is *the coherence paradigm*. In one study, participants read coherent stories and scrambled stories in which sentences are unrelated to each other. At the behavioral level, participants have a better memory performance when sentences are coming from coherent stories than scrambled stories, which demonstrates their engagement in situation model construction when reading coherent stories and their failure in constructing a situation model when reading disconnected sentences. At the neural level, when participants are reading coherent stories, there is an increase in activity in dmPFC, posterior parietal cortex, and perisylvian language areas, which, however, are all modality-nonspecific brain areas [27]. This original result fails to support the claim of embodied situation models. But Kurby & Zacks make a refinement to the experiment years later. Using the same materials, they code the texts into high-imagery clauses (auditory, motor, and visual) and low-imagery clauses before the experiment. They find that reading motor imagery clauses are associated with increases in activity in the secondary somatosensory and premotor cortex and reading auditory imagery clauses are associated with increases in activity in the secondary auditory cortex. Importantly, the modality-specific effects only occur when participants are reading coherent stories [21]. Combined with behavioral results, Kurby & Zacks' research provides evidence that participants are generating simulations to construct a situation model.

The second paradigm is *the narrative shifts paradigm*. In short stories, some clauses imply a change in one or more situational dimensions (i.e. space, time, entity, causality, and motivation). In face of such narrative shifts, participants have to update their situation models. Experimenters ask participants to read short stories containing narrative shifts and have found an increase in

activity in the precuneus, cingulate cortex, and temporal cortex [28]. A similar activation pattern is found when participants listen to the stories rather than read them visually [29]. Importantly, this neuron network is largely overlapped with the network that responded to event boundaries while viewing movies of everyday events [30], which suggests that the same neural representations underlie narrative event comprehension and real event perception. Speer and his colleagues further investigate which region is responsible for changes in a particular situational dimension. They have found that, for example, in response to changes in objects, participants' motor and somatosensory cortex in the hemisphere controlling the dominant hand are more activated [31]. These results demonstrate that certain dimensions of situation models are represented in an embodied way.

The third paradigm is *the inconsistency paradigm*. Short stories are presented that contain information contradicting the global text content. Participants spend more time reading those inconsistent word(s), indicating the existence of a situation model and the detection of situation model violation [32]. The detection is quickly followed by an attempt to integrate the inconsistent information and update the situation models in particular dimensions. Researchers have found that, compared with a consistent version of the stories, stories with inconsistent emotional information elicit activation in the ventromedial prefrontal cortex and the extended amygdala complex, while stories with inconsistent temporal information elicits activation in the precuneus and frontoparietal network [33]. Note that these are all modality-nonspecific areas. What's worse, with the same paradigm but shortened two-sentence stories, researchers fail to find any significant effect of inconsistency [34]. Possible reasons for this null result could be the shallow processing and the failure of constructing a strong situation model within two sentences. Anyway, the inconsistency paradigm doesn't support the embodiment of situation models.

To summarize, behavioral evidence cannot rule out the disembodied view in principle. Only neuroimaging experiments meeting our criterion can potentially prove the embodiment of situation models, and we do identify some qualified paradigms. However, not all the paradigms give out expected outcomes. Evidence for the embodiment of situation models is still relatively few and not robust enough.

#### **4. FUTURE DIRECTIONS**

Given the current researches on the representational format of situation models, some problems arise. First, how could we find more evidence to test the embodiment of situation models? Second, most experiments so far have been fMRI studies, which are unable to demonstrate any causal relationship. Consequently, we might doubt that embodiment is epiphenomenal, namely that modality-specific areas do not play a functional role in situation

model construction. Third, we might doubt that situation models are not necessarily embodied but depend on tasks. In response to these problems, we suggest some directions that future experiments might adopt.

##### ***4.1. Another Possible Strategy: from Q2 to Q1***

The strategy for the three paradigms discussed above is to test embodiment in discourse comprehension, answering Q2 on the solution of Q1. But the other way round might also be a good choice. We have got plenty of evidence of embodiment of isolated words and short sentences (see [21] for a review). What if we can demonstrate situatedness for those embodied isolated words and short sentences?

The theory of situated conceptualization backs up this idea. According to Barsalou, we are constructing a situation model even when comprehending a single concept. When facing the referent of a concept, we not only encode the particular referent as an exemplar of the concept but also the situational information associated with this exemplar, including objects, agents, setting, actions, self-relevance, and internal states. Later, when comprehending the concept again, the exemplar with the fittest situation will be activated as an instantiation of the concept [35], [36].

However, there is little concrete evidence for situated conceptualization yet. One of the difficulties is how to test the situation model for a single word. Individual differences matter a lot here: the situation that a participant generates is almost random. For example, when I comprehend the "kick", I might think of the situation that I'm kicking the football on the playground, but others might think of the situation that someone is kicking a dog on the street.

One possible solution is to have participants experience the same situation before the experiment. For example, we could arrange for an experimenter to shake hands with the participants in front of a red door before the study. Then we will test if reading "shake" could fasten the recognition of "red". If so, we could conclude that participants are constructing a situation model containing information about the color red. Combined with the evidence that "shake" also activates the motor cortex in the brain, it would be a piece of evidence for the embodiment of situation models.

##### ***4.2. Application of TMS***

fMRI results can at best show a correlation between modality-specific activation and situation model construction. To demonstrate that modality-specific areas play a functional role rather than an epiphenomenon in situation model construction, we shall interfere with the modality-specific areas in the experiment. One possible technique is Transcranial Magnetic Stimulation (TMS),

by which magnetic fields are applied to the scalp to affect the neural activity of specific brain regions [37]. For example, we apply TMS to influence participants' motor and somatosensory cortex, and then to see if there is a behavioral difference in their memory performance and recognition tests, which indicates changes in situation models.

### 4.3. Depth of Understanding

Why can't we find positive evidence for embodied situation models in all paradigms discussed above? Surely we could blame it on experimental designs or contingent factors. However, an alternative explanation is a combined theory of disembodied and embodied cognition, suggesting that concepts are abstract and symbolic at some level, but are instantiated in perceptual simulation in a situated way [38]. According to this theory, situation models might be abstract when we only make a shallow understanding of the narrative events. But when we comprehend the story deeply as if we get "immersed" into the fictional world, we would generate perceptual simulations. In other words, simulations can make understanding richer, but they're not necessary. To test this hypothesis, we could manipulate conditions where participants comprehend texts shallowly or deeply. fMRI data can reveal whether the activation of modality-specific areas is necessary or not, and a TMS study discussed above can provide insights into the role of modality-specific areas in the depth of understanding as well.

## 5. CONCLUSION

This essay gives a systematic review on the existence and the representational format of situation models. According to Zwaan's theory of event-indexing model, a situation model can be analyzed in five dimensions: space, time, entity, causation, and motivation. In all these five dimensions there is robust behavioral evidence demonstrating the existence of situation models. Nevertheless, the representational format of situational models, i.e. whether they are embodied or disembodied, is still controversial. Behavioral evidence is unable to answer this question, and only neural evidence testing both situation model construction and modality-specific activation can solve this problem. We generalize three promising paradigms from the literature (the coherence paradigm, the narrative shifts paradigm, and the inconsistency paradigm), but they only provide relatively weak evidence in favor of the embodiment of situation models. For future experiments, we suggest that experimenters could focus on situated conceptualization, application of TMS, and depth of understanding.

## REFERENCES

[1] T. A. Van Dijk and W. Kintsch, *Strategies of discourse comprehension*. New York: Academic

Press, 1983.

- [2] R. A. Zwaan, "Situation models: The mental leap into imagined worlds," *Current directions in psychological science*, vol. 8, no. 1, pp. 15–18, 1999.
- [3] R. A. Zwaan and G. A. Radvansky, "Situation models in language comprehension and memory.," *Psychological bulletin*, vol. 123, no. 2, p. 162, 1998.
- [4] R. A. Zwaan, "Situation models, mental simulations, and abstract concepts in discourse comprehension," *Psychonomic bulletin & review*, vol. 23, no. 4, pp. 1028–1034, 2016.
- [5] A. M. Glenberg, M. Meyer, and K. Lindem, "Mental models contribute to foregrounding during text comprehension," *Journal of memory and language*, vol. 26, no. 1, pp. 69–83, 1987.
- [6] D. G. Morrow and H. H. Clark, "Interpreting words in spatial descriptions," *Language and Cognitive Processes*, vol. 3, no. 4, pp. 275–291, 1988.
- [7] M. de Vega, "Backward updating of mental models during continuous reading of narratives.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 21, no. 2, p. 373, 1995.
- [8] R. A. Zwaan, "Processing narrative time shifts.," *Journal of Experimental Psychology: Learning, memory, and cognition*, vol. 22, no. 5, p. 1196, 1996.
- [9] M. Carreiras, N. Carriedo, M. A. Alonso, and A. Fernández, "The role of verb tense and verb aspect in the foregrounding of information during reading," *Memory & Cognition*, vol. 25, no. 4, pp. 438–446, 1997.
- [10] K. Ohtsuka and W. F. Brewer, "Discourse organization in the comprehension of temporal order in narrative texts," *Discourse Processes*, vol. 15, no. 3, pp. 317–336, 1992.
- [11] A. J. Sanford, S. C. Garrod, and S. Garrod, *Understanding written language: Explorations of comprehension beyond the sentence*. John Wiley & Sons, 1981.
- [12] T. P. Truitt and R. A. Zwaan, "Verb aspect affects the generation of instrument inferences," 1997.
- [13] M. Carreiras, A. Garnham, J. Oakhill, and K. Cain, "The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish," *The Quarterly Journal of Experimental Psychology Section A*, vol. 49, no. 3, pp. 639–663, 1996.
- [14] R. A. Stanfield and R. A. Zwaan, "The effect of implied orientation derived from verbal context on picture recognition," *Psychological science*, vol. 12, no. 2, pp. 153–156, 2001.

- [15] R. A. Zwaan, R. A. Stanfield, and R. H. Yaxley, "Language comprehenders mentally represent the shapes of objects," *Psychological science*, vol. 13, no. 2, pp. 168–171, 2002.
- [16] M. Singer, M. Halldorson, J. C. Lear, and P. Andrusiak, "Validation of causal bridging inferences in discourse understanding," *Journal of Memory and Language*, vol. 31, no. 4, pp. 507–524, 1992.
- [17] P. Van den Broek and R. F. Lorch Jr, "Network representations of causal relations in memory for narrative texts: Evidence from primed recognition," *Discourse processes*, vol. 16, no. 1–2, pp. 75–98, 1993.
- [18] D. E. Keefe and M. A. McDaniel, "The time course and durability of predictive inferences," *Journal of Memory and Language*, vol. 32, no. 4, pp. 446–463, 1993.
- [19] M. F. Lutz and G. A. Radvansky, "The fate of completed goal information in narrative comprehension," *Journal of Memory and Language*, vol. 36, no. 2, pp. 293–310, 1997.
- [20] T. Trabasso and S. Suh, "Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory," *Discourse processes*, vol. 16, no. 1–2, pp. 3–34, 1993.
- [21] C. A. Kurby and J. M. Zacks, "The activation of modality-specific representations during discourse processing," *Brain & Language*, vol. 126, pp. 338–349, 2013.
- [22] R. A. Zwaan, "Embodied cognition, perceptual symbols, and situation models," *Discourse Processes*, vol. 28, no. 1, pp. 81–88, 1999.
- [23] L. W. Barsalou, "Perceptual symbol systems," *Behavioral and brain sciences*, vol. 22, no. 4, pp. 577–660, 1999.
- [24] L. W. Barsalou, W. K. Simmons, A. K. Barbey, and C. D. Wilson, "Grounding conceptual knowledge in modality-specific systems," *Trends in cognitive sciences*, vol. 7, no. 2, pp. 84–91, 2003.
- [25] D. Casasanto and T. Gijssels, "What makes a metaphor an embodied metaphor?," *Linguistics Vanguard*, vol. 1, no. 1, pp. 327–337, 2015.
- [26] C. A. Kurby and J. M. Zacks, "Situation models in naturalistic comprehension," in *Cognitive Neuroscience of Natural Language Use*, 2015, pp. 59–76.
- [27] T. Yarkoni, N. K. Speer, and J. M. Zacks, "Neural substrates of narrative comprehension and memory," *Neuroimage*, vol. 41, no. 4, pp. 1408–1425, 2008.
- [28] N. K. Speer, J. M. Zacks, and J. R. Reynolds, "Human brain activity time-locked to narrative event boundaries," *Psychological Science*, vol. 18, no. 5, pp. 449–455, 2007.
- [29] C. Whitney, W. Huber, J. Klann, S. Weis, S. Krach, and T. Kircher, "Neural correlates of narrative shifts during auditory story comprehension," *Neuroimage*, vol. 47, no. 1, pp. 360–366, 2009.
- [30] J. M. Zacks *et al.*, "Human brain activity time-locked to perceptual event boundaries," *Nature neuroscience*, vol. 4, no. 6, pp. 651–655, 2001.
- [31] N. K. Speer, J. R. Reynolds, K. M. Swallow, and J. M. Zacks, "Reading Stories Activates Neural Representations of Visual and Motor Experiences," *Psychological Science*, vol. 20, no. 8, pp. 989–999, 2009.
- [32] M. Rinck, A. Hähnel, and G. Becker, "Using temporal information to construct, update, and retrieve situation models of narratives.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 27, no. 1, p. 67, 2001.
- [33] E. C. Ferstl, M. Rinck, and D. Y. von Cramon, "Emotional and temporal aspects of situation model processing during text comprehension: An event-related fMRI study," *Journal of cognitive Neuroscience*, vol. 17, no. 5, pp. 724–739, 2005.
- [34] E. C. Ferstl and D. Y. von Cramon, "Time, space and emotion: fMRI reveals content-specific activation during text comprehension," *Neuroscience letters*, vol. 427, no. 3, pp. 159–164, 2007.
- [35] L. W. Barsalou, "Situated conceptualization: Theory and applications," in *Foundations of embodied cognition: Perceptual and emotional embodiment*, New York, NY, US: Routledge/Taylor & Francis Group, 2016, pp. 11–37.
- [36] L. W. Barsalou, "Situated conceptualization: a framework for multimodal interaction (keynote)," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017, pp. 3–3.
- [37] C. A. Kurby, "Research methods: Neuroscientific methods to study discourse processes," in *The Routledge Handbook of Discourse Processes, Second Edition*, 2017, pp. 131–138.
- [38] B. Z. Mahon and A. Caramazza, "A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content," *Journal of physiology-Paris*, vol. 102, no. 1–3, pp. 59–70, 2008.