Advances in Social Science, Education and Humanities Research, volume 615
Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)

Situation Model and Embodiment

Yuji Yao[1*] Shizhe Wang [2] Fangqi Chen[3]

[1]*University of California, Santa Barbara, Santa Barbara, CA 93107*
[2] *Xiamen University Malaysia, Selangor, Sepang, 43900, Malaysia*
[3]*Wenzhou-Kean University, Wenzhou, China*
**Corresponding author. Email: Victoria.yao@outlook.com*

ABSTRACT
There are decades of evidence based on the fundamental language logic, showing that people create situation models when they understand text, which will be presented in more detail in the second part of this paper. However, after half a century of discussion among psychologists, it remains unclear whether these situation models are implemented in modality-specific simulations. To be able to sort out more clearly the clues whether the situation models are embodied or disembodied, in this work, we briefly review behavioural evidence for situated text understanding, and then evaluate the available evidence that bears on the question of whether situation models are "embodied" or "disembodied". We try to suggest further ways, such as using PET and fMRI technology, in which this question might be investigated effectively in the end of this paper.

Keywords*: Situation model, modality-specific simulations, embodied cognition, disembodied cognition*

1. INTRODUCTION

How do people understand language? This question causes lots of thinking and debates. Since the 20th century, with the discovery of the fundamental language logic, psychologists began to explore the relationship between this logic and people's understanding of the language:

In 1972, Bransford, Barclay, and Franks ran a study, trying to see whether participants' memory and understanding of a text was based on the "deep structural relations" of the sentences, but they failed to find related results [1]. As such experiments became more and more, researchers gradually realized that the logic and sentence structure were not the critical points for understanding the sentence. They discovered something else: In the same paper written by Bransford et al. in 1972, their results showed that instead of the "semantically interpreted" sentence structure, it was the whole situation described by the sentence that helped the participants to understand and memorize. As a result, a new theory came out, which was one of the main theories discussed in this paper: Situation Model theory.

Based on Zwaan and Magliano's report in 1995 [2], Situation Model theory described that when people understood words, they did not simply analyze the sentence structure or language logic; instead, people established a "situation" depicted by these words. Situation Model theory is considered one of the essential theories in learning language understanding. It helped explain that language sometimes helped people establish a "mental representation" when reading words [3]. In the first part of this paper, several studies will be presented to support the evidence of the Situation Model theory.

After recognizing that language relies on building situational models to be understood, psychologists have gone a step further to explore whether situational model are embodiment or disembodiment. The main task of psychologists who hold the embodiment view is to figure out whether the mental representations built in contextual models activate pattern-specific visual areas in the brain. However, some experiments have shown no clear evidence for corresponding activation of visual areas, so this relationship remains controversial. These experiments will be described more in the third part of this paper.

In this paper, we try to clarify the experimental process of the embodiment view, summarize the existing experimental results of psychologists, and try to propose a solution to the problem in the fourth part of the article.

2. SITUATION MODEL THEORY

2.1. Famous experiments and their results

Since the discovery of the Situation Model theory, some researchers had discovered the importance of constructive theory, and they began to run studies to see which helped people understand and memorize texts compared with the sentence structure.

In 1972, Bransford et al. ran a study trying to see whether sentence memory was based on the "semantically interpreted" structure of the sentences received or the overall situations created by the sentence. To run the study, they applied two types of sentences: PI (Potential Inference) and NI (Non-Inference). Potential Inference sentence meant that there was a hidden description in the sentence, while non-Inference sentence did not have [1]. They also used a third type of sentence where one of the pronouns was changed. Participants first heard several sentences describing a situation, and then they were asked to read another list of sentences. The task was to choose the one that was heard before, and participants also needed to rate their confidence about their choices [1]. Based on Bransford et al.'s hypothesis, if people were thinking about the internal structure of the sentence, they should notice the change of pronouns. The result was consistent with their hypothesis: people could not distinguish those sentences, especially in PI condition [1].

Therefore, Bransford et al. concluded that people used the constructive method (i.e., understanding the sentence by its semantic situation described) to understand and form memory [1]. This research laid the foundation for the emergence of the situation model. In 1983, first pointed out the notion of the Situation Model theory, where they described the theory as forming a mental representation when reading and understanding the sentence [4].

In 1987 and 1989, Morrow et al. [5,6] had shown that people were able to construct spatial situation models under experimental conditions, while in 1993, Zwaan and Oostendorp investigated whether people could construct detailed spatial situation models or not when they read in a general condition. They hypothesized that participants in the general condition would acquire spatial information more quickly and construct a weaker spatial situation model [7]. In this experiment, there were two phases: In the first phase, participants were asked to read as usual, and in the second phase, participants were asked to focus especially on spatial information [7].

Although overall the result said that people actually had a weak spatial representation of the situation in a normal reading condition, when subjects were asked to make a spatial inference, they were still able to form a strong mental representation of space [7].

Later in 1996, another experiment was done on the construction of the time and situation model. The purpose of this study is to explore whether time will affect the construction of the model when reading the text and whether the change of time will affect the speed of understanding the text [8]. In this experiment, the participants were asked to understand the sentences with a narrative time shift (e.g., one hour later) and without a narrative time shift [8]. The research results showed that in statements with time intervals, information was harder to extract, and participants spent more time understanding the statement [8]. Events were continuous in the model of everyday life, but some irrelevant events were often ignored in the text, which would affect the construction of the situational model. In other words, it also provided evidence for situation modeling from the side.

Overall speaking, researchers have found lots of evidence supporting the Situation Model theory, and by today, this theory has become an important foundation for other researchers to go deeper in the language understanding area.

2.2. Conclusion

As discussed above, people remember texts more easily under situational conditions than under non-situational conditions, which means that the theory of the situational model holds true.

the Situation Model says that people will form a mental representation while processing the text they read or hear, and there was quite a lot of evidence supporting this definition. Instead of only focusing on the Situation Model theory, some researchers tried to link this model to another theory, the embodiment theory, to see whether there was a relation between them. In the next part, this relation will be evaluated.

3. SITUATION MODEL AND EMBODIMENT

In this article, we focus on the relationship between the situation model and embodiment theory and evaluate the related research methods. In the next section, we will also try to suggest how to improve the research methodology in terms of its shortcomings.

The key to figuring out this relation is to see whether the mental representation created in the Situation Model activates the modality-specific visual area in the brain; if so, it will be reasonable to say a relation between them. Overall, researchers who supported this relation believed that if people were processing the mental image they created, they were using their vision to process and understand the mental representation. Thus the corresponding visual areas should be activated, which meant the mental image was embodied [9]. However, some experiments also showed no clear evidence for the

corresponding activation of the visual area, and thus this relation was still debatable.

3.1 Related Experiments and Their Results

Dating back to 1992, researchers already started their journey on constructing visual images and verbal descriptions. Denis and Cocude [10] ran a study to see whether the "final stage" of the representation built from a verbal source was very similar to that from participants' visual memory. They measured the reaction time and created a correlation diagram with time vs. distance, and they discovered that the mental images created from verbal descriptions were functionally alike with that from visual memory, which meant that the mental images from verbal descriptions were embodied even if there was no direct visual input [10]. However, although Denis and Cocude showed an almost similar trend of the time vs. distance correlation in verbally learning, there was still no direct evidence showing that the mental image was visually embodied.

Therefore, in 1997, D'Esposito et al. started a series of experiments [11], where one of their goals was to see the modality specificity of the visual cortex when creating the mental image. Based on this goal, they hypothesized that the "modality-specific visual cortex" would be activated when generating the mental representation of the object [11]. To better supervise the brain activity, they applied the fMRI technique with A "standard radiofrequency (RF)," which helped control the head motion to run the experiment. There were concrete words condition and abstract words condition in the study, representing "easy to image" words and "hard to image" words, respectively. Participants would be asked to image the concrete words they hear and "passively" listen to the abstract words [11]. The result was consistent with their hypothesis: The individual subject analysis showed that several brain areas were more activated in the concrete words condition, especially the left inferior temporal lobe, which showed the most steady and robust activation during the experiment.

Through the fMRI technique, D'Esposito et al. provided evidence showing that the mental representation of words was embodied when participants were processing the words in their minds.

However, as mentioned earlier, there was still research that failed to find the activation of the primary visual cortex in their results.

For example, one study [12] pointed out that in their “route finding” test, which is participants imagine that they are walking, alternatively take roads to the left and right, and then, when they are asked where they are using positron emission tomography (PET) scan them. In this experiment, they failed to find the “absolute rCBF”, which is Regional cerebral blood flow, has any changes. They also did other experiments, such as the “alphabet-scrutinizing task.” In this task, the participants close their eyes and listen to the sound of a letter. Then, in a similar condition, the subjects imagine the letter as large as it fills the whole visual field or as small as it fills the central field of view only. And then, they are asked to answer a question about the characteristic about the letter. Still, other than an area in the “lingual gyrus close to the parieto-occipital sulcus,” there was no other activation discovered in the primary visual cortex [12].

As a result, Roland et al. believed that the early visual cortex was not engaged in the visual imagery [12]; in other words, those images created were not visually embodied. Later, Roland and Gulyas set up three conditions for the participants: (1) visually learning the figures presented on a screen, (2) the mental image created based on the figures, (3) the visual recognition of the figure among other pictures [13]. They measured the rCBF in the resting phase and the experiment phase, but the primary visual cortex was not activated in the mental image condition; instead, the other two groups successfully highlighted the primary visual area [14]. However, as Roland and Gulyas mentioned in their paper, the results might be due to the methods and stimulus, and it might change when using other techniques [12].

In 2000, Mellet et al. used another technique to help illustrate this relation: the positron emission tomography (PET) technique [9]. They hypothesized that high-resolution visual mental images would activate the early visual cortex, wherever it came from verbal learning or visual learning. They recruited participants to experiment, and they applied a particular stimulus in their study: scenes that contained different objects arranged on a base, each differed only in the order of the objects placed on the base [9]. The participants would be arranged into either "visually-learning group" or "verbally-learning group": People in the visually-learning condition needed to learn and memorize the scenes before the test truly began (learning phase), and people in the verbally-learning condition needed to hear the words, form a mental representation and memorize the arrangements of segments [9]. After learning, participants were asked to imagine one of the scenes; and then, they would hear "two positions on the base" and the relative height of the objects in those positions. The accuracy and reaction time were recorded, and the PET and rCBF were recorded to vividly record brain activities [9].

The pattern in the two learning conditions was very similar to each other: there was no difference between the two conditions (verbal and visual), as well as the reaction time. Their follow-up study [15] continued to use the PET method. They hypothesized that acquiring topographic knowledge through verbal information can differ in some ways from acquiring topographic knowledge through visual experience. In their experiment, they scanned the participants' brains when they received verbal information and visual information about the same map

[15]. The results suggested that although both tasks involve visuospatial internal representation, the brain still learns in a modality-specific way [15].

4. CONCLUSION

The purpose of this paper is to compile the research results of the situation model, list the evidence of the situation model and explore its relationship with embodiment theory in order to facilitate further research or to assess whether the situation model is more embodiment or disembodiment.

In conclusion, it is noticeable that there were lots of articles that supported the link between mental representation and visual cortex, even using PET and fMRI technology, which strongly support for the embodiment theory. Although some researchers were opposed to this topic. In their experiments, the establishment of mental representations did not significantly correlate with the activation of the visual cortex. However, their results might be due to the stimulus used because other brain areas were activated, such as the "superior occipital and posterior parietal cortices" in Roland and Gulyas's paper [12]. Therefore, we conclude from the above discussion that mental imagery is visually embodied, and so when people are processing the situation model, the mental representation may be embodied as well.

In this paper, we mainly talk about the Situation Model theory and the Embodiment theory. We also evaluate the relationship between the two theories: after research, we think there is a link between these two theories: People are visually embodied while creating the mental representation from the text. We also notice some limitations in the experiments we read during the research, which we believe can be the future research direction.

First, we discover that very few experiments covered participants' eyes during the investigation. The activation of the visual cortex might be due to the visual input participants receive from the environment. As a result, future researchers could set up a control group or cover participants' eyes during the mental imaging group so that no other visual input is processed. If the mental representation of the situation is visually embodied, then covering eyes should not make a big difference from the results in the previous articles.

Second, in Mellet et al.'s paper in 2000 [9], they mentioned an interesting point in the introduction part: they noticed that when the mental image was vivid and clear, the early visual cortex was more activated. We believe this can be another future direction to further illustrate this relation, where researchers can set up a "blurred description" group for the participants and a "clear and vivid description" group to see whether the mental image's intactness will affect the participants' status of the visual cortex. We hope that by presenting the limitations of the experiments mentioned above in this paper, we can help to eliminate more confounding factors and obtain more effective experimental results when conducting experiments on the situation model and embodiment theory in the future. We further conclude that the situation model is modality specific in the process of building or not.

Overall, we anticipate future research to run more experiments using other techniques so that more data can be illustrated.

REFERENCES

[1] Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive psychology, 3(2), 193-209.

[2] Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. Journal of experimental psychology: Learning, memory, and cognition, 21(2), 386.

[3] Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological bulletin, 123(2), 162.

[4] Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension.

[5] Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and language, 26(2), 165-187.

[6] Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. Journal of memory and language, 28(3), 292-312.

[7] Zwaan, R. A., & Van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension?. Discourse processes, 16(1-2), 125-143.

[8] Zwaan, R. A. (1996). Processing narrative time shifts. Journal of Experimental Psychology: Learning, memory, and cognition, 22(5), 1196.

[9] Mellet, E., Tzourio-Mazoyer, N., Bricogne, S., Mazoyer, B., Kosslyn, S. M., & Denis, M. (2000). Functional anatomy of high-resolution visual mental imagery. Journal of Cognitive Neuroscience, 12(1), 98-109.

[10] Denis, M., & Cocude, M. (1992). Structural properties of visual images constructed from poorly or well-structured verbal descriptions. Memory & Cognition, 20(5), 497-506.

[11] D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., & Farah, M. J. (1997). A functional MRI study of mental image generation. Neuropsychologia, 35(5), 725-730.

[12] Roland, P. E., & Gulyas, B. (1994). Visual imagery and visual representation. Trends in neurosciences, 17(7), 281-287.

[13] Mellet, E., Petit, L., Mazoyer, B., Denis, M., & Tzourio, N. (1998). Reopening the mental imagery debate: lessons from functional anatomy. Neuroimage, 8(2), 129-139.

[14] Roland, P. E., & Gulyás, B. (1995). Visual memory, visual imagery, and visual recognition of large field patterns by the human brain: functional anatomy by positron emission tomography. Cerebral Cortex, 5(1), 79-93.

[15] Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. Cerebral Cortex, 12(12), 1322-1330.