

# The Effect of Multiple Perspectives Information on the Characteristics of Human's Spatial Cognition in the Human-Human Interaction of Spatial Cognition Tasks

Xianliang Mu, Lifen Tan, Yu Tian, and Chunhui Wang<sup>(✉)</sup>

National Key Laboratory of Human Factors Engineering,  
China Astronaut Research and Training Center, Beijing 100094, China  
1078042576@qq.com, {wluya.o9802, chunhui\_89}@163.com,  
cctian@126.com

**Abstract.** Spatial cognition is an important branch of cognition science. The behavior in the spatial cognition interactions between people was studied in our research. And the results will contribute to the design of robot's cognitive system that similar to those of human which will make the robot much more intelligent to aids in human-robot cooperative work. We have designed several experimental scenes with different degree of ambiguity. In different scenes, the assistants will communicate with a commander through language and gesture on a single perspective or double perspectives. Based on these experiment, we can master the feature of human's spatial language expression and the rule of choosing spatial reference frame in the human-human interaction of spatial cognition tasks. The following three results are summed up. (1) The assistants are inclined to use the "egocentric" and "exocentric" frames of reference. (2) The frequency of choosing reference frames increases with the improvement of scene ambiguity. (3) Compared with single perspective, the assistants need much less number of reference frames on the double perspectives.

**Keywords:** Spatial cognition · Cooperative work · Spatial reference frame · Perspective-taking

## 1 Introduction

Spatial cognition refers to the processing of spatial information that perceived by people, such as the three-dimensional objects' size, shape, location, distance and the relationship between them in the psychological space or physical space. The spatial cognition tasks and human-human interaction are very common in our daily life. When people working together in spatial cognition tasks, the commander may communicate with the assistant or the assistant can observe the world on the commander's perspective to finish the work well. Actually, the assistant may be human or robot.

In the traditional human-robot interaction, human makes decisions through the analysis of information from his own perspective and the robot's perspective, and then he will interact with the robot through mouse, keyboard, handles, pedals, etc. The robot

will fulfill specific task after it has received the commands. Under this interactive mode, the robot with a low intelligence level cannot understand the human's intention well. And the human's capacity of perceiving, attending the scene information has not been fully employed. People have a really high workload when they interact with robots with low intelligence level.

In the future, we hope that robots can communicate with people and have the capability of perspective-taking to help people fetch tools or assemble machines.

## 2 Method

In the human-human interaction, the key issue is how to eliminate ambiguity and it's the same with human-robot interaction. So the assistant should get the commander's intention as soon as possible to fulfill a task efficiently. In our study, we focus on the feature of human's spatial language expression and the rule of choosing spatial reference frame in the human-human interaction of spatial cognition tasks. According to literature [1–6], there are five frames of reference: exocentric (world-based, such as "Go north"), egocentric (self-based, "Turn to my left"), addressee-centered (other-based, "Turn to your left"), deictic ("Go here [points]"), and object-centric (object-based, "The fork is to the left of the plate").

We have designed several experimental scenes with different degree of ambiguity. The assistants will observe the scene on two different perspectives. One is the perspective of themselves and the other is combining the perspectives of the commander and the assistant. In different scenes, the assistants will communicate with a commander through language and gesture on a single perspective or double perspectives. The participants' utterances were collected. Then we extracted the categories of reference frames that participants used and recorded how many times they had used.

Based on the experiment, we can master the feature of human's spatial language expression and the rule of choosing spatial reference frame in the human-human interaction of spatial cognition tasks.

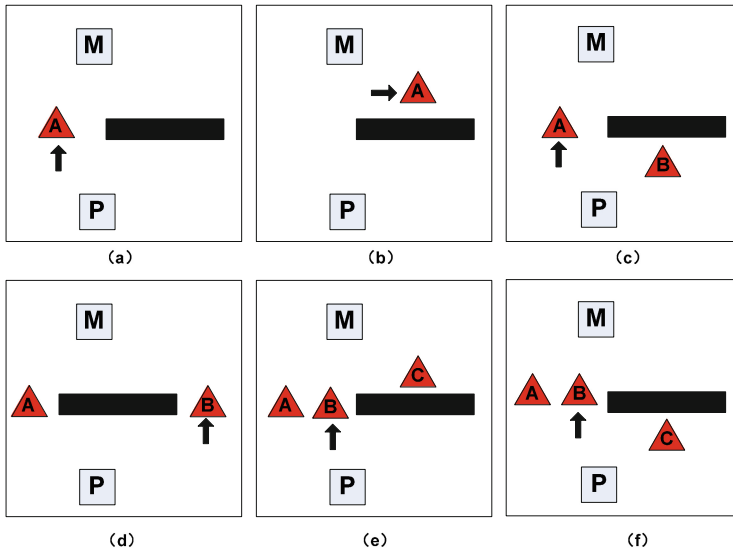
## 3 Experiment

### 3.1 Subjects

We recruit 12 participants who are between 20 and 40 years old serving as assistants, and they are not suffering from disease or color blindness. Besides, they should have a normal ability of expression.

### 3.2 Experiment Scene

As shown in Fig. 1, we design 6 scenes which have different levels of ambiguity. "P" stands for assistant, "M" stands for commander, "A", "B", "C" stand for goals, and the "arrow" points at the goal that the commander wants to get. The black rectangle stands for an obstacle. In the first three scenes, the assistant can identify the target "A" through



**Fig. 1.** Experimental scene: (a) target “A” can be seen by M and P, (b) M can see the target “A” but P can’t see the it, (c) M can see the target but P can see “A” and “B”, so P should exclude “B” to choose “A” automatically, (d) identify target “B” through communication, (e) M can see “A”, “B” and “C”, while P can see “A” and “B”, so the target “B” should be identified through communication, (f) M can see “A” and “B”, while P can see “A”, “B” and “C”, so the target “B” should be identified through communication.

perspective-taking easily. In the other three scenes, each one has a different degree of ambiguity. Once the commander sent out a verbal order, the assistant should find the goal as soon as possible through communicating with commander and perspective-taking. In different scenes, the participants’ utterances were collected. And we extracted the categories of reference frames that participants used and recorded how many times they had used, as shown in Fig. 3.

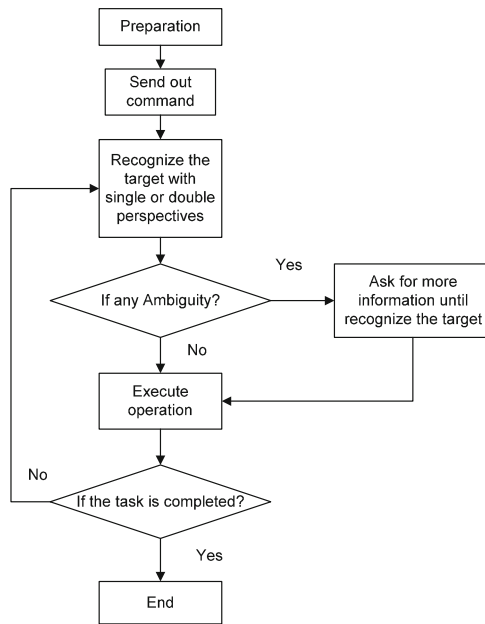
### 3.3 Experiment Progress

The 12 assistants will accomplish the experiment in turn, and all the 6 experiment scenes should be done as shown in Fig. 1. What’s more, in order to avoid the influence of order, we arrange the assistants’ order through a method called “balanced Latin square design”, as shown in Table 1.

The assistants named P7 ~ P12 will follow the same order designed above. In addition, the former 6 assistants will do the spatial cognition experiment on a single perspective. Then, it will take some time for the assistants to do a questionnaire test. We will make use of the data to analyze the subjective factors that affect experimental results. Later, these 6 assistants will do the spatial cognition experiment on two perspectives, namely, the assistant’s own perspective and the commander’s perspective. Also, the other questionnaire test should be accomplished by the assistants. The later 6

**Table 1.** The assistants' experimental order in different scenes

| Assistants | Experimental order |         |         |         |         |         |
|------------|--------------------|---------|---------|---------|---------|---------|
|            | Scene a            | Scene b | Scene c | Scene d | Scene e | Scene f |
| P1         | 1                  | 2       | 6       | 3       | 5       | 4       |
| P2         | 2                  | 3       | 1       | 4       | 6       | 5       |
| P3         | 3                  | 4       | 2       | 5       | 1       | 6       |
| P4         | 4                  | 5       | 3       | 6       | 2       | 1       |
| P5         | 5                  | 6       | 4       | 1       | 3       | 2       |
| P6         | 6                  | 1       | 5       | 2       | 4       | 3       |



**Fig. 2.** A flow chart of the experiment

assistants will do the double perspectives test first and then accomplish the single perspective test.

The detailed procedure of single or double perspectives is shown below.

- a. Prepare to conduct experiment.
- b. The commander send out the order, "Please fetch the red screw driver to me."
- c. The assistant starts to search the red screw driver.
- d. The assistant can see the experimental scene through his own perspective or double perspectives. (The assistant can get the commander's perspective by a PC displayer.)
- e. If there are any ambiguities, the assistant will ask for more information.
- f. The assistant will identify and deliver a target.

- g. The commander will tell the assistant if he has got a correct target. And if the choice is wrong, the assistant have to repeat step c ~ g. If the choice is correct, the assistant will follow the next step.
- h. The experiment end.
- i. The assistant will accomplish the questionnaire text.
- j. The man who is responsible for data collection should record the data as soon as possible.

The Fig. 2 gives a description in a flow chart form.

## 4 Results

### 4.1 Analysis Between Assistants

We processed the data and got the correlation matrix of different reference frames. As shown in Table 2, a significant negative correlation was found between egocentric and exocentric. So people who chose the egocentric would seldom choose the exocentric and vice versa. The reason is that P3 and P4 are northerners who prefer exocentric and the others are southerners who prefer egocentric. What’s more, the data showed that assistants preferred to use these two frames of reference. So we added the egocentric to exocentric and got a new figure, Fig. 3. It is much more significant to see this preference. Besides, all of the assistants chose the object-centric quite often, whether northerners or southerners.

**Table 2.** Correlation matrix

|                    | exocentric | egocentric    | addressee-centered | object-centric | deictic     |
|--------------------|------------|---------------|--------------------|----------------|-------------|
| exocentric         | 1          | <b>-0.946</b> | -0.138             | -0.198         | -0.157      |
| egocentric         | -          | 1             | -0.058             | 0.011          | -0.053      |
| addressee-centered | -          | -             | 1                  | 0.34           | <b>0.98</b> |
| object-centric     | -          | -             | -                  | 1              | 0.375       |
| deictic            | -          | -             | -                  | -              | 1           |

As we know, different people may have different capability of spatial cognition. So we calculated the total amount of each assistant’s choice of spatial reference frames, as shown in Fig. 5. The statistic of the result is shown in Table 3, in which the low variance indicates that the assistants have similar capabilities of spatial cognition (Fig. 4).

### 4.2 Analysis of Different Scenes

In Fig. 6, the frequency of choosing reference frames increases with the improvement of scene ambiguity. However, the ambiguity of the scene is hard to be explained quantitatively. We just focus on the trend in the figure. What’s more, assistants preferred to use egocentric and exocentric in each scene.

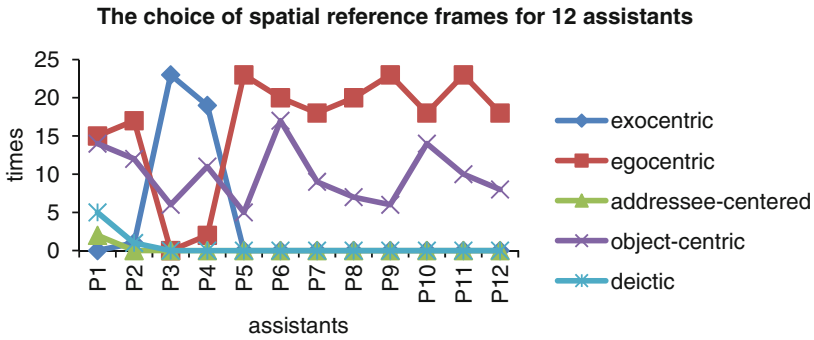


Fig. 3. The choice of spatial reference frames for 12 assistants throughout the experiment

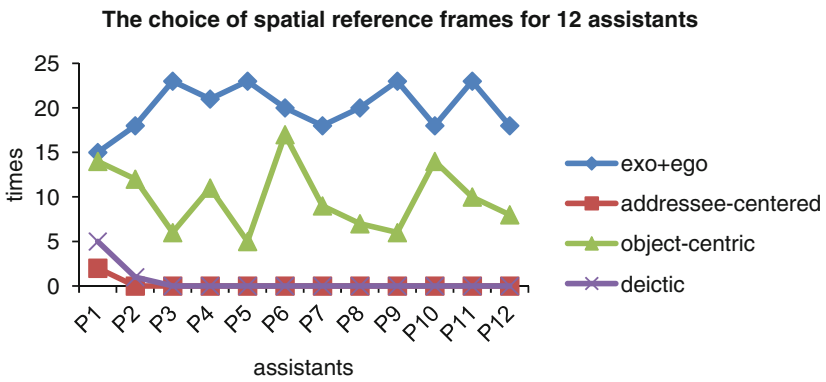


Fig. 4. Add exocentric to egocentric

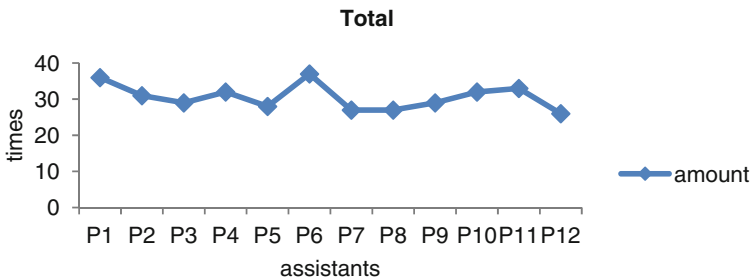
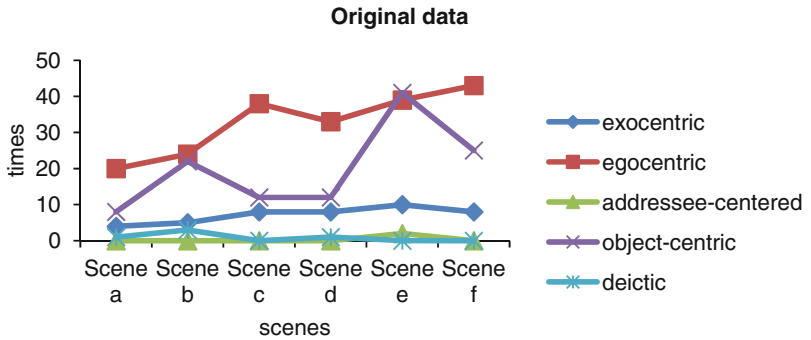


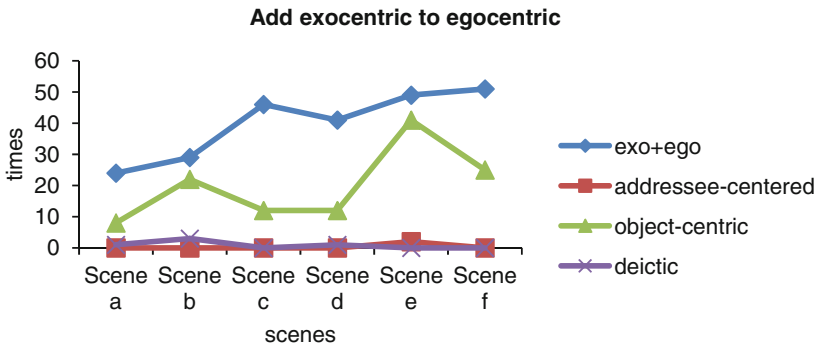
Fig. 5. The total amount of each assistant's choice of spatial reference frames

**Table 3.** The statistic of the total amount of each assistant’s choice of spatial reference frames

| Mean  | Standard Error of Mean | Standard deviation | Variance | Skewness | Standard Error of skewness |
|-------|------------------------|--------------------|----------|----------|----------------------------|
| 30.58 | 1.026                  | 3.554              | 12.629   | 0.558    | 0.637                      |



(a)

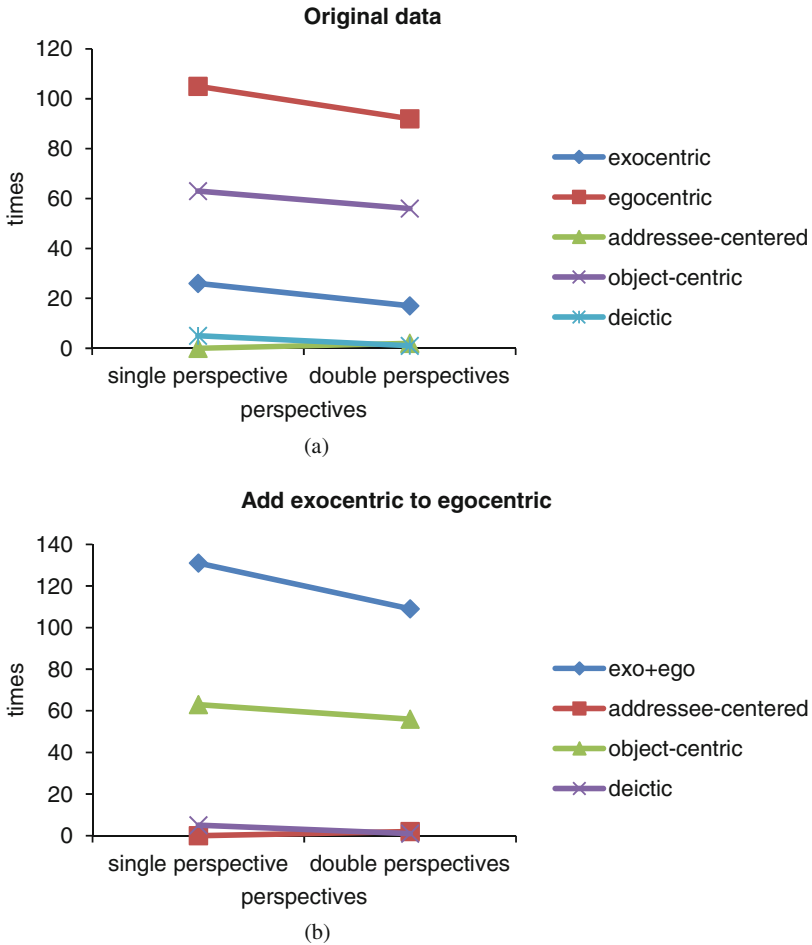


(b)

**Fig. 6.** The choice of spatial reference frames for 12 assistants in different scenes: (a) original data, (b) add exocentric to egocentric

### 4.3 Analysis of Single and Double Perspectives

In Fig. 7, compared with single perspective, the assistants need much less number of reference frames on the double perspectives. It means that abundant visual image information will improve the performance of human-human interaction. What’s more, assistants preferred to use egocentric and exocentric in each case.



**Fig. 7.** The choice of spatial reference frames for 12 assistants on single perspective and double perspectives: (a) original data, (b) add exocentric to egocentric

## 5 Discussion

There are various characteristics of human’s spatial cognition. In our study, we just focus on human’s spatial language expression and the rule of choosing spatial reference frame. Besides, the effect of single and double perspectives on human’s choice of spatial reference frame was studied here, and the effect of the third perspective needs further research.



## 6 Conclusion

We can get some conclusions from this human-human interaction experiment in the spatial cognition task.

- a. When we design the robot's cognitive system that similar to those of human, we need to take the difference between southerners and northerners into consideration. The northerners prefer exocentric while the southerners prefer egocentric. In addition, these two kinds of spatial reference frames and the object-centric reference frames are frequently-used in the human-human interaction of spatial cognition tasks. So these three reference frames should be the key point of cognitive modeling task. Besides, the deictic reference frame has a strong relevance with gesture, such as "Go here [points]". Assistant can easily understand the commander's command through combining the gesture and language information. If you want to improve the naturalness of the human-robot interaction system, the deictic reference frame should be considered.
- b. In the scene c, d, e, f, the ambiguities of scenes are much higher. There will be much more interactive behavior between people. So we should figure out the human's spatial language expression pattern and the rule of choosing spatial reference frame in these scenes. And these results will contribute to cognitive modeling work.
- c. Multiple perspectives information will improve the performance of human-human interaction. Particularly, the commander will take much less time to communicate with the assistant. A second perspective can improve the efficiency of human-human cooperative work.

**Acknowledgments.** This study was supported by the Experimental Technique Youth Fund Project (No. SYFD1400618) and the Foundation of National Key Laboratory of Human Factors Engineering (No. SYFD140051802). Author Xianliang Mu, Lifen Tan, Yu Tian, Chunhui Wang were supported by the foundation of National Key Laboratory of Human Factors Engineering (No. HF2013-Z-B-02, NO.HF2011Z-Z-B-02). In addition, the authors would like to thank anonymous reviewers for their comments about the research and suggestions for improving this paper.

## References

1. Levelt, W.J.M.: Some perceptual limitations on talking about space. *Limits in perception: essays in honour of Maarten A. Bounman*, pp. 323–358 (1984)
2. Carson-Radvansky, L.A., Radvansky, G.A.: The influence of functional relations on spatial term selection. *Psychol. Sci.* **7**, 56–60 (1996)
3. Carson-Radvansky, L.A., Logan, G.D.: The influence of functional relations on spatial template construction. *J. Mem. Lang.* **37**, 411–437 (1997)

4. Goldin-Meadow, S.: When gestures and words speak differently. *Curr. Dir. Psychol. Sci.* **6**, 138–143 (1997)
5. McNeill, D.: *Hand and Mind: What Gestures Reveal About Thought*. University of Chicago Press, Chicago (1992)
6. Mintz, F.E., Trafton, J.G., Marsh, E., Perzanowski, D.: Choosing frames of reference: perspective-taking in a 2-D and 3-D navigational task. *Hum. Factors Ergon. Soc. Meet.* **48**, 1933–1937 (2004)