Available online at www.sciencedirect.com
SCIENCE DIRECT
COGNITION
Cognition 94 (2004) 185–192
www.elsevier.com/locate/COGNIT

Action, verbal response and spatial reasoning

Ranxiao Frances Wang*

Department of Psychology and Beckman Institute, University of Illinois, 603 E. Daniel St, Room 533, Champaign, IL 61820, USA

Received 20 November 2002; received in revised form 6 October 2003; accepted 12 May 2004

Abstract

Studies have shown that perception of distance, orientation and size can be dissociated from action tasks. The action system seems to possess more veridical, unbiased information than the perceptual/verbal system. The current study examines the nature of the distinction between action and verbal responses in a spatial reasoning task. Participants imagined themselves facing different orientations and either pointed to where other objects would be, or verbally reported their egocentric directions (e.g., "50 degrees to my left"). When using pointing responses, RT and error increased as a function of the angular disparity between the imagined heading and their actual heading. However, when using verbal responses, performance was not affected by angular disparity, suggesting that participants knew the direction of the targets from the imagined perspective but could not point to them directly. The verbal and action systems have fundamentally different information or processes rather than quantitatively different ones.

Keywords: Perception and action; Perspective change; Spatial reasoning

1. Action, verbal response and perspective change problem

Various studies have suggested a distinction between representations underlying perception and action, showing that a more accurate and veridical representation is associated with actions but not available to the perceptual/knowledge system (Bridgeman, Kirch, & Sperling, 1981; Bridgeman, Peery, & Anand, 1997; Creem & Proffitt, 1998; Goodale & Milner, 1992; De Vega & Rodrigo, 2001; Loomis, Da Silva, Philbeck, & Fukusima, 1996; Proffitt, Bhalla, Gossweiler, & Midgett, 1995; etc.). For example, Bridgeman et al. (1981, 1997) showed participants a target dot on the screen surrounded by

* Tel.: +1-217-244-3664; fax: +1-217-244-5876.
E-mail address: francesw@s.psych.uiuc.edu (R. Frances Wang).

0022-2860/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2004.05.001

a rectangular frame. When the frame was displaced, the dot was perceived to jump in the opposite direction. However, when asked to point to the dot with their unseen hand, participants pointed correctly, unaffected by the illusion. Thus, these studies suggested that the action system relies on veridical information of the world that is not available to the perceptual/knowledge system.[1]

A similar distinction is found on perception of geographical slant. Proffitt and colleagues (Creem & Proffitt, 1998; Proffitt et al., 1995) asked participants to stand at the foot of a hill and judge the slope. They either made verbal estimates in terms of degrees, or perceptual judgments using a display of the cross-section of the hill, or action judgments by adjusting a palm board with their unseen hand. The perceptual and verbal judgments significantly overestimated the slope, but the action judgments remained accurate. Again, the action system relies on veridical information while the perceptual/verbal system does not.

Action is also shown to be more accurate than perception/verbal reports in distance estimation. Loomis et al. (1996) asked participants to adjust two sets of markers on the ground in front of them, one pair laid out along the front/back axis and the other pair paralleled the left/right axis, until the distance between markers of one pair equalled that of the other pair. Participants adjusted the front/back interval significantly larger in order to match the left/right interval perceptually, suggesting that the visual space for perception is distorted. In contrast, when asked to walk the interval between the two front/back markers, the distance they walked did not show the same systematic bias, suggesting the positions of the two markers are represented correctly in the action system.

The two systems interact under certain conditions. Proffitt et al. (1995) asked participants to judge the slant either at the foot of the hill, or away from the hill in a lab room. When the judgments were made according to memory, the action responses showed similar biases as the verbal responses. Bridgeman et al. (1997) also showed that when the motor response was delayed, pointing responses were subject to the Roelofs effect as the cognitive responses. These studies suggested that the veridical information in the action system is relatively short-lived.

The current study examines the nature of the representations underlying the action and verbal systems using a spatial reasoning task. Participants imagined facing a different orientation and reported the egocentric direction of other targets from that imagined perspective. It has been shown that performance in a spatial reasoning task is affected by the angular disparity between the imagined heading and the participants' actual heading. For example, Rieser (1989; also see Easton & Sholl, 1995; Farrell & Robertson, 1998; May, 1996; Presson & Montello, 1994; Shelton & McNamara, 1997; Wang, 2003; Wraga, Creem, & Proffitt, 2000) asked participants to imagine themselves turn to face a different orientation and "point to object **X** as if you were facing object **Y**". Performance decreased as the angular disparity increased, both in terms of RT and pointing errors, suggesting that information about the egocentric direction of the target direction

[1] Several distinctions have been proposed between action and perceptual/knowledge tasks, e.g., dorsal vs. ventral (Goodale & Milner, 1992), sensorimotor vs. cognitive (Bridgeman et al., 1997), egocentric vs. allocentric (Creem & Proffitt, 1998), unconscious vs. conscious (Creem & Proffitt, 1998).

from the imagined perspective is not directly available to the participants for making a pointing response.

It remains unclear, however, whether information about the imagined perspective is available to the verbal system. Research has shown that linguistic systems tend to be more flexible in encoding and using spatial information, selecting different reference frames when needed (Carlson & Logan, 2001; Carlson-Radvansky, & Irwin, 1993; Levinson, 1996). Thus, if the angular disparity effect in a spatial reasoning task is specific to the action system, then one might expect a different pattern using verbal responses, even though both tasks required the participants to report the egocentric directions.

2. Methods

2.1. Participants

Thirty undergraduate students from an introduction to psychology class at the University of Illinois participated for course credit.

2.2. Apparatus

The experiment was conducted in a rectangular room with five target objects (see Fig. 1). A swivel chair was placed in the middle of the room fixed to the floor. A video camera was mounted directly above the chair, which sent an overhead image of the room to a VCR. The pointing responses were recorded and coded after the experiment was completed. A Gateway E4200 PC computer was used to randomize the sequence of the trials for each participant.

2.3. Design and procedure

Half of the participants were tested in the verbal task, and half were in the pointing task. For participants in the verbal task, the experimenter stood at 45°, 90° and 180° relative to

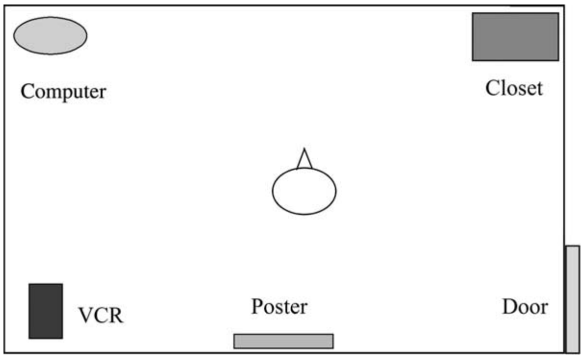

Fig. 1. An overhead view of the testing room.

the participants to show them these angles before they entered the testing room. The participants were instructed to report target directions in terms of the egocentric angles (e.g., straight ahead is "zero degree", straight behind is "180 degrees", straight to the left is "90 degrees left", and so on). Participants were asked to report the angles as accurately and quickly as possible, giving the exact angles to the precision of 5° if possible. For participants in the pointing task, they were asked to point to the direction of the targets with whichever hand was most convenient.

The participants first learned the target locations by sitting in the swivel chair and looking around as the experimenter named the five targets. They were allowed as long as they needed to remember the target locations. Then they were blindfolded and sound masked, and turned to face one of the targets. Then they completed 20 trials in a random sequence. In each trial, the participant was asked to "imagine facing **X** where's **Y**?" Each participant imagined facing each of the five targets and reported the egocentric direction of each of the remaining ones, either by pointing or by verbal description, yielding 20 trials in total (5 imagination headings $\times$ 4 targets each). Thus, the angular disparity between their actual heading and the imagined heading was systematically varied (0, 60, 120, and 180 degrees). Different participants physically faced different objects. For both verbal and pointing tasks, the same number of participants faced each object (i.e., 3 participants physically faced the Poster during testing, 3 faced the VCR, and so on). Thus, the egocentric responses were counterbalanced across participants for different perspectives in both tasks.

2.4. Coding and data analysis

Responses were coded from the videotape after the experiment was completed. Response latency was measured from the end of the target name to the end of the response, which was indicated as the participants' hand stabilized. Because not all participants had angular disparity of 180° (those who faced the Poster did not have an imagination heading at 180°), the ANOVA tests were based on 3 imagination angles only (0°, 60°, and 120°).

3. Results

Participants showed different patterns of performance in the two tasks. In the pointing task, participants showed a significant angular disparity effect, both in RT and error ($Fs(2, 32) > 8.4$, $Ps < 0.01$). In contrast, there was no significant effect of angular disparity in the verbal task ($Fs(2, 28) < 1.1$, $Ps > 0.34$), although participants made their verbal estimations in specific angles, using values such as 10°, 50°, 155°, and so on, as in the pointing task. Thus, although pointing to an object from an imagined perspective took extra time comparing to from the observer's actual perspective, verbal responses were equally fast and accurate regardless of the perspective, despite the fact that the participants were reporting the same information (i.e., the target's egocentric direction).

The overall RT in the verbal task was longer than in the pointing task, however. This is not surprising because the verbal response itself takes more time than the motor response, and converting a known egocentric direction into an angle in degrees also takes extra time. However, this difference raises a potential alternative explanation for the lack of angular disparity effect in the verbal task. It is possible that the overall slowing absorbed some of the processing time. To test this possibility, participants in both tasks were divided into a fast group and a slow group based on a median split on their overall RT.[2] In the verbal task, the two groups differed only in their overall RT ($F(1, 13) = 28.4$, $P < 0.001$), and neither group showed an angular disparity effect ($F\text{s}(2, 26) < 1.0$, $P\text{s} > 0.37$), nor was there an interaction between angular disparity and group ($F\text{s}(2, 26) < 0.48$, $P\text{s} > 0.63$). In the pointing task, the two groups differed in overall RT ($F(1, 15) = 16.6$, $P = 0.001$), and there was an angular disparity effect ($F\text{s}(2, 30) > 8.2$, $P\text{s} < 0.001$) and an interaction between angular disparity and group (significant for RT: $F(2,30) = 19.6$, $P < 0.001$; Non-significant for error: $F(2,30) = 0.3$, $P = 0.74$), suggesting that the faster responders tended to show less angular disparity effect than the slower ones, possibly due to their ceiling performance.

More importantly, a direct comparison was conducted between the slow group in the pointing task and the fast group in the verbal task (see Fig. 2). The overall RT was matched between the two groups ($F(1,14) = 1.8$, $P = 0.20$). However, there was still a significant interaction between angular disparity and task ($F\text{s}(2, 28) > 4.6$, $P\text{s} < 0.02$). More specifically, there was a significant effect of angular disparity in the pointing task ($F\text{s}(2, 14) > 7.4$, $P\text{s} < 0.01$), but not in the verbal task ($F\text{s}(2, 14) < 2.4$, $P\text{s} > 0.13$). Thus, even when the two groups were matched in overall performance, there was still a significant difference in the angular disparity effect in the two types of tasks: a strong angular disparity effect in the pointing responses but not in the verbal responses.

4. Discussion

Previous studies on the dissociation between perception/communication and action suggested that the action system often has access to more accurate information and outperforms the perceptual/ verbal system. However, the current study showed an opposite pattern in a spatial reasoning task: although verbal reports of egocentric target directions from an imagined perspective were as fast and accurate as those of the actual perspective, the pointing responses showed a large cost and did not seem to have access to information that was clearly there. That is, even though both tasks required participants to report the same information, namely the direction of a target relative to their body, participants were able to describe the direction of a target from the imagined perspective but cannot make a pointing response readily.

These results suggest that the verbal system has flexible access to information that the action system does not. The lack of angular disparity effect using verbal reports suggests that information about egocentric target directions from imagined perspectives is as readily available as that of the actual perspective. However, this information cannot be used to directly guide pointing responses. Thus, additional processing is needed, leading to performance costs in both RT and accuracy. In contrast, these representations seem to be equally accessible to the verbal system and therefore verbal responses required no

[2] Although there might be differences between the groups such as processing speed, motivation, overall spatial ability, and so on, the analysis allows the match of the overall performance across tasks.

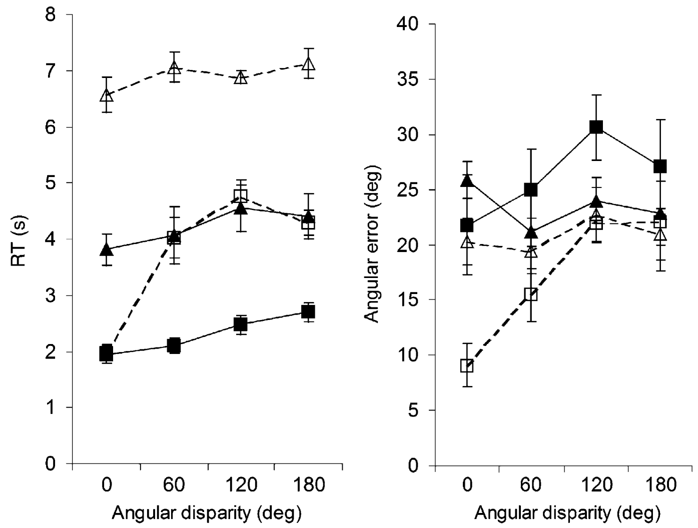

Fig. 2. A direct comparison between the verbal and pointing tasks by dividing participants into fast and slow groups according to their overall RT. The left panel shows the RT and the right panel shows the angular errors. The filled symbols show the fast groups and the open symbols show the slow groups. The square symbols show the pointing task, and the triangle symbols show the verbal task. The critical comparison is between the slow group in the pointing task and the fast group in the verbal task. The error bars are between-subject standard errors.

additional processing for the imagined perspective compared to the participants' actual perspective, leading to equal performance for all perspectives.

The difference between the two tasks can potentially be a result of different types of representations, or different types of processes. For example, one possible strategy for the verbal task is to verbally encode the direction of the five targets, and then use numerical subtraction to solve the perspective change task. For example, if the participants encode that the Computer is at 65°, the VCR is at 10°, and the Poster is at $-60°$, then the direction of Poster relative to the Computer can be calculated by subtracting the direction of Poster from the direction of Computer: $(-60°) - 65° = -125°$. This type of numerical subtraction may not be affected by the angular disparity between the imagined heading and the actual heading, while the pointing task relies on a simulated rotation strategy and performance decreases as the angular disparity increases. Thus, the pointing task and the verbal task may involve different representations (i.e., visual vs. verbal coding of the target directions) and different processes (i.e., simulated rotation vs. numerical subtraction). Although the verbal subtraction strategy is plausible, most participants reported that they directly estimated the angles between the imagined heading and the target, rather than maintaining the angles of each target and then doing subtraction for each trial.

Another possibility is that the two tasks depend on different types of spatial representations with different reference frames. For example, Biederman and colleagues (Biederman, 1987; Biederman & Gerhardstein, 1993; Cooper, Biederman, & Hummel, 1992) suggested that certain types of object-centered coding of spatial relationships may allow an object be recognized from different viewpoints equally well. Similar types of orientation-free representations were shown by Evans and Pezdek (1980; Sholl, 1987; etc.)

for navigation-scale environments, where participants were able to directly recall the spatial relationships from different perspectives. Thus, it is possible that the verbal task relies on a view-independent representation while the pointing task relies on a view-dependent one, leading to an angular disparity effect in the pointing task but not in the verbal task.

A third possibility comes from recent research (e.g., Brockmole & Wang, 2003; May, 1996, 2004; Wang, 2003) suggesting that the angular disparity effect in a pointing task is not due to simulated rotation of one's body, but due to interference from representations of the actual target locations. Thus, it is possible that both verbal and pointing tasks rely on similar representations. However, the action system is strongly connected to the representation of objects' actual locations, which limits its ability to use information from the representation of the imagined perspective. In contrast, a verbal system may not be tied to a specific view and thus can choose more flexibly with little interference, leading to a strong angular disparity effect in the pointing task but not in the verbal task.

Despite the apparent difference in the pattern of results, the current findings are consistent with theories of action and perception. The current study suggests that different representations and processes underline action and judgment/communications. This distinction clearly has evolutionary significance: although one would like to be able to imagine other perspectives flexibly and make judgments and descriptions, it will be a disaster if our legs and hands are operated by those imaginations instead of reality. Therefore products of imagination may be used for judgments and descriptions freely, but these representations cannot be used as readily as the actual representations for actions such as pointing. Perception/communication and action rely on fundamentally different types of representations or processes.

Acknowledgements

Thanks to Elizabeth Spelke, Mark May and Steffen Werner for discussions on this project, and to Heather Bauer and Rashad Abdul-Salaam for data collection. Supported by NSF Grant BCS 03-17681 to R.F. Wang. Comments should be sent to Frances Wang (francesw@s.psych.uiuc.edu).

References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics*, *29*, 336–342.

Bridgeman, B., Peery, S., & Anand, S. (1997). Interaction of cognitive and sensorimotor maps of visual space. *Perception and Psychophysics*, *59*, 456–469.

Brockmole, J. R., & Wang, R. F. (2003). Changing perspective within and across environments. *Cognition*, *87*, B59–B67.

Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory and Cognition*, *29*, 883–892.

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, *46*, 223–224.

Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and further evidence. *Canadian Journal of Psychology*, *46*, 191–214.

Creem, S. H., & Proffitt, D. R. (1998). Two memories for geographical slant: Separation and interdependence of action and awareness. *Psychonomic Bulletin and Review*, *5*, 22–36.

De Vega, M., & Rodrigo, M. J. (2001). Updating spatial layouts mediated by pointing and labeling under physical and imaginary rotation. *European Journal of Cognitive Psychology*, *13*, 369–393.

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 483–500.

Evans, G. W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 13–24.

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 227–233.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 20–25.

Levinson, S. C. (1996). Frames of reference and Molyneux's's question: Crosslinguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press.

Loomis, J. M., Da Silva, J. A., Philbeck, J. W., & Fukusima, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, *5*, 72–77.

May, M. (1996). Cognitive and embodied modes of spatial imagery. *Psychologische Beitraege*, *38*, 418–434.

May, M. (2004). Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology*, *48*, 163–206.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, *23*, 1447–1455.

Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin and Review*, *2*, 409–428.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1157–1165.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin and Review*, *4*, 102–106.

Sholl, M. J. (1987). Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 615–628.

Wang, R. F. (2003). Spatial representations and spatial updating. In D. E. Irwin, & B. H. Ross (Eds.), *The psychology of learning and motivation* (*vol. 42*) (pp. 109–156). *Advances in research and theory: Cognitive vision*, San Diego, CA: Academic Press.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 151–168.