Title

Twists and Oliver Twists in Mental Rotation: Complementary Actions as Orphan Processes

Permalink

https://escholarship.org/uc/item/56r172cd

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN

1069-7977

Authors

Athreya, Dilip
Chandrasekharan, Sanjay
Srinivasan, Narayanan

Publication Date

2006

Peer reviewed

Twists and Oliver Twists in Mental Rotation: Complementary Actions as Orphan Processes

Sanjay Chandrasekharan (sanjayen@gmail.com)
Dilip Athreya (athreyadilip@yahoo.com)
Narayanan Srinivasan (ammuns@yahoo.com)
Centre for Behavioural and Cognitive Sciences
University of Allahabad, Allahabad 211002, India

Abstract

A growing body of work shows that compatible actions executed in parallel with cognitive tasks contribute beneficially to cognition, compared to incompatible actions. We investigate how such complementary actions are generated. Two models from imitation research, Associated Sequence Learning (ASL) and Active Intermodal Matching (AIM), were extended to develop models of complementary action generation. ASL postulates a general generation process based on learning, whereas AIM postulates a specialist process. Using a mental rotation task where participants tended to spontaneously generate parallel actions, we conducted two experiments to test the predictions of the extended models. Surprisingly, the results show that when compared to no-actions, complementary actions do not improve accuracy. The two experiments do not provide clear validation for either model of generation, but there is more support for the generalist model than the specialist one. We propose a revision to the generalist model based on this trend.

Keywords: Complementary Actions, Epistemic Actions, Imitation, Mental Rotation, Situated Cognition

Introduction

Actions compatible with cognitive tasks such as mental rotation and counting have been shown to contribute beneficially to cognition (Kirsh & Maglio, 1994; Kirsh, 1995; Kosslyn, 1994; Wexler, Kosslyn & Berthoz, 1998; Goldin-Meadow & Wagner, 2005). Recent work also argues that such actions may play a beneficial role in perception (Wexler & van Boxtel, 2005; Noe, 2004).

What is the mechanism underlying the generation of such complementary actions? This is the question we address in this paper. The paper is organized as follows: Section 1 reviews some of the evidence that supports the beneficial role of action in cognition. Section 2 examines two possible models of the mechanisms underlying such actions, and their predictions. Section 3 presents our experiments and results. Section 4 discusses how the results relate to the models. We conclude with future work.

Action Supporting Cognition

Most studies examining the link between action and cognition report that actions compatible with cognitive tasks play a beneficial role in cognition. The most influential study in this area is Kirsh and Maglio (1994), which showed that even in a fast-paced task environment like the Tetris video game, players use actions to lower computational load. Tetris involves maneuvering falling shapes (zoids) into specific arrangements on the screen. Players execute actions on the falling zoids, to expose information early, to prime themselves to recognize zoids faster, and to perform external checks and verifications to reduce the uncertainty of judgments. The point of taking such actions is "is not for the effect they have on the environment as much as for the effect they have on the agent" (Kirsh & Maglio, 1994).

The authors term such actions '*epistemic actions*', which are defined as "physical actions whose primary function is to improve cognition by: 1) reducing the memory involved in mental computation; 2) reducing the number of steps involved in mental computation; 3) reducing the probability of error in mental computation" (Kirsh & Maglio, 1994).

The primary computations involved in Tetris are mental rotation of the zoids and matching of zoids to available slots. The participants physically rotate the zoids to significantly lower the amount of mental rotation required to judge the 'fit' of a zoid to available slots. This involves a visual comparison between slots and the physically rotated zoids. However, a visual comparison is *not* required for actions to aid in mental rotation. Wexler et al (1998) show that *unseen* motor rotation in the Cooper-Shepard mental rotation task (Cooper & Shepard, 1973) leads to faster reaction times and fewer errors when the motor rotation is compatible with the mental rotation than when they are incompatible. They also report that in some cases motor rotation made complex mental rotations easier. Also, speeding up the motor rotation speeded up the mental rotation, while slowing the motor action slowed down the mental one. Similar effects have been shown to exist in children (Frick, Daum, Walser & Mast, 2005). Manipulating virtual objects have also been reported to improve subsequent mental rotation and recognition of such objects (Wexler & van Boxtel, 2005).

Besides the above direct evidence, Kosslyn (1994) reports extensive indirect evidence for the role of action in mental rotation, including a study that showed participants need more time to perform mental rotations that are physically awkward, and another one where incompatible movements disrupted memory. Kosslyn (1994) also refers to a brain-damaged patient who consistently reached up to the screen

and pretended to 'twist' the stimulus in a rotation task (as in Tetris), and participants in the classic Shepard and Metzler experiment reporting "kinesthetic imagery" in their hands.

On a different vein from mental rotations, Kirsh (1995) reports higher accuracy in a coin-counting task when participants pointed at the stimulus, compared to a no-pointing condition. Gestures during cognitive tasks have been shown to lower cognitive load and promote learning (Goldin-Meadow & Wagner, 2005). Humans and other animals exploit head and eye movements to better perceive depth, absolute distance, heading and 3D objects (Wexler & van Boxtel, 2005). Bergen (2004) reports that processing time for sentences involving actions increases when participants perform incompatible actions in parallel.

All the actions reported in the above brief review do not meet the epistemic action criteria set out by Kirsh & Maglio (1994), so we will use the more general term 'complementary actions' to refer to such compatible actions generated during cognitive tasks.

Complementary Actions and Imitation

How are such actions generated? We will extend two models from imitation research to investigate this question.

A recent review (Brass & Heyes, 2005) succinctly captures the central problem in imitation: "When we observe another person moving, we do not see the muscle activation underlying the movement, but rather the external consequences of that activation. So how does the observer's motor system 'know' which muscle activations will lead to the observed movement?"

This last question can be used to reframe the generation question for complementary actions: how does the participant's motor system 'know' which muscle activations will lead to 'compatible' actions in a task? Further, how does it 'know' when to generate such actions?

One possible answer is: it doesn't 'know'. In imitation, this view is termed Associative Sequence Learning (ASL), where the visual and motor components are considered to become linked through Hebbian learning, and imitation is an *automatic* activation of motor representations when observing an action (Brass & Heyes, 2005).

A large body of imaging evidence shows the automatic activation of motor representations while observing actions (for reviews see Metzinger & Gallese, 2003; Svenson & Ziemke, 2004; Brass and Heyes, 2005, Gallese, 2005). It has also been shown that motor areas are activated more while participants observe human hands than robotic hands, and motor areas are not activated when humans watch actions not part of human repertoire (such as barking). Related studies show more motor activation for dancers while watching dance and pianists while watching piano playing.

Behaviorally, there is only indirect evidence for the automatic activation model. Most experiments are based on an interference paradigm similar to the one used by Wexler et al. (1998). An example is the finger-tapping paradigm that illustrates a variant of the Simon Effect (Simon, Sly, & Villapakkam, 1981), where movement execution is faster when accompanied by observation of a congruent movement than with an incongruent movement (Brass, Bekkering & Prinz, 2002).

The generalist view of action generation would predict that such activation of motor representations is automatically triggered, therefore "*they are not expected to be restricted to situations where imitation is intended*." (Brass & Heyes, 2005) This is in contrast to a specialist view, termed Active Intermodal Matching (AIM) which postulates a special mechanism mediating imitation, where a supra-modal representation of the action to be imitated is generated. This mechanism would allow the "switching on" of the motor module only when imitation is intended (Brass & Heyes, 2005, Heyes, Bird, Johnson, & Haggard, 2005).

These two models of the mechanisms underlying imitation can be applied directly to the question of how complementary actions are generated. A generalist model, based on the learned link between visual and motor components, would predict that compatible actions would be automatically activated while observing visual stimuli involving movement. Therefore this activation *would not* be limited to situations where the actions contribute beneficially to the tasks. In contrast, a specialist model would predict a "switching on" of the motor module only when the compatible action is beneficial.

Two experiments were conducted to test these two models of complementary action generation using a mental rotation task. Briefly, the experiments consisted of showing participants a rotation operation, which they had to remember. They were then presented a target pattern, along with four rotated versions of the same pattern (answers). The participants were then asked to execute the remembered rotation operation on the target pattern, and choose from the four options the right answer, i.e. the result of the rotation. The rotation operation had two levels of complexity, low and high. Pilot studies showed that participants tended to significantly generate hand or head rotations during the task.

Experiment 1

In the first experiment (Voluntary action condition), we presented participants with the stimuli, keeping track of the trials where participants rotated their hands (and heads), and the accuracy for the action and no-action cases. This experiment had two objectives: one, see how often actions were generated and when; two, see how the actions interacted with accuracy in the rotation task.

On objective one, if actions were generated in most trials, that would indicate an automatic mechanism. But if they were executed mostly in the high complexity trials, that would indicate a specialist mechanism. On objective two, if the participants who used more actions had more accuracy compared to participants who used actions less, that would indicate actions are beneficial, and are "switched on" because they are beneficial. This would suggest a specialist module directing the action. If the participants who used actions had lower accuracy compared to the participants who used actions less, that would indicate an automatic

mechanism triggering the movement. This would support the generalist view.

One significant caveat need to be added to the above logic. If the activation is not inhibited and is fully automatic (as we provisionally assume here), there need not be any link between activation of action and complexity, same with accuracy (Heyes, personal communication). Some inhibition of the automatic process needs to be assumed for such a link to exist, and the ASL model does assume this. We also eventually reach this conclusion (see general discussion). However, our experiments were based on the *overt* action we observed during pilot studies, which requires the generation process to be not inhibited. The fully automatic model we assume above is a *provisional* assumption based on this pilot data, intended only to frame the experiments and discussion.

It is also worth noting here that most of the other rotation studies compare compatible actions with incompatible ones, and not action with no-action.

Method

Participants: Twenty three student volunteers from University of Allahabad, with normal or corrected-to-normal vision. None had laboratory experience with mental imagery.

Apparatus: A computer screen, microphone and keyboard, placed on a table in front of the subjects. The screen was parallel to participants' frontal plane, at eye level and approximately 75 cm from the participant.

Stimuli: A set of four small 2D patterns within a white square (frame) were prepared on a 3x3 matrix with only five cells being filled, as illustrated in Fig. 1. The visual angle was 1.5 o x 1.5o. With each of these patterns, three more patterns were generated by rotating the original four patterns by 90°, 180° or 270°. Any one of the four orientations of a particular stimulus pattern was randomly used as a stimulus in a particular trial.

Figure 1: The four basic patterns used in the study

There were eight rotational operations (see Fig. 2) with two levels of complexity (low or high). Each level of complexity had four operations. Low complexity operations were rotations of 90° (right and left) and 180° (right and left). High complexity operations were vertical and horizontal flips followed by a rotation of 90° (left or right).

The rotational task was given a reference by providing an empty-blank white square (frame). To demonstrate the operations, video clips were created using Flash (samples at: http://www.sce.carleton.ca/~schandra/CAflash). Each

rotation in the low complexity condition took twenty seconds of display time. In the high complexity condition, each flip operation took twenty seconds in addition to each rotation operation, which also took twenty seconds to complete. There was a two second gap between flip and rotation. The end position (frame), after the rotational operation completed, stayed for five seconds.
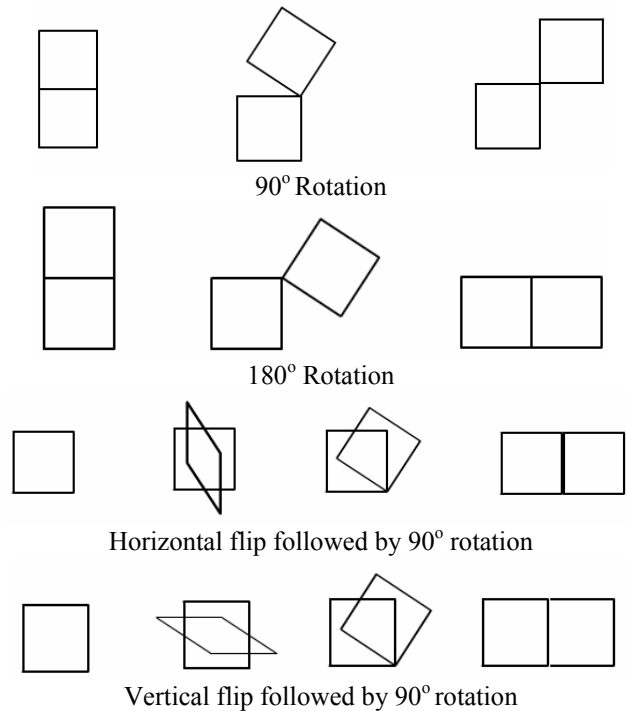

90° Rotation

180° Rotation

Horizontal flip followed by 90° rotation

Vertical flip followed by 90° rotation

Figure 2: Snapshots of the rotation operations

Procedure: The experiment consisted of thirty two trials (8 operations x 4 patterns), presented randomly. Each trial had two phases. In the first phase, a rotation was demonstrated using a video clip. Participants were asked to remember the rotation they saw, apply the same operation on the pattern coming up in the second phase and select the answer that best fitted the mentally rotated pattern.

The second phase started after four seconds, during which the screen was blank. This phase presented a pattern to be mentally rotated, along with four possible answers (as shown in Fig. 3), which remained on screen until participants produced a voice response. Participants first said their choice (1, 2, 3 or 4) aloud into the microphone, and then typed their choice in the textbox that appeared following the voice response. They then pressed the Enter key to initiate the next trial, which started after two seconds. Commercially available software (DirectRT, running on a PC with a VGA monitor) was used for stimuli presentation and data collection.

The experimenter sat beside the participant and used a chart to document the trials in which the participant generated complementary actions. There were three kinds of

actions -- finger, wrist and elbow movements. Only movements with an arc of roughly more than 45 degrees were considered as actions. There were head movements as well, but mostly associated with hand movements.
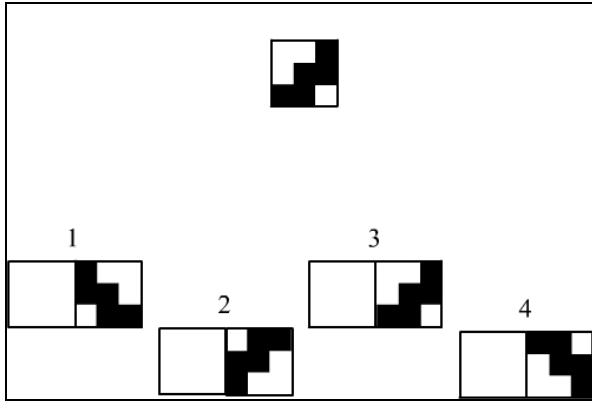

Figure 3: The screen during the second phase

Results and Discussion

Use of hands: All participants generated actions. However, out of twenty three participants, seventeen used hands extensively (in more than 50% of trials: high-action-generation group) and six did not use hands significantly (in more than 50% of trials: low-action-generation group). The high-action-generation group used their hands in 83% of the trials while the low-action-generation group did not use their hands in 84% of the trials. Head movements were ignored in this analysis, as only one participant used them in isolation (i.e. without parallel hand movements), and those head movements were minute.

A one-way within ANOVA (complexity: low, high) was performed on the percentage of trials in which hands were used by the participants. Among the high-action-generation group, participants used hands mostly in the high complexity condition compared to the low complexity condition $F(1,16) = 28.509$ $p < 0.001$. Within the high complexity condition, they used their hands in 96% of trials, and within the low complexity condition, they used their hands in only 70% of trials.

Among the low-action-generation group, participants used hands in 18.7% of trials in the high complexity condition and 13.5% of trials in the low complexity condition. The difference indicates a similar trend $F(1,5) = 3.049$ $p = 0.141$, although the effect was not significant given the small number of subjects.

These results present a mixed bag for the generalist model. The high rate of use indicates a generalist process, but a focused mechanism is implied by the way the use of hands went up significantly (in the high-action-generation group) as the task became harder (see general discussion).

Accuracy: The accuracy results are shown in Figure 4. The accuracy for the high-action-generation condition (only trials with use of hands) and the low-action-generation condition (only trials without the use of hands) were taken for further statistical analysis. A 2 between (action: high-action-generation, low-action-generation) x 2 within (complexity: low, high) ANOVA was performed on the accuracy values from all the participants. The results show that the performance without hands in the low-action-generation condition (0.812) was significantly better than with hands in the high-action-generation condition (0.604) $F(1,21) = 6.322$, p<.05. The results also show that performance in the low complexity condition (0.781) was significantly better than performance in the high complexity condition (0.536) $F(1,21) = 20.576$, $p < 0.001$.
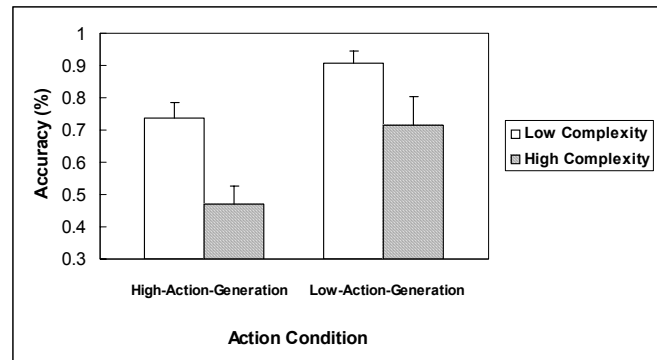

Figure 4: Accuracy with High-action-generation and Low-action-generation

The results indicate that in spite of the fact that the majority of the participants used their hands, their performance was significantly lower than those participants who did not use their hands. Even in the high complexity condition where hands were used on almost all the trials (the high-action-generation group) the performance was lower compared to those who used hands less. These results show that *not all* complementary actions generated during a task lead to better performance, some may actually be interfering with the task.

What about the mechanism underlying generation? Since there was extensive use of hands, but the use lowered accuracy, the results support the generalist model, i.e. automatic activation of motor components. The next experiment tested this further, exploring how enforcing action and curtailing action affected performance.

Experiment 2

If compatible actions are activated automatically (as the generalist view holds), being forced to use hands should not make any difference in accuracy, compared to the voluntary action condition. On the other hand, if actions get activated only in a guided manner (as the specialist view would hold), being forced to use hands would interfere with the task and lower accuracy, compared to the voluntary action condition.

Similarly, if compatible actions are activated automatically (as the generalist model holds), then

restricting all action should lower accuracy compared to the voluntary action condition, as participants have to spend effort to not move their hands. However, if actions are activated in a guided manner, there will not be any effect on accuracy compared to the voluntary action condition, as the action is under voluntary control anyway.

To explore these possibilities, two task conditions were used with two groups of participants. In one condition, we curtailed all hand (and head) movements (action-curtailed condition). In the other condition (action-enforced condition), participants were required to use their hands.

Method

Participants: Twenty eight volunteers with normal or corrected-to-normal vision from the University of Allahabad. None had laboratory experience with mental imagery. They were randomly assigned to one of the groups, action-curtailed or action-required.
Apparatus and Stimuli: Same as in Experiment 1.
Procedure: The stimulus presentation was same as in Experiment 1. For the action-curtailed condition, participants were asked to keep their hands flat on the table, and not move once the trial started. After providing a voice response to the four choices, they could move, to type their choice into the textbox. For the action-enforced condition, participants were asked to use their hands in some way, but the way in which to use hands was left open to their choice.
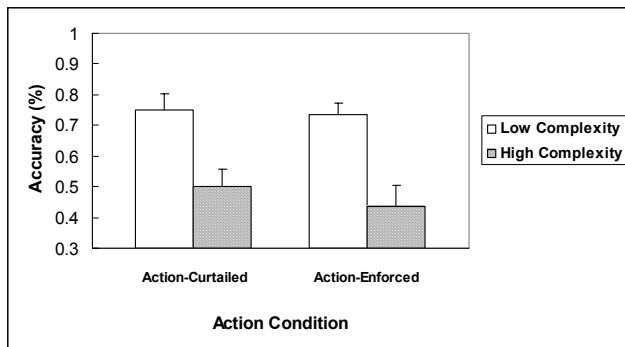
Results and Discussion

Figure 5: Accuracy with action-restricted and action-required conditions.

Figure 5 shows the mean accuracies for both the conditions. A 2 between (action-enforced, action-curtailed) x 2 within (complexity: low, high) ANOVA was performed with the accuracy values. There was no significant difference in accuracy between the action-curtailed and action-enforced conditions. The use of action did not result in better performance, compared to the no-action condition. Similar to the first experiment, complexity had a significant effect $F(1,26) = 26.46$, p <0.001 with better performance in the low complexity condition (0.73) compared the high

complexity condition (0.48). The results show that complementary actions are not always advantageous.

Table 1 captures the accuracy results for the three major conditions. There is very little difference between the three cases. This means the results are mixed for the two models. The enforcement of action has no effect compared to the voluntary action condition (high-action-generation condition from Experiment 1), and this supports the generalist view, as enforcing actions is equivalent to automatic action. But the curtailment of action also has no effect on accuracy, which goes against the generalist model, as we would expect the curtailing of an automatic action to take effort, and to interfere with the task.

Table 1: Accuracy results for the three major conditions

Level	Voluntary action	Action-enforced	Action-curtailed
Low complexity	0.736	0.725	0.735
High complexity	0.472	0.437	0.524

The Orphaned-Process Model

Of the four experimental variables we considered (voluntary-action-activation, voluntary-action-accuracy, enforced-action-accuracy and curtailed-action-accuracy) the results from the first three support the generalist view. The last result and the complexity effect seem to support the specialist model. This indicates that an in-between mechanism likely underlies complementary action generation. To understand this mechanism, Table 2 presents the results as supporting the generalist model and otherwise.

Table 2: Results in relation to the generalist model

Experiment	*Support for the generalist model*	*Anomalous results for the generalist model*
Voluntary action (activation)	More action than not	More action in high complexity
Voluntary action (accuracy)	Less accuracy (compared to no-action)	
Enforced action (accuracy)	Accuracy unchanged	
Curtailed action (accuracy)		Accuracy does not go down

The activation of action in high complexity cases indicates that there is a link between processing load and action generation. However, this does not necessarily mean a focused, specialist generation process. It could be the case that action components are automatically activated while viewing moving stimuli (as indicated by imaging studies), but this action is inhibited. However, this 'caretaker' inhibition process takes up processing resources. During

high-processing-load tasks, this process could lose resources, leaving the automatically activated motor component 'orphaned'. Such a reallocation of processing resources would lead to compatible actions being activated overtly during high-processing-load tasks.

This 'orphaning' mechanism extends the ASL model beyond plain inhibition to cognitive load-modulated inhibition, and explains the two puzzles raised in section 3: why actions are generated mostly in high processing conditions, and why mostly compatible actions are generated.

The orphan model also explains the second anomalous result for the generalist model – why accuracy does not go down in the action-curtailed condition. In this condition, the hands are tabled overtly before the beginning of the task. But since the action is expressed overtly, the motor component does not require a covert 'caretaker' process. An analogy would be a person putting his hands in his pocket to control a gesture obsession, instead of trying to remember not to gesture. The overt action makes use of a physical control, so no cognitive resources are allocated to inhibit the action. The tabling of hands is a similar overt inhibition. This means no extra resources are taken up by the tabling, which explains why there is no lowering of accuracy.

Future Work

The orphan model does not fully explain the results, however. For instance, why were actions rarely activated in six participants during the voluntary condition? Individual differences in mental rotation abilities could explain this, but this needs investigation. Our analysis indicates such differences would not affect the results and models reported here, but may help refine them (low mental rotation abilities may lead to more automatic action generation).

Secondly, during the action-curtailed condition, some people found it very hard to not use their hands. Their fingers would start twitching, or their neck would start moving imperceptibly. This suggests that some effort is required to keep actions from expressing overtly. It is unclear whether this takes processing resources away from the mental rotation task and requires more investigation.

We are now implementing an ERP study based on the above experiments, as one way to test the orphan process model would be to isolate the time course of the actions, and identify the possible cortical areas involved in the three experimental conditions. We are also running a set of dual-task experiments to test the postulated link between processing complexity and the overt expression of action.

Acknowledgments

References

Bergen, B., Chang, N., Narayan, S. (2004). Simulated Action in an Embodied Construction Grammar. In K. D. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago. Hillsdale, NJ: Lawrence Erlbaum.

Brass, M. & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*, 489-495

Brass, M., Bekkering, H., & Prinz, W. (2002). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*.

Cooper, L.A. & Shepard, R. N. (1973). The time required to prepare for a rotated stimulus. *Memory and Cognition*, *1*, 246-250.

Frick, A., Daum, M.M., Walser, S., Mast, F.W. (2005). Developmental changes in the interference of motor processes with mental rotation. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, CogSci2005, Stresa, Italy.

Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, *4*. 23-48

Goldin-Meadow, S. & Wagner, S.M. (2005). How our hands help us learn. *Trends in Cognitive Sciences*, *9*. 234-240

Heyes, C., Bird, G., Johnson, H., Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, 22, 233-240

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513-549.

Kirsh D. (1995). Complementary Strategies: Why we use our hands when we think. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

Kosslyn, S.M. (1994). *Image and Brain*. Cambridge, MA: MIT Press

Metzinger, T & Gallese, V. (2003) The emergence of a shared action ontology: Building blocks for a theory. *Consciousness and Cognition*, *12*, 549 – 571, Special Issue on Grounding the Self in Action.

Noe, A. (2004). *Action in Perception*. Cambridge, MA. MIT Press

Simon, J.R. Sly, P.E., Villapakkam, S. (1981). Effect of compatibility of S-R mapping on reactions toward the stimulus source. *Acta Psychologica*, 47, 63-81

Svenson, H. and Ziemke, T. (2004). Making Sense of Embodiment: Simulation Theories and the Sharing of Neural Circuitry Between Sensorimotor and Cognitive Processes. In K. D. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago. Hillsdale, NJ: Lawrence Erlbaum.

Wexler, M., & van Boxtel, J.J.A., (2005). Depth perception by the active observer, *Trends in Cognitive Sciences*, *9*, 431-438.

Wexler, M., Kosslyn S.M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*, 77-94.