2015 Special Issue

Generalisation, decision making, and embodiment effects in mental rotation: A neurorobotic architecture tested with a humanoid robot

Kristsana Seepanomwan [a], Daniele Caligiore [b], Angelo Cangelosi [a], Gianluca Baldassarre [b,*]

[a] University of Plymouth, Centre for Robotics and Neural Systems, United Kingdom
[b] Consiglio Nazionale delle Ricerche, Istituto di Scienze e Tecnologie della Cognizione, Italy

ARTICLE INFO

ABSTRACT

Mental rotation, a classic experimental paradigm of cognitive psychology, tests the capacity of humans to mentally rotate a seen object to decide if it matches a target object. In recent years, mental rotation has been investigated with brain imaging techniques to identify the brain areas involved. Mental rotation has also been investigated through the development of neural-network models, used to identify the specific mechanisms that underlie its process, and with neurorobotics models to investigate its embodied nature. Current models, however, have limited capacities to relate to neuro-scientific evidence, to generalise mental rotation to new objects, to suitably represent decision making mechanisms, and to allow the study of the effects of overt gestures on mental rotation. The work presented in this study overcomes these limitations by proposing a novel neurorobotic model that has a macro-architecture constrained by knowledge held on brain, encompasses a rather general mental rotation mechanism, and incorporates a biologically plausible decision making mechanism. The model was tested using the humanoid robot iCub in tasks requiring the robot to mentally rotate 2D geometrical images appearing on a computer screen. The results show that the robot gained an enhanced capacity to generalise mental rotation to new objects and to express the possible effects of overt movements of the wrist on mental rotation. The model also represents a further step in the identification of the embodied neural mechanisms that may underlie mental rotation in humans and might also give hints to enhance robots' planning capabilities.

1. Introduction

Mental imagery concerns cognitive processes for the creation and manipulation of mental images, and for decision making tasks on visual object matching (Kosslyn, 1996; Lamm, Windischberger, Moser, & Bauer, 2007). In a typical mental rotation experiment of cognitive psychology, a participant has to mentally rotate an object perceived in a picture to decide if it is the same as a target object or different from it (i.e. a flipped version of it), and then indicate the answer by pressing one of two buttons (Shepard & Metzler, 1971; Wexler, Kosslyn, & Berthoz, 1998). In this kind of task, participants normally report that in order to make the decision they mentally rotate one object, clock-wise or counter clock-wise, until it visually matches or mismatches the target object. The actual existence of this process is supported by the main result of mental rotation experiments: the reaction time to press one of the two buttons, and the error rate of the answers, increase with the angular disparity between the rotated object and the target object.

Mental rotation has been widely investigated not only in cognitive psychology, but also in cognitive neuroscience and computational modelling (Kosslyn, 1996; Zacks, 2008). Initially, it was proposed that the brain mechanisms underlying mental rotation mainly involve visual and spatial perception systems (Corballis & McLaren, 1982; Shepard & Metzler, 1971). More recently, behavioural (Wexler et al., 1998; Wohlschläger, 2001) and neuroscientific experiments (Georgopoulos, Lurito, Petrides, Schwartz, & Massey, 1989; Lamm et al., 2007) have suggested the idea that mental rotation relies on a mentally simulated action (Michelon, Vettel, & Zacks, 2006) rather than on a purely visual and spatial imagery skill. Brain-imaging evidence on the brain areas most involved in mental rotation supports the idea that mental rotation indeed depends on a strong integration of sensorimotor

* Correspondence to: Via San Martino della Battaglia 44, 00185 Roma, Italy. Tel.: +39 06 44595231.
E-mail addresses: kristsana.seepanomwan@plymouth.ac.uk (K. Seepanomwan), daniele.caligiore@istc.cnr.it (D. Caligiore), a.cangelosi@plymouth.ac.uk (A. Cangelosi), gianluca.baldassarre@istc.cnr.it (G. Baldassarre).
URL: http://www.istc.cnr.it/people/gianluca-baldassarre (G. Baldassarre).

processes and covert mental simulation of motor movements (see Zacks, 2008, for a review). In particular, based on refined experimental paradigms, mental rotation has been proposed to involve the following processes (Lamm et al., 2007): (a) encoding of stimuli and the generation of mental images; (b) planning and execution of mental rotation; (c) comparison (matching) of the rotated stimulus with the target stimulus; (d) decision and performance of the same/different response. The next subsection presents a focused review of the empirical evidence on the brain areas involved in these processes.

Computational models of mental rotation have been proposed to replicate some limited aspects of human mental rotation processing. Most of these models tend to focus on the rotation of objects having different rotation axis and shapes to support object recognition, and have been developed using artificial neural network design (Fukumi, Omatu, & Nishikawa, 1997; Fukumi, Omatu, Takeda, & Kosaka, 1992; Inui & Ashizawa, 2011; Kulkarni, Yap, & Byars, 1990; Rowley, Baluja, & Kanade, 1998; Sasama, Mitsumoto, Yoneda, & Tamura, 2009). For example, Sasama et al. (2009) proposed an error back-propagation neural network model that takes two images as input and produces one binary vector as an output that encodes the angular disparity between the two input images together with a response answer (match/mismatch). Similarly, Inui and Ashizawa (2011) proposed a radial basis function neural network to mentally rotate 3D objects. These models use neural networks but these networks have not been designed to reproduce brain mechanisms suggested to underlie mental rotation in humans. Moreover, not being used to control whole embodied systems, they do not refer to the involvement of motor mechanisms in mental rotation, and do not consider other important processes supporting the mental rotation core processes (e.g. the monitoring of the overall mental-rotation success and the decision making process needed to produce the final answer). As a consequence, they reproduce only a few aspects of the overall mental rotation processes and give little contribution to the growing neuroscientific literature aiming to understand the brain neural mechanisms underlying mental rotation.

Computational *robotic* models have also been used to study mental rotation. The embodied nature of these models allows a better representation of the sensorimotor aspects involved in mental rotation. These aspects have indeed been neglected by prior computational models of mental rotation (e.g., Inui & Ashizawa, 2011; Sasama et al., 2009). Although cognitive robotics models of mental simulation have been recently proposed (Di Nuovo, De La Cruz, & Marocco, 2013), these do not directly address mental rotation capabilities, but rather mental simulation for motor planning tasks. They also do not propose hypotheses on the brain mechanisms that may underlie them (e.g., Di Nuovo, Marocco, Cangelosi, De La Cruz, & Di Nuovo, 2012). Robotic models of mental rotation also have a technological valence. Indeed, endowing robots with mental rotation capabilities could increase their planning abilities in relation to the manipulation of objects, in particular, as planning in robots has mainly focused on navigation (Baldassarre, 2001, 2003; Dissanayake, Newman, Clark, Durrant-Whyte, & Csorba, 2001; Meyer & Filliat, 2003) and reaching tasks (Khatib, 1986; Masehian & Sedighizadeh, 2007) rather than on object manipulation. Planning applied to object manipulation could allow robots to mentally simulate in advance the consequence of potential actions, e.g. to perform only useful actions, to best concatenate actions to accomplish complex goals, and/or to generalise actions to new objects and contexts. As further discussed in Section 4 (Conclusions), mental rotation seen as a form of planning has also some peculiarities related to the low number of actions involved, and to the fact that it involves spatial transformation processes that are invariant with respect to the object involved (Terekhov & O'Regan, 2013), which may suggest interesting solutions for planning problems with this type of features.

We (Seepanomwan, Caligiore, Baldassarre, & Cangelosi, 2013a, 2013b) recently proposed a neural-network model whose macro architecture was linked to brain macro areas. This model was able to solve a simple mental rotation task of 2D visually-perceived objects in a simulated humanoid robot, the iCub (Tikhanoff et al., 2008). This model, a predecessor to the model proposed here, was developed within an "embodied cognition" theoretical framework for which high-level cognition processes rely on the same areas of the brain used to process analogous sensorimotor information (Borghi & Cimatti, 2010). According to this view, off-line cognition, such as mental rotation and imaging, is body based: "even when decoupled from the environment, the activity of the mind is grounded in mechanisms that evolved for interaction with the environment—that is, mechanisms of sensory processing and motor control" (Wilson, 2002). The model departed from another model – TRoPICALS – developed within the "computational embodied neuroscience" framework and aiming to establish detailed links between embodied cognition and behaviour and the brain system-level mechanisms underlying them (Caligiore, Borghi, Parisi, & Baldassarre, 2010; Caligiore, Borghi et al., 2013; Caligiore, Pezzulo, Miall, & Baldassarre, 2013; TRoPICALS focused on compatibility effects, Tucker & Ellis, 2001, and affordance processing, Gibson, 1986; Rizzolatti & Craighero, 2004). The model (Seepanomwan et al., 2013a, 2013b) reproduces some mechanisms, possibly performed in the parietal–premotor brain circuits, implementing the object mental rotation processes and some other mechanisms, possibly performed in prefrontal–premotor circuits, implementing the decision making processes involved in mental rotation (see Zacks, 2008, and Section 1.1, for more details on the related biological mechanisms). To our knowledge, this model represents the first instance of a neuro-robotic model of mental rotation, and a first hypothesis of the brain mechanisms that may underlie this process. However, despite its ability to solve a typical mental rotation task, the model has significant limitations. First, it lacks mental rotation generalisation capability for novel objects. Second, it generates error rates related to mental rotation tasks that do not reflect the inherent difficulty of the tasks themselves: this is due to the decision making component of the model being based on a rigid non-biologically plausible mechanism that leads the model to an abrupt drop in performance when the images to be rotated become increasingly complex. Third, it does not fully exploit the sensorimotor possibilities rendered by its robotic embodied nature, for example to investigate the interesting interference/synergy effects that current proprioception and gestures can have on mental rotation (e.g., see Wexler et al., 1998; Wohlschläger & Wohlschläger, 1998). Finally, the model was tested with the iCub simulator but not with the hardware robot.

In this study we propose a new neuro-robotic model of mental rotation that builds upon the prior model proposed in Seepanomwan et al. (2013a, 2013b) and overcomes its limitations discussed above. Specifically, the new model has generalisation capabilities to transfer the mental rotation processes acquired with a small set of 2D visual training stimuli to novel 2D visual objects. Moreover, it employs a flexible decision making mechanism, based on biologically plausible models of decision making (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Usher & McClelland, 2001), that reproduces an error rate that varies gradually with the difficulty of the mental rotation. Further, its mental rotation capabilities could be challenged with overt movements of the robot, congruent or incongruent with the covert mental rotation process, to investigate their interactions (we anticipate that the model led us to formulate a hypothesis for which proprioception signals that are congruent with mental rotation can improve its performance, as in empirical experiments, only if the mental rotation task is made difficult by some noise sources). The model is tested through the hardware iCub humanoid robot (Metta,
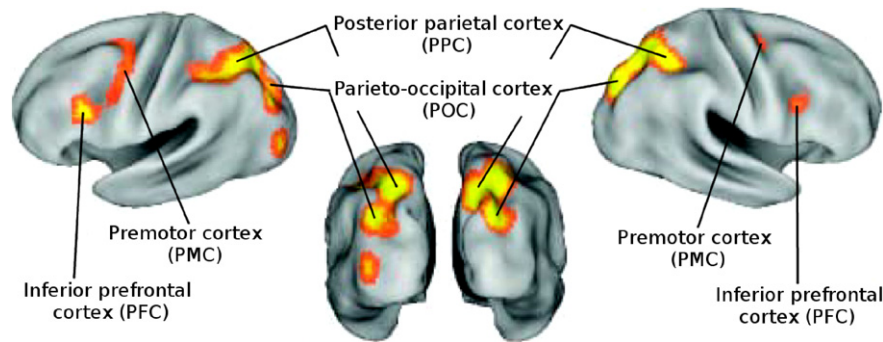
Fig. 1. The key brain areas involved in mental rotation and considered in the model. The red–yellow–green colouring highlights increasingly active areas. Left: brain lateral left hemisphere. Centre: posterior brain view. Right: brain lateral right hemisphere. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Reprinted with permission from Zacks (2008).

Sandini, Vernon, Natale, & Nori, 2008; Sandini, Metta, & Vernon, 2007) alongside the iCub simulator (Tikhanoff et al., 2008). This is relevant not only to facilitate the inclusion in the model of some issues relating to embodied cognition, but also to test the robustness of the model to the variable conditions of the environment and of the robot. For example, in the tests presented, the images from the robot camera changed in different trials due to luminance changes within the environment, the variable response of the camera and the accuracy limitations of the camera motors.

We anticipate that the model still contains various limitations, further discussed in Section 4, that represent important challenges for future work. These for example involve a largely non-neural management of the information flow between the system components, the coarse granularity of the mental rotation steps, and the simplicity of the object images used to test the model. Notwithstanding these limitations, to our knowledge the model currently represents the most articulated operational hypothesis of the neural mechanisms that might underlie mental rotation.

The rest of the paper is organised as follows. The next subsection illustrates the relevant neuroscientific evidence on the brain areas involved in mental rotation, used here to constrain the overall model architecture. Section 2 presents the robotic setup followed by a detailed explanation of the mental rotation task used to validate the model and finally a description of the main features of the model and the learning algorithms used to train it. Section 3 presents and discusses the results and Section 4 draws the final conclusions of the study.

1.1. Brain areas and neural mechanisms involved in mental rotation

Various areas of the human brain have been shown to be involved in mental rotation through functional magnetic resonance imaging (fMRI) techniques. A meta-review (Zacks, 2008) summarises the main areas that several studies have found to play a relevant role (Fig. 1).

Most brain imaging studies scanning the human brain during the performance of the mental rotation task show a prominent activity of the posterior parietal cortex and posterior-occipital cortex. In particular, the areas around the intraparietal sulcus (more specifically, the superior parietal lobule, Brodmann Area BA7, and the inferior parietal lobule, BA 40), and the areas surrounding the parieto-occipital sulcus (parieto-occipital arcus, BA 19) (Carpenter, Just, Keller, Eddy, & Thulborn, 1999; Harris & Miniussi, 2003; see Zacks, 2008, for a review). The activity of some of these areas also correlates with the amount of mental rotation requested in the different task trials and dependent on the object–target orientation disparity. Posterior parietal cortex receives input related to both visual and somatosensory information (Rizzolatti, Luppino, & Matelli, 1998), and on this basis it is

capable of elaborating information about the location and orientation of target objects in peripersonal and extrapersonal space, and their relation to own body (Andersen & Buneo, 2002; Colby & Goldberg, 1999), in large part employing eye-centred coordinate frames modulated by own body postures (Snyder, Grieve, Brotchie, & Andersen, 1998). Posterior-occipital cortex includes high-level visual areas encoding complex visual features, in particular related to movement (e.g. involving global and own movement, Braddick et al., 2001). Based on this evidence, these areas are thought to play a key role in implementing the proprioceptive and visual information integration and transformation supporting the core processes of the dynamic mental rotation processes (Zacks, 2008).

Other brain regions that consistently activate during the mental rotation experiment involve the supplementary motor area and the premotor cortex, in particular involving the medial precentral gyrus (BA6) (Cohen & Bookheimer, 1994; Johnston, Leek, Atherton, Thacker, & Jackson, 2004; Lamm et al., 2007; Zacks, 2008). These areas encode a repertoire of actions at a more abstract level with respect to primary motor cortex, and play important functions in motor planning and execution (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995). The activation of these areas strongly supports the involvement of motor processes in mental rotation, putatively to implement motor mental simulation. This possibility is corroborated by the fact that the supplementary motor area has been strongly involved in motor imagery (Stephan et al., 1995). Some studies also reveal an activation of primary motor areas, primarily linked to the production of the final response (button press) rather than to the main mental simulation processes (Richter et al., 2000).

Kosslyn, Digirolamo, Thompson, and Alpert (1998) and Zacks (2008) have also shown the activation of prefrontal areas, in particular the inferior lateral prefrontal cortex (inferior precentral sulcus, BA44/45). This region, part of Broca's area responsible for speech production, is involved in motor production and action recognition (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Given its high-level within the motor hierarchy, this area might orchestrate mental rotation at a high-level, as suggested by its role in motor imaging (Grafton, Arbib, Fadiga, & Rizzolatti, 1996).

Several components of the model are formed by neural maps using, in specific or abstract ways, *population codes*. Neural maps are suitable to model cortical areas as they capture their important 2D topological organisation and also facilitate the analysis and visualisation of the processes occurring within them (Caligiore, Parisi, & Baldassarre, 2014). Population codes (Pouget, Dayan, & Zemel, 2003) are based on the idea that information (on stimuli and actions) is encoded in the brain on the basis of the activation of populations of units organised in neural maps having a broad response field. In particular, each unit responds maximally to a
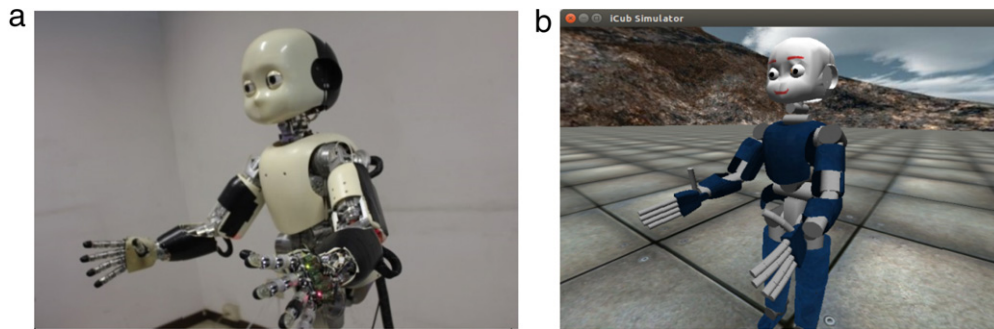
Fig. 2. The iCub humanoid robot used to reproduce the behaviour of a participant of mental rotation experiments. (a) Real robot. (b) Simulated robot.

certain value of the variables to encode and then progressively less intensely to more distant values. This response can be obtained with short-lateral excitatory connections and long-lateral inhibitory connections, or in a more abstract fashion (as for most maps used here) with Gaussian functions.

To implement the decision making process involved in the mental rotation task, the model uses a *mutual inhibition model* (Bogacz et al., 2006; Usher & McClelland, 2001). In this model (closely related to the architecture and neural competition that can be implemented by population-code maps) different decision options are represented by neural units that accumulate over time the evidence (support) on the goodness of the different options, compete through reciprocal inhibitory connections of the units, and finally produce a decision when the activation of one of them reaches a given threshold. This model (together with other analogous models, e.g. Bogacz et al., 2006) is very important, as it allows the reproduction of the reaction times often recorded in psychological experiments (Caligiore et al., 2010, 2008; Erlhagen & Schöner, 2002). It is one of the most accredited models of decision making processes taking place in the human brain (Bogacz, 2007).

In the brain, several processes needed to acquire and express mental rotation (e.g., learning from experience, and selection of cortical contents) are putatively implemented by cortical areas working in close cooperation with sub-cortical regions, in particular basal ganglia and cerebellum with whom they form whole integrated systems (Alexander, DeLong, & Strick, 1986; Baldassarre, Caligiore, & Mannella, 2013; Caligiore, Borghi et al., 2013; Caligiore, Pezzulo et al., 2013; Middleton & Strick, 2000). For simplicity, the model reproduces in abstract ways such processes, e.g. to implement the decision making processes and the mapping of the object representations to the corresponding arm postures, without explicitly simulating these sub-cortical systems.

2. Methods

2.1. The iCub humanoid robot

Fig. 2 shows the real and simulated iCub humanoid robot we used to reproduce the behaviour of the participants of the mental rotation experiment. The iCub is a many degrees of freedom (DOFs) robot built for studying cognitive development in humans (Cangelosi & Schlesinger, 2015; Sandini et al., 2007). In the tests of the model, the robot used one "eye" (i.e., one of its two 640 × 480 RGB cameras) to perceive the visual stimuli and the joint 5 of the right arm to control the wrist angle. The model guided the learning and performance of the robot during the mental rotation task. The robot did not hold real objects, but it could move its wrist in order to assume the wrist orientation corresponding (or not) to the orientation of the mentally rotated object, allowing us to study how the mental rotation process can be affected by overt action (i.e., the wrist movement and posture) similar to that of experiments involving humans (Wexler et al., 1998; Wohlschläger, 2001; Wohlschläger & Wohlschläger, 1998).

2.2. The stimuli

Fig. 3 shows the three sets of 2D abstract objects, broadly similar to those employed by Hochberg and Gellman (1977), used as stimuli during the mental rotation tasks. The stimuli were coloured in red to improve their detection by the iCub's camera. They were designed to create different levels of difficulty in the mental rotation task. Each set (A, B, C) consisted of three objects which could assume six orientations (90°, 60°, 30°, 0°, −30°, −60°) and could have a basic appearance (the one shown in Fig. 3) or an appearance corresponding to the "mirror" image of the basic appearance. The stimuli of the sets A and B, three for each set, contained a clear main axis (we will see this is important to perform mental rotation). However, the stimuli of set B were formed by more features than the stimuli of set A. Stimuli from set A were used for training the model and for a recognition test (Recog; Fig. 3(a)) whereas those from set B were used in a test directed to measure the generalisation capabilities of the model (Gen1; Fig. 3(b)). Stimuli from set C represented a more complex dataset that did not contain a strong orientation axis as in sets A and B. Set C was further used for a second, more challenging generalisation test (Gen2; Fig. 3(c)). Fig. 3(d) gives some examples of pairs of stimuli shown to the robot during the tests Recog, Gen1 or Gen2 (each test involved showing multiple pairs in different trials). In each object-pair image the object on the left was the target object (henceforth called "target object") and the object on the right was the one to be mentally rotated (henceforth "rotated object"). Each object was used to generate 144 object-pair images ($144 = 2^2 \times 6 \times 6$, where 2^2 is the number of the possible combinations of the basic and mirror appearance of the target and the rotated objects, and 6 are the initial possible orientations of the target and the rotated objects). The number of object-pair images used in each test was hence 432 (144×3, where 3 is the number of objects for each set), and the overall total number of object-pair images used over all three tests was 1296 (432×3).

2.3. The mental rotation task

The robot was tested in a mental rotation task similar to those typically used in mental rotation experiments with humans (e.g., Shepard & Metzler, 1971; Wexler et al., 1998). In this task, two visual stimuli having different orientation (90°, 60°, 30°, 0°, −30°, −60°) and appearance ("basic" or "mirror") combinations, are shown to the robot on a computer screen as illustrated in Fig. 4 (bottom right). The robot has to compare the two stimuli and decide if they are the same or different object. The model does not have attention control, so the behaviour of scanning the target and the rotated objects with the camera is hardwired and performed in sequence, thus allowing the robot to take a snapshot of each of the two objects.
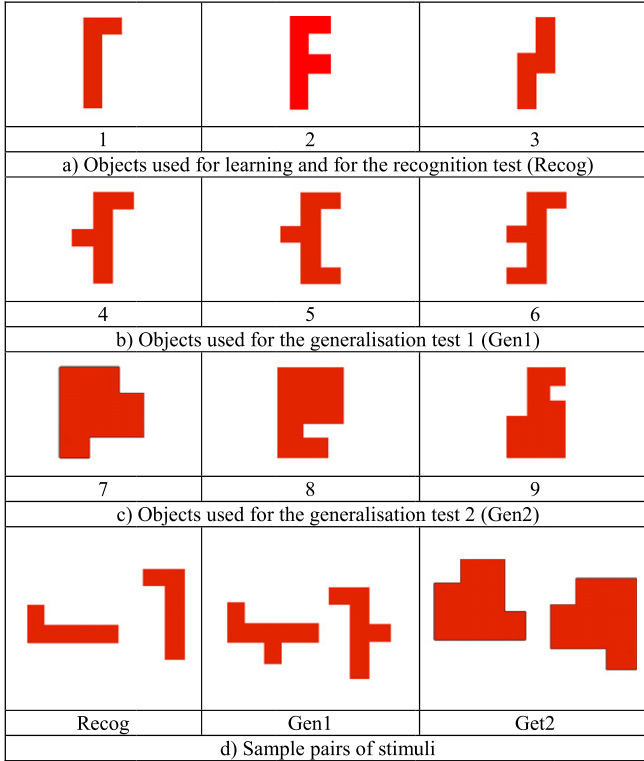
Fig. 3. (a) Stimulus set A, used for training and for the recognition test (Recog). (b, c) Stimulus sets used for the generalisation tests (respectively Gen1 and Gen2). (d) Three object-pair images used during the tests. In the three examples, the left object is the target object, here rotated 90° to the left, whereas the right object is the object to be rotated, here having a 0° rotation (in all three examples the rotated object is different from the target object).

Fig. 4. The neural network model controls the iCub robot during the mental rotation task. The bottom right picture shows the robot in front of the screen where it sees the target and rotated objects. Thin arrows indicate hardwired connections. (Fixed connections in most cases are set to 1.) Bold arrows indicate trained connections. Dashed arrows indicate information flows.

2.4. Model architecture and functioning of its components

The model architecture (Fig. 4) is formed by several components corresponding to the main brain areas involved in the mental rotation processes (as illustrated in Section 1.1, see Fig. 1). In this subsection we overview the model components and their role within the whole system, whereas in the following subsections we detail the mechanisms behind their functioning and learning and the reasons for their choice. The components of the model architecture are the early visual cortical areas (VC), the parieto-occipital cortex (POC), the posterior-parietal cortex (PPC), the premotor cortex (PMC), the prefrontal cortex (PFC), and the primary motor cortex (M1). Each component is formed by subcomponents (mainly neural maps) performing different functions. VC is an image-processing component that extracts the edges of objects from the current image in a way which is reminiscent of early visual cortex processes (Hubel, 1988). POC is formed by five neural maps of 32×32 units each: POCi encodes the current imaged orientation of the rotated object during mental rotation; POCl, POCs, and POCr anticipate the image of the rotated object if a left/still/right mental rotation of respectively $-30°, 0°, 30°$ of the current image encoded in POCi is performed. Based on the planned movement supplied by PMCm, only one of these three possible rotations is performed (for example leading to the rotated object image encoded in POCr). The resulting image is relayed to POCp which encodes the predicted rotated image depending on the performed rotation.

PPC is formed by three components: PPCp, which is formed by six units and encodes the proprioceptive signal related to the robot wrist orientation corresponding to the current actual or imaged orientation of the mentally rotated object encoded in POCp; PPCt, which is formed by six units and encodes the target orientation of the wrist corresponding to the orientation of the target object

encoded in PFCt; PPCc, which is formed by 6×6 units and combines the signals from the current imaged wrist orientation (PPCp) and its desired orientation (PPCt) to select a desired movement in PMCm.

PMC is formed by two components: PMCm, which is formed by three units that encode three possible movements (taking place in M1, here not explicitly simulated), i.e. respectively: Ml $= -30°$ corresponding to an anti-clockwise "left" rotation of the wrist; Ms $= 0°$ corresponding to a "stay", or null, rotation; Mr $= 30°$ corresponding to a clockwise "right" rotation; and PMCd, which is explained in detail below as it depends on PFC.

PFC is formed by two neural maps of 32×32 units each: PFCt, which represents a working memory map used to store the visual appearance of the target object; PFCtm, which computes the amount of overlap (matching) between the rotated object image (POCp) and the target image (PFCt) to support the decision making process implemented in PMCd.

Finally, PMCd is formed by two units activated by "evidence" from PFCtm on the current matching of the predicted rotated image (POCp) with the target object (PFCt). Broadly speaking, a large overlap provides evidence for a "YES" answer and a small overlap provides evidence for a "NO" answer. On this basis, PMCd implements a decision making process (neural competition) selecting a "YES" or "NO" response mimicking the decision to press one of the two response buttons of the experiments with humans (further details are given below). Table 1 summarises the main features of the neural maps used in the model, whereas Table 2 summarises the main features and functions of the connections between the maps.

2.5. Model functioning

The robot tests, mimicking the mental rotation laboratory experiment, were performed both with the iCub Simulator and

Table 1
Key features of the neural maps of the model.

Map	Area	Encoding	Number of neurons
POCi	POC	Current object mental image	32×32
POCl, POCs, POCr	POC	Possible rotated object image	32×32
POCp	POC	Predicted object image	32×32
PFCt	PFC	Target object image	32×32
PFCtm	PFC	Target/object match	32×32
PPCp	PPC	Proprioception of wrist	6
PPCt	PPC	Target wrist orientation	6
PPCc	PPC	Target-actual orientation combination	6×6
PMCm	PMC	Planned wrist movement	3
PMCd	PMC	Decision "YES/NO"	2

Table 2
Key features of the connections of the model.

Connection	Type	Weights values	Function	Areas
C1, C2, C3	All-to-all	Trained with delta rule	Mental image rotation	POC
C4, C5, C6, C7	One-to-one	1.0	Relays of information	POC
C8	One-to-one	1.0	Mental image for matching	POC $\rightarrow$ PFC
C9	One-to-one	1.0	Target image for matching	POC
C10	One-to-one	1.0/num. of target image pixels	Size of matching	PFC
C11	One-to-two	-10.0	Inhibition of decision	PFC–PMC
C12	One-to-one	1.0	Evidence for YES answer	PFC–PMC
C13	One-to-one	-1.0	Evidence against YES (NO)	PFC–PMC
C14	All-to-all	-10.0	Inhibition of decision	PMC
C15	One-to-one	-0.5	Dynamic competition	PMC
C16	All-to-all	Trained with Kohonen	Orientation $\rightarrow$ proprioception	POC $\rightarrow$ PPC
C17	All-to-all	Trained with Kohonen	Target $\rightarrow$ desired wrist angle	PFC$\rightarrow$ PPC
C18	All-to-all	{0.0, 1.0}	Info on predicted angle	PPC
C19	All-to-all	{0.0, 1.0}	Info on desired angle	PPC
C20	All-to-all	{0.0, 1.0}	Action selection	POC $\rightarrow$ PMC
CN1, CN2, CN3	All-to-all	1.0	Selection of rotation	PMC $\rightarrow$ POC

the actual iCub robot. The robot training, mimicking learning processes undergone by participants before the mental rotation experiment, was performed on the basis of the processing of simple images not involving the robot. The robot tests that used the actual iCub robot were more challenging than those performed with the iCub Simulator as they involved noisy/distorted images captured with the robot cameras, camera movements, and wrist movements generating information on proprioception based on the robot encoders. This is shown in Section 3 by comparing the performance of the model from both the simulated and real robot tests.

Each trial of the robot test is divided in succeeding time steps. At each step the robot performs a mental rotation, perceiving the image of the target object and of the rotated object in the computer screen via its left eye (camera). To this purpose, the eye gaze (centre of the camera) is first focused on the target object, and then on the rotated object, with hardwired movements. The two images are red-colour filtered and the edges of each object are extracted with the Canny edge detection technique (Canny, 1986; OpenCV library). The output from the edge detection process is used to activate the input units of VC. The target-object image activates the PFCt map that stores the target image as a working memory, then triggering the activation of PPCt, which encodes the current target (desired) wrist orientation corresponding to the current target object orientation.

Within POC, in the first mental rotation step POCi activates POCs corresponding to an anticipated image after a planned rotation of 0° (corresponding to no rotation/still object). From the second mental rotation onward, POCi activates one of the POCl, POCs, POCr maps depending on the planned movement (PMCp) corresponding to a possible wrist rotation of respectively +30°, 0°, −30°.

Within PPC, PPCt is activated by PFCt encoding the target image of the mental rotation. This causes an activation of PPCt representing the desired posture of the wrist corresponding to the target object orientation. Similarly, PPCp is activated by the predicted object image in POCp representing the wrist posture

corresponding to the imagined object. In detail, units of PPCt (and similarly PPCp) have a Gaussian activation computed as follows:

$$PPCpA_j = \sum w_{ji} POCp_i \tag{1}$$
$$I_{win} = \max (PPCpA)$$
$$PPCp_j = G \left(\text{dist} \left(I_{win}, I_j \right) \right)$$

where $PPCpA_j$ is the activation potential of unit j of PPCp, $POCp_i$ is the activation of unit i of POCp, w_{ji} is the connection weight between the two units, max is a function returning the index I_{win} of the unit of PPCp with maximum activation (winning unit), dist is a function computing the distance (in the neural space) between the winning unit and a given unit with index Ij.

PPCc combines the signal from PPCp, encoding the wrist posture corresponding to the current orientation of the rotated object with the signal from PPCt, encoding the target (desired) wrist orientation, to suitably select a movement within PMCm. This integration is performed with a signal multiplication as it is done in the gain-field models to capture what happens in parietal cortex (Pouget et al., 2003; Pouget & Sejnowski, 1997). In gain-field models, the output map is obtained in a hardwired fashion as the neurons of the map, here PPCc, encode all possible combinations of the input vectors, here PPCt and POCp, so that learning is not needed. In brain, neural maturation mechanisms and refining local learning processes could lead to form such combinatorial encoding. To obtain this combinatorial activation of PPCc, each unit in PPCt is connected to all units in the corresponding unit row of PPCc, whereas each unit in PPCp is connected to all units of the corresponding unit column in PPCc. The activation of a PPCc unit is then obtained by multiplying the activation of the two input signals (see Fig. 4; all the connections of C18 and C19 are equal to 1 when present; see Salinas & Abbott, 1996, for a neural-network implementation using only standard additive neural operations to obtain gain-field effects as those obtainable, as here, with multiplication). Here PPCp and PPCt are one-dimensional

population codes, encoding the object orientation in terms of wrist posture, and PPCc is a two dimensional population code encoding the combination of information from PPCp and PPCt. In detail:

$$PPCc_{ji} = PPCt_j * PPCp_i. \tag{2}$$

We now focus on PMC. PMCm units receive the following activation from PPCc:

$$AM_k = \sum \left(w_{kji} * PPCc_{ji} \right) \tag{3}$$

where AM_k is the activity of units Ml, Ms and Mr of PMCm and w_{kji} are the connection weights from PPCc to PMCm. The connection weights w_{kji} have a particular configuration: the 6 main-diagonal units of PPCc are equal to one only towards unit Ms in PMCm, and zero otherwise; the 15 units at the top right of the main-diagonal are equal to one only towards unit Mr, and zero otherwise; the 15 units at the bottom left are equal to one only towards unit Ml, and zero otherwise. This connectivity biases the activation of a certain unit, and hence the selection of a certain movement, within PMCm. In real participants the motor skills corresponding to these connections might form in early life on the basis of reinforcement learning processes (Barto & Sutton, 1998) not simulated here for simplicity. PMCm units are activated on the basis of a softmax function, ensuring an output that sums up to one for the three units, and hence can be interpreted as a probability and can be used to randomly select one of the three corresponding actions (Ml, Ms, Mr):

$$p(M_k) = \exp(AM_k/T) \Big/ \sum \exp(AM_q/T) \tag{4}$$

where $p(M_k)$ is the probability of selecting action M_k, $\exp(\cdot)$ is the exponential function, and T is the temperature parameter of the softmax function (set to 0.1) that regulates the sharpness of the selection. The stochastic process of action selection used here abstracts the winner-take-all action selection mechanisms possibly implemented in basal ganglia-motor cortex loops (Alexander et al., 1986; Baldassarre, Caligiore et al., 2013) used in several models of these loops (Doya, 2000).

The signal from PMCm is used as input to POC to select the suitable direction of the mental rotation, through the selection of either POCl, POCs, or POCr predictions. To this purpose, the units of these maps activate only if they receive an input from the corresponding units of both PMCm and POCi (neurally implemented with a summation of the two signals and a threshold of 1.5). The selection of the activation pattern of one of these maps causes the prediction of a rotated image in POCp. In the next rotation step, the rotated image in POCp is fed back to POCi and to PPCp (and hence PPCc and PMCm) to cause the next mental rotation, thus implementing a repeated reverberation of information through POC–PPC–PMC, implementing the visual and motor mental rotation processes. Reverberation mechanisms similar to these and pivoting on *forward models* (here implemented by the POCi–POCl/s/r neural networks) have been extensively used as a proxy to represent planning based on mental imaging in bioinspired computational models (Baldassarre, 2003; Butz, Sigaud & Gérard, 2003; Grush, 2004; Ziemke, Jirenhed, & Hesslow, 2005).

We now focus again on PMC, in particular on the decision making process implemented in PFC/PMCd. This process allows the model to decide if any internally generated image of the rotated object does or does not match the target object image. To this purpose, PFCt acts as a working memory, storing the target object image from VC. PFCtm is formed by units that activate only when the units of the mentally-rotated object image in POCp match the units of the target object image encoded in PFCt. To this purpose, PFCt and POCp are connected to PFCtm through one-to-one connection weights and only the units of PFCtm receiving an activation above a threshold of 1.5 are activated.

The activation of the units of PFCtm is summed up (and normalised with the activation of PFCt) by a PFC unit encoding the size of the matching between the imaged object and the target object. In turn this unit excites the unit of PMCd representing a YES reply action (PMCd_yes) and inhibits the unit of PMCd representing a NO reply action (PMCd_no has a fixed bias of 0.85). The units of PMCd, forming a *reciprocal inhibition model* of decision making, implement a neural dynamic competition based on the following differential equations (Bogacz et al., 2006; Usher & McClelland, 2001):

$$\tau PMCd_{yes} = -k PMCd_{yes} - w PMCd_{no} + PFC_{tm} \\ - i(Ml + Mr) - i(PFC_{tm} < th) \tag{5}$$

$$\tau PMCd_{no} = -k PMCd_{no} - w PMCd_{yes} + (0.85 - PFC_{tm}) \\ - i(Ml + Mr) - i(PFC_{tm} < th) \tag{6}$$

where τ ($\tau = 0.2$) is a time constant, k ($k = 0.2$) is a decay rate of PMCd units (PMCd_yes, PMCd_no), w ($w = 0.2$) is the inhibitory connection between the two PMCd units (the equations consider the activations of the PMCd units at the previous time step for this purpose), PFCtm represents the overall activation of the PFCtm map, computed as the sum of the activities of its units divided by the sum of the activity of the units of PFCt (so, the value of this variable ranges in [0, 1]), Ml + Mr is an inhibition from the movement units of PMCm, implying that a decision is not made if the system is still mentally rotating the object, and (PFCtm < th) is a variable equal to 1 when PFCtm < th, and equal to 0 otherwise. This ensures an inhibition of the decision making process when the matching is below a threshold th (th = 0.15; $i = -10$). The differential equations were integrated with the Euler method with a time-step equal to 0.02.

For each mental rotation step, this dynamic competition makes ten cycles that allow the units of PMCd to accumulate evidence for the YES or NO reply, and to compete between them, for each mental rotation step. When one of the two units achieves the threshold of 1.0, the system is considered to have made a decision in favour of the YES/NO reply corresponding to it. The time needed to solve this competition, measured from the beginning of the mental rotation trial, is considered as the reaction time taken by the system to perform the mental rotation and to make a decision (cf. Bogacz et al., 2006; Caligiore et al., 2010; Caligiore et al., 2008; Erlhagen & Schöner, 2002; Usher & McClelland, 2001). The maximum number of mental image rotation steps for trial was set to 20 in all simulations: if the model did not give an answer within this time window, it was forced to give a random answer (which happened rarely). The accumulation of evidence lasted the whole trial, i.e. the units of PMCd were reset at the beginning of each trial, but not during it. This allowed the system to accumulate evidence for a NO reply during the whole trial, and to rapidly accumulate evidence for a YES reply in the presence of a large overlap of the rotated object with the target object. The balancing of these two processes was an important element to produce a correct behaviour of the model.

2.6. Model learning

The model underwent two learning processes before being tested. These processes allowed the system to respectively acquire the core capacity needed to predict/imagine the visual appearance of objects after a step of mental rotation (forward models: connections C1, C2, C3), and to associate to a certain object image the corresponding object orientation encoded, in an embodied fashion, in terms of corresponding wrist orientation (connections C16 and C17). These learning processes are intended to capture the processes of acquisition of the general capability of rotating objects that human participants acquired during life, before

undergoing the psychological experiments on mental rotation. The first learning process, involving the connections C1, C2, and C3, allowed the system to form the forward models need to mentally rotate images. The second learning process, involving the connections C16 and C17, allowed the system to associate object images with the corresponding object orientations, encoded by proprioception in terms of wrist posture. This process captures the acquisition of associations between proprioceptive and visual information reaching PPC. For simplicity, we did not perform the two training processes with the robot directly rotating objects with a hand. Indeed, the use of actually rotating objects on one side would have posed difficult artificial vision problems (as the hand would have occluded the object sight) going beyond the scope of this work, and on the other side would have not produced advantages in terms of explanatory power of the model with respect to the target mental rotation experiments.

We now consider the two learning processes in detail. The training of the forward models needed to mentally rotate objects (connections C1, C2, C3) was done on the basis of images each formed by 3 dot-points randomly located in the image. This was done because preliminary experiments showed that training the robot with standard objects (e.g., as those used to during tests) was computationally very demanding and progressively converged to abilities as those acquired by the robot with the three-dot simpler images. This choice was also supported by the result of previous models showing that transformations in space can be generalised over the specific objects involved (Terekhov & O'Regan, 2013). The images were projected on a computer screen and perceived by the robot through its camera. Each forward model was thus trained by presenting 3-dot patterns as input, then by rotating them through the rotation angle corresponding to the forward model, and then by presenting the resulting 3-dot pattern as desired output to it. Based on pilot experiments we could see that this procedure ensured a fast training of the forward models. Figs. 5 and 6 illustrate the training strategy in detail. The filled dots shown in Fig. 5 represent the image encoded in POCi whereas empty dots represent the desired image that POCl/s/r units should encode after one rotation step. This latter image was obtained by using "cvWarpAffine", an image transformation function of the OpenCV library. As shown in Fig. 5, the images of the three random dots were formed by 32×32 pixels: the three dots were drawn within a smaller 23×23 pixel area of the image to ensure that after rotation possible dots drawn on corner locations of this 23×23 pixel area were located within the 32×32 pixel image.

During training, PMCm selects a rotation action at random, and the selected rotation then denotes which of the forward models C1, C2 and C3 to train. The three-dot images (Fig. 5, black dots) were encoded in POCi as an input pattern for the forward model, and the predicted rotated three-dot images (Fig. 5, white dots) were used to denote the desired output of POCl/s/r. Training was based on the delta rule:

$$\Delta w_{ji} = \eta \left(d_j - y_j \right) x_i \qquad (7)$$

where Δw_{ji} is the change of the connection weight between units i and j, η is a learning rate ($\eta = 0.1$), d_j is the desired activation of output unit j, y_j is the actual activation of output unit j, and x_i is the activation of input unit i.

We now focus on the training of connections C16 and C17. This training led the system to develop suitable associations between the object images and their corresponding wrist posture. The training was based on an abstraction of "motor babbling" (Caligiore et al., 2008), a process where random movements are exploited to acquire motor competencies. In particular, the robot was assumed to hold in hand an object and to perform random wrist rotations with it. The shape of the object was projected on a computer screen, perceived by the robot through its camera, and represented in
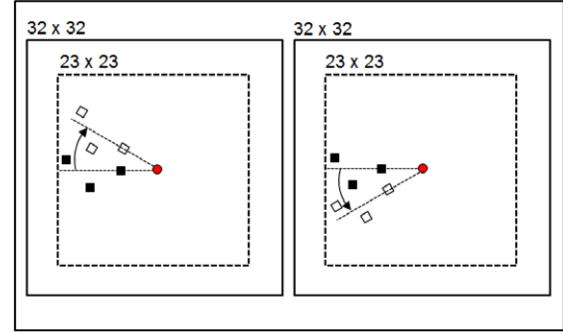
Fig. 5. Illustration of the procedure used to train the forward models of the system. The left panel illustrates the creation of a 30° clock-wise rotated image related to three random dots, whereas the right panel shows the effect of the opposite −30° rotation. Arrows and dashed lines indicate the direction of rotation: the three full dots represent the image to be rotated whereas the three empty dots represent the resulting rotated (predicted) image. Small circles represent the centre of rotation. The three random dots were generated within the dashed square areas (23×23 pixels) to keep the image within the larger square image area (32×32 pixels) after rotation, as the random dots could be drawn on the corner locations of the 23×23 area.
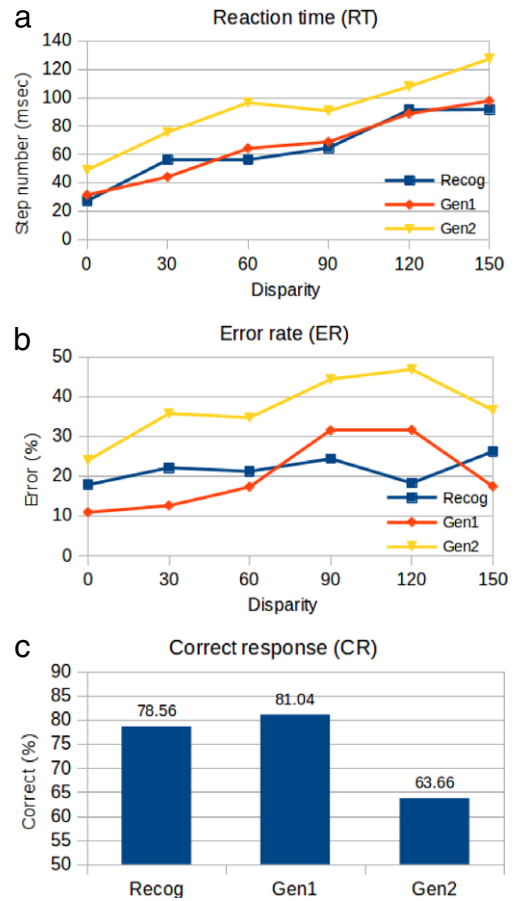
Fig. 6. Performance of the simulated robot during the three mental rotation tests: (a) Reaction time (RT); (b) Error rate (ER); (c) Correct response (CR).

POCp and PFCt. As mentioned above, we used the clean images of the objects, without the occlusion of the robot hand. In Section 3 we will also present the results of a simulated version of the experiment where the images used by the model will be those sent to the monitor. These images represent a "perfect vision" and allow measurement of the effect, on the robot performance, of the noise and distortions of the images projected on the real screen when perceived through the real robot camera.

To activate PPCt and PPCp, necessary for the training, we performed the wrist rotations only once, recorded the related wrist postures, and used these measures during the whole training process. This was performed to lessen the training time, as the rotation during the whole training on one side would have taken very long and on the other side was not necessary as, given the discrete number of possible wrist postures, we could assume a stable activation of POCt and POCp. In other words, the encoder continuous value could be mapped onto few discrete values needed to activate POCt and POCp. To ensure the equivalence of this procedure to the actual use of the real robot wrist rotation, the results reported in Section 3 on the effect of proprioception on mental rotation were performed with the real robot actually rotating its hand.

A supervised learning rule was used to train the connections C16 and C17. This rule exploits some of the mechanisms of the (unsupervised) Kohonen learning rule (Kohonen, 2001) but is used in a supervised-learning fashion to ensure the acquisition of the link between images and their corresponding wrist orientations. The idea is that, during learning, the available proprioception can furnish a suitable desired output of the mapping. Specifically, instead of selecting the unit with the highest activation potential as the winning unit of PPCp (and PPCt), as prescribed by the unsupervised Kohonen learning rule and as done during the model functioning (see Eq. (1)), we selected as the winning unit the unit encoding the particular wrist posture corresponding to the current object image. Once the winning unit was established, the activation of the PPCp units was performed as in Eq. (1) (population code). Based on this activation, we updated the connection weights C16 (similarly C17) as follows:

$$\Delta w_{ji} = \eta \text{PPCp}_j \left(\text{POCp}_i - w_{ji} \right) \tag{8}$$

$$w_{ji} = w_{ji} \Big/ \sum w_{jq} \tag{9}$$

where Δw_{ji} is the weight update, η is the learning rate set to 0.1, PPCp_j is the Gaussian activation of output unit j, POCp_i is the activation of the input unit i, and w_{ji} is the current weight value. Eq. (8) ensures that the connection weights reaching PPCp (or PPCt) units that are highly activated by the current posture become progressively associated with the active units of POCc (or PFCt) encoding the current object image. Eq. (9) normalises the connection weights after each update. After learning, when an object image with a certain orientation is perceived, it tends to cause a concentrated activation of a few PPCp units encoding the wrist posture corresponding to it; instead, when a more difficult object is perceived the activation tends to be more equally distributed ("flat") over all units.

3. Results and discussion

This section illustrates the results of the training and testing of the model. All results reported here refer to averages from data obtained by training and testing the model ten times (for each training and test, the trained connection weights of the model were assigned small random initial values). This might be considered equivalent to testing ten simulated participants with the mental rotation tasks. Three types of data were recorded during the mental rotation tests (data were also averaged over the trials corresponding to different object-pair images): (1) the response times (RT): this is the number of steps used by the decision making process to trigger a YES/NO answer: RT were measured for each disparity angle between the target and the rotated objects (recall that for each mental rotation cycle the neural competition underlying the decision making process performs ten cycles); (2) the error rates of the answers (ER): this is the number of times the

model gives a wrong answer (i.e. it replies YES in correspondence to different objects or NO in correspondence to same objects), measured for each disparity angle between the target and the rotated objects; (3) the percent of correct responses (CR), averaged over all disparity angles.

Note that at this stage of the model development we aimed at reproducing the behavioural target data from the experiments with human participants only qualitatively. This, together with the knowledge on brain areas (see Section 1.1), allowed us to impose constraints on, and hence guide, the construction and progressive improvement of the architecture, functioning and learning mechanisms of the model. Aiming to reproduce the data quantitatively would have required a large number of experiments with the robot in order to suitably tune the model parameters: this was practically very difficult and would have produced little additional knowledge at this stage of model development.

3.1. Mental rotation and generalisation

Fig. 6 illustrates the results of the tests mimicking the mental rotation experiment with the simulated robot. The figure illustrates the RT and the ER in correspondence to different disparity angles between the target and the rotated objects, and the CR, for the three tests Recog, Gen1, and Gen2. The figure shows that, for all tests, the RT and ER increases with the stimuli disparity. These results qualitatively agree with data obtained from experiments with humans (Shepard & Metzler, 1971; Wexler et al., 1998).

Fig. 6(a) and (b) show that the Recog and Gen1 conditions lead to similar RT and ER, implying that the model has a good generalisation capability. This capability relies on the capacity of the forward models of the model to mentally rotate different possible objects even when they have not been previously experienced, and represents an advancement with respect to previous models that could only rotate objects with which they were trained (Seepanomwan et al., 2013a, 2013b).

Fig. 6(c) even shows that the CR of the condition Gen1 is better not only in comparison to Gen 2 (t-test, mean of Gen1 = 81.04, mean of Gen2 = 63.66, $p < 0.001$) but also in comparison to Recog (t-test, mean of Recog = 78.56, mean of Gen1 = 81.04, $p < 0.001$). Recog is also better than Gen2 (t-test, mean of Recog = 78.56, mean of Gen2 = 63.66, $p < 0.001$). The reason for which Gen1 is even better than Recog is possibly that the objects used in Gen1 contain a clearer information on orientation: this leads to a clearer proprioceptive simulation within PPC and benefits the matching decision making process.

The greater RT for Gen2 with respect to Recog and Gen1 for most disparity values, including the zero disparity value not requiring mental rotation, is due to a smaller difference between the two inputs to the "YES" and "NO" neurons in PMCd. This smaller difference leads to a slower accumulation of evidence in favour of one of the two answers. This also leads to a higher ER. A direct inspection of the behaviour of the model revealed the nature of the errors with Gen 2. In some cases, the model gives an answer even when the two images of the target and rotated object do not have the same orientation. This can happen when the number of overlapping units in PFCtm is high enough to make the activity level of one of the two units in PMCd exceed the threshold. In other words, in these cases two images are very similar, notwithstanding they have a different orientation, and therefore trigger a false recognition. On the other hand, in some other (more rare) cases the model is able to correctly rotate the object to match it with the target but the number of overlapping neurons in PFCtm or PFCfm is small. In these cases, the images are not similar enough leading to a lack of recognition as the model arrives to the trial time-out, causing a random, and possibly incorrect answer (this
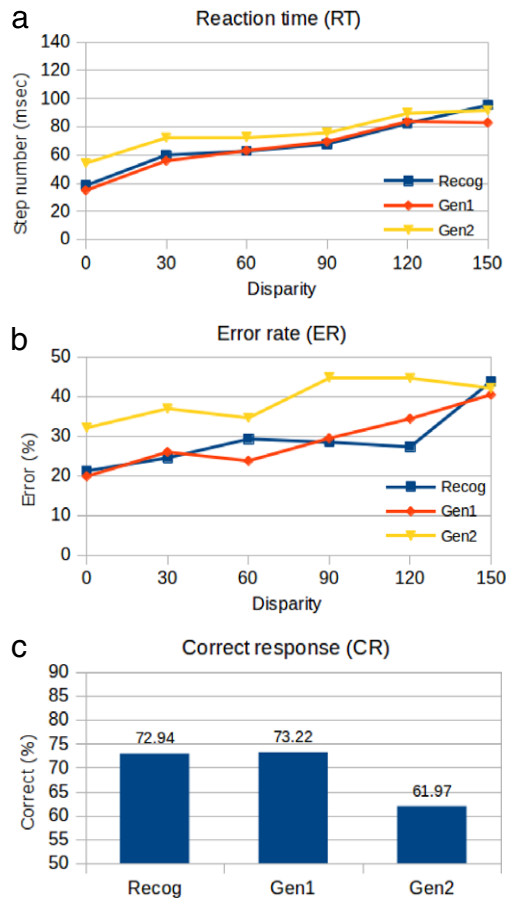
Fig. 7. Performance of the model in the real robot (iCub) during the three mental rotation tests: (a) Reaction times (RT). (b) Error rates (ER). (c) Correct response (CR).

may represent cases where the experiment participants produce an answer although they are not sure of their choice).

Fig. 7 shows the results of a similar reproduction of the mental rotation experiment obtained with the real robot. The results are qualitatively similar to those obtained in simulation, as illustrated above, with the exception that now the slight increase in performance (CR) of Gen1 with respect to Recog is no more statistically significant (t-test, mean of Recog $= 72.94$, mean of Gen1 $= 73.22$, $p = 0.615$). Recog and Gen1 conditions are confirmed to be better than Gen2 (t-test, mean of Recog $= 72.94$, mean of Gen2 $= 61.97$, $p < 001$; mean of Gen1 $= 73.22$, mean of Gen2 $= 61.97$, $p = 0.001$). The performance of the model measured as CR is higher in the simulated tests than in those carried out using the real robot (t-test: mean of simulated Recog $= 78.56$, mean of real Recog $= 72.94$, $p < 0.001$; mean of simulated Gen1 $= 81.04$, mean of real Gen1 $= 72.94$, $p < 0.001$; mean of simulated Gen2 $= 63.66$, mean of real Gen2 $= 61.97$, $p < 0.001$). This is due to the challenges posed by the real robot tests using noisy/distorted images from the real robot camera.

Overall, these results show that the model was able to generalise the mental rotation ability to the never seen objects of Gen1 and Gen2 image sets. This ability was acquired on the basis of the Recog images used to train the image-wrist posture mappings (C16 and C17 connections), and the simple three-dot images used to train the forward models (C1, C2, and C3 connections). The generalisation capability of the forward models to rotate any object (here represented with edges) ultimately relies on the important fact for which translations and rotations in space are transformations that are independent of the features of the specific object being translated or rotated (Terekhov & O'Regan, 2013). The results of the tests with Gen1 and Gen2 specify and quantify the

generalisation capabilities of the system. With Gen1, whose novel objects have many distinct features and a clear orientation axis, the model is able to rotate and compare the rotated and target objects in a good way, thus achieving RT and ER similar to those of the training set (Recog). Instead, the objects of Gen2 are much more difficult to rotate and match as they have many features matching the target image even when the orientation of the object and target differs. The system thus performs more erratic rotations, resulting in longer RT, and several matching and decision making errors, resulting in higher ER (we will see specific examples of these errors in Section 3.3). Last, the results confirm the higher challenges posed by the tests carried out with the actual robot vs. those carried out in simulation.

3.2. The role of overt movements during mental rotation and a prediction of the model

The model was also used to investigate the possible effects of performing overt movements on the mental rotation processes. In this respect, empirical data (e.g., Chu & Kita, 2011) show that if the direction of the overt movement performed during the mental rotation task is congruent with the direction of the mental rotation, participants of the experiments are facilitated to solve the task (lower RT). Vice versa, if the direction of the overt movement is opposite with respect to the direction of the mental rotation, people are hampered to solve the task (higher RT). Here we analysed, using the real robot, how the overt wrist movement affected the mental rotation processes during three tests using the Recog, Gen1, and Gen2 images. In the tests, the robot performed wrist movements that were either congruent ("match proprioception") or incongruent ("mismatch proprioception") with respect to the movement direction of the rotated object. To this purpose, we assumed that proprioception causes the activation of anterior PC proprioceptive areas in turn activating other associative areas (such as PPCp, Rizzolatti et al., 1998) and we hypothesised that the mentally simulated rotation causes the activation of the same areas. The latter hypothesis is in line with the embodied cognition framework adopted here, described in Section 1, and with the direct evidence reported in Section 1.1.

In particular, during mental rotation, the real robot wrist movement signals were recorded through the wrist encoder and mapped into one of the six possible orientations, with the resulting angle used to add an additional activation to the corresponding units of PPCp (the activation had the shape of a Gaussian over PPCp units: the Gaussian had a height equal to 1 and was centred on the PPCp unit encoding the wrist angle). In the "match proprioception" condition, if the model mentally rotated the current image to the left, the robot moved its wrist to the left. In this case the same unit in PPCp was activated by two signals encoding respectively the orientation of the mentally-rotated object, encoded in POCp, and the orientation of the wrist, based on the wrist-angle derived from the robot wrist-posture encoder signal. In the "mismatch proprioception" condition, if the model mentally rotated the current image to the left, the robot moved its wrist to the right, and vice versa. In particular, two opposite units in PPCp were activated by the two signals, the predictive and the proprioceptive signal. For example, if the mental rotation process led to activate the POCp unit number 1, the proprioceptive signal caused the activation of the opposite unit, number 6; if the mental rotation process led to activate the unit number 2, the proprioceptive signal activated the opposite unit, number 5, and so on. In line with the empirical experiments run with humans, we expected that when the movement direction of the wrist matched the direction of mental rotation, the resulting RT and ER would have decreased while CR would have increased.

The results of the tests, shown in the graphs at the left side of Fig. 8, indicate that indeed the mismatch condition led to

(a) Reaction times.

(b) Error rates.

(c) Correct response.
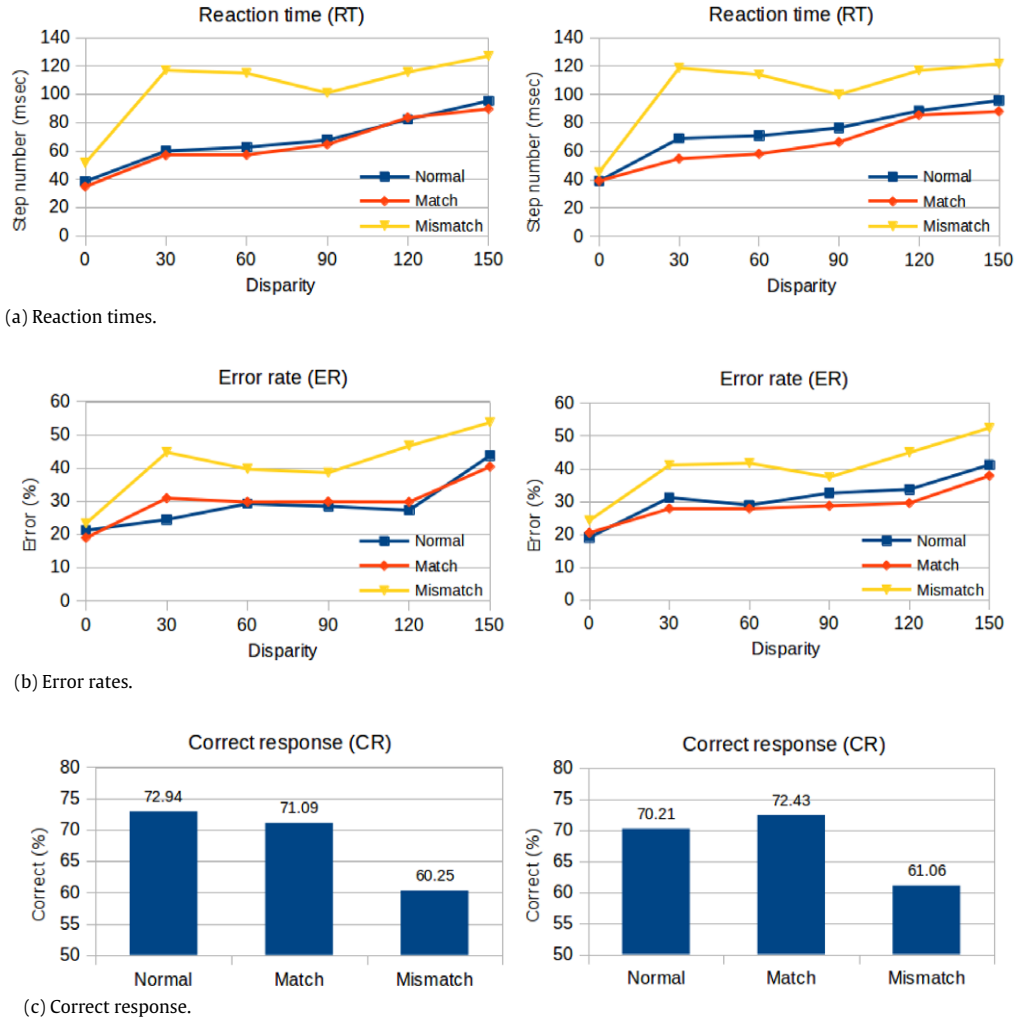
Fig. 8. Recog dataset: results when supplying the model with an additional proprioceptive input from the robot wrist that matches ("Match") or mismatches ("Mismatch") the mental rotation direction ("Normal" refers to the model without additional proprioception). Left graphs: results when adding the proprioceptive input to the "Normal" model considered in the previous sub-section. Right graphs: results when adding the proprioceptive input in a condition where noise has been added to PPCp.

deteriorate the performance of the system with respect to the baseline condition, in particular producing a longer RT (Fig. 8(a)), a higher ER for the cases with high disparity (Fig. 8(b)), and a lower CR (Fig. 8(c)). Instead, contrary to our expectation, the matching condition did not lead to a relevant benefit. A closer observation of Fig. 8(a) indicated that the latter result was due to a "ceiling effect" for which the performance of the baseline system was close to optimal, and hence could not be improved by a congruent proprioception. In particular, this graph shows that the duration of the RT for different disparity values requiring a certain number n of mental-rotation steps is only slightly above $n * 10 + 40$ (where 40 is the RT with a zero disparity angle, roughly indicating the time needed by the YES and NO units to accumulate the initial evidence needed to reply). As the neural competition process underlying the system decision making runs for 10 cycles for each mental rotation cycle, this indicates that the system performance is indeed close to being optimal.

So, why do empirical experiments show that a congruent proprioception can support mental rotation? We formulated the hypothesis for which congruent proprioception can improve mental rotation when this is made difficult by different disturbing factors, such as noisy initial images, unreliable mental rotation processes, or noisy mappings from images to proprioception. To test this hypothesis we re-ran the experiments adding noise to PPCp to capture in an abstract way such disturbances. In particular,

we added a flat noise ranging in $[-0.5, +0.5]$ to each unit of PPCp at each rotation step (the activation of the units was cut within $[0, 1]$ to keep them in the usual range). The expectation of these further tests was that noise added to PPCp would have deteriorated the performance with respect to the baseline condition, and that in this case a congruent proprioception could indeed improve the mental rotation process.

Fig. 8 displays the results of the tests without noise (the condition considered so far) and with noise. The graphs show that noise increases RT and ER in the normal condition while not affecting them much in the mismatch and match conditions. Moreover, now congruent proprioception improves the mental rotation process in terms of CR (t-test: Normal mean $=70.21$, Match mean $=72.43$, $p < 0.01$; Normal mean $=70.21$, Mismatch mean $=61.06$, $p < 0.001$), thus confirming our hypothesis.

To test the robustness of the results on the effects of proprioception, we ran again the tests just described (noise added to PPCp) with Gen1 and Gen2 datasets. The results, shown in Fig. 9, confirm the overall effects found with the Recog dataset. In particular, for the Gen1 dataset, a congruent and incongruent proprioception tends to respectively improve and deteriorate the performance of the model in terms of RT, ER and CR (t-test: Normal mean $=67.48$, Match mean $=75.25$, $p < 0.01$; Normal mean $=67.48$, Mismatch mean $=56.13$, $p < 0.001$).

(a) Reaction times.

(b) Error rates.

(c) Correct response.
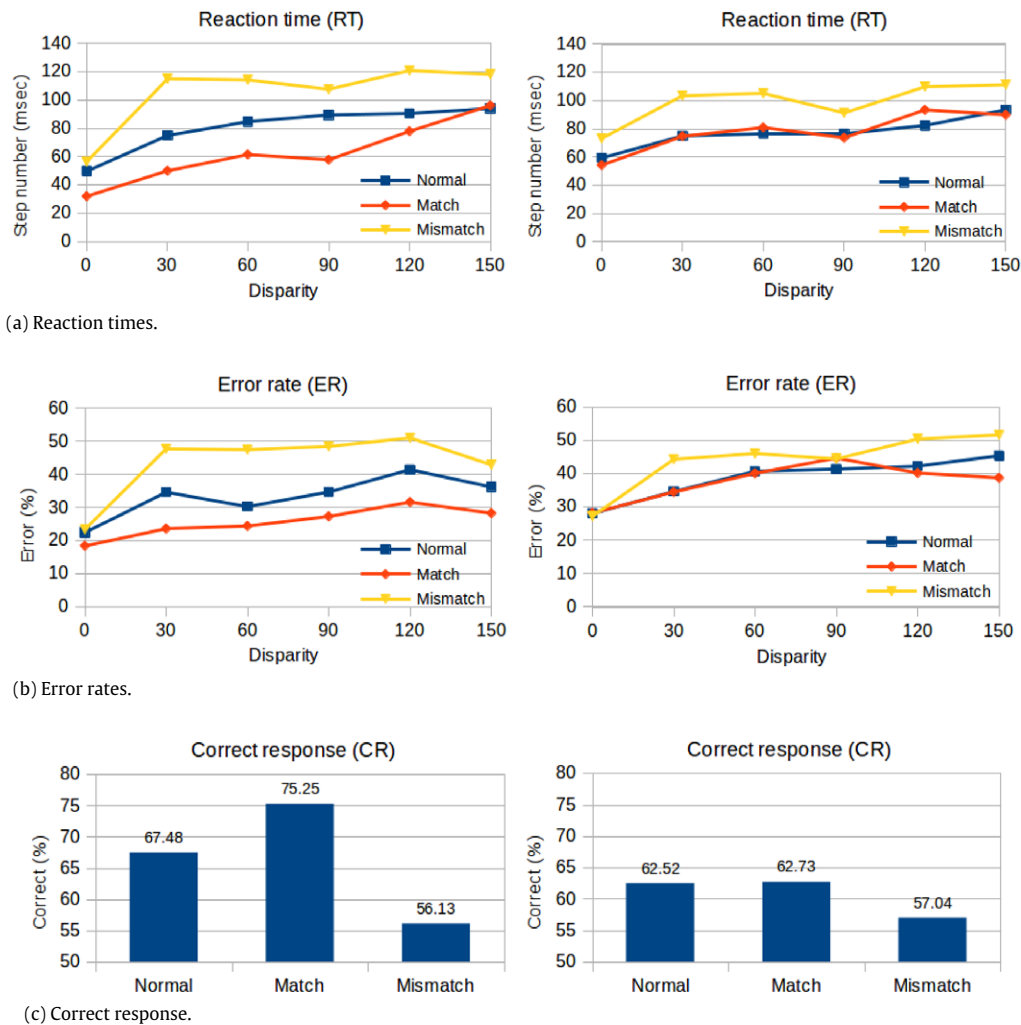
Fig. 9. Behaviour of the model when noise has been added to PPCp and the model receives an additional proprioceptive input that is congruent ("Match") or incongruent ("Mismatch") with the mental rotation direction ("Normal" refers to the model without additional proprioception). Left graphs: Gen1 object images. Right graphs: Gen2 object images.

For Gen2, only incongruent proprioception has a (detrimental) effect: this is present in terms of RT, ER and CR (*t*-test: Normal mean = 62.52, Match mean = 62.73, p = 7554; Normal mean = 62.52, Mismatch mean = 57.04, $p < 0.001$). The reason why proprioception does not help in Gen2 is likely that most difficulties encountered by the model are not related to mentally rotating the objects, but to the discrimination of the matching/mismatching points of the rotated object with the target (see Section 3.3 for further details).

The result for which *a proprioception coherent with mental rotation improves it only in cases where mental rotation is made difficult by noise sources* represents, to the best of our knowledge, a prediction of the model. This prediction could be tested in future psychological experiments by tuning the difficulty of the mental rotation task (e.g. by affecting the object images with noise).

3.3. Analysis of the internal functionality of the model

This sub-section presents some analyses of the internal functioning of the model that produced the performance illustrated in the previous sections. Fig. 10 illustrates the activation of key areas of the model when it perceives sample images, containing different target and rotated object couples, drawn from the Recog, Gen1 and Gen2 datasets. Fig. 10(a) shows a trial where the model gives an answer for an image of the Recog dataset using five steps of mental rotation. The graphs of the figure allow the visualisation of key aspects of the functioning of the model during mental rotation. PFCt encodes the edges of the target object image (in this example it has a horizontal axis). POCi encodes the input to the forward models: at step 1 this corresponds to the image of the rotated object. POCp represents the predicted image after the mental rotation (recall that at step 1 this corresponds to a no-rotation movement, so the predicted image is as the one of POCi). PPCc encodes a combination of the desired wrist posture corresponding to the target object orientation and of the wrist posture corresponding to the rotated object: this combination is the basis to trigger the proper movement at the level of PMCm. In the example, at step 1 PMCm encodes a left (anti-clockwise) mental rotation. At step 2 this mental rotation results in a predicted image of the "rotated object" (POCp) now actually rotated anti-clockwise by 30°. Notice the effects of the following rotations (PMCm) on the mental image of the system encoded in POCp, e.g. at step 2 the model does not rotate the object as it should. While the model is performing these mental rotations, PFCtm computes the matching elements of the mental image (POCp) with the target (PFCt). Notice how from step 1 to step 4 PFCtm involves a similar number of active units, and so the system decision making process (PMCd) does not produce any response. When at step 5 the PFCtm reveals a substantial matching, the evidence in support of the YES reply accumulates, and once the decision threshold is surpassed the related action is triggered (see the activation of PMCd).

Fig. 10. Activation of different areas of the model, indicated in the top row of each graph panel, while the system mentally rotates objects from the three datasets: (a) Recog. (b) Gen1. (c) Gen2.

Fig. 10(b) shows an example of mental rotation of an image taken from Gen1. The mental-rotation process and the final decision made by the system are fully correct based on the capacity of the system to properly rotate the object (see the sequence of states of POCp during the mental rotation).

Fig. 10(c) reports an example with an image from Gen2. In this case, the image is much more challenging, and after some attempts to mentally rotate the object the system sees a strong resemblance between the rotated object and the target image, and so produces a "YES" reply before successfully rotating the object (in this case the decision is correct by chance).

Fig. 11 reports the visualisation of some key areas of the system in the case of mental rotation, utilising the same object considered in Fig. 10(a), but this time with the addiction of proprioceptive information to PPCp that is congruent or incongruent with the mental rotation process. Fig. 11(a) shows how when proprioceptive information is congruent it can help the system to more reliably perform mental rotation, so taking four steps (Fig. 11(a)) rather than five (Fig. 10(a)) to successfully accomplish the mental rotation. This better performance is due to the fact that through congruent proprioception the system performs a better mental simulation of the object/wrist rotation (comparing the activation of PPCc in the two cases). Instead, Fig. 11(b) shows how an incongruent proprioception leads the system to accomplish the full rotation in a less efficient way (six steps) as the mental rotation processes of the model are more erratic.

4. Conclusions

This work has presented a novel neuro-robotic model to study the neural mechanisms possibly underlying mental rotation in humans. The model presents some innovations with respect to previous models that further refine the current hypotheses on such mechanisms. First, starting from the approach followed in Caligiore et al. (2010), the model macro-architecture was constrained with knowledge on the areas of the human brain involved in mental rotation obtained with brain imaging studies and other neuroscientific studies suggesting the mechanisms operating within them. In this respect, the model was based on a more accurate analysis of the involved brain areas. This led to the isolation of four key brain areas forming the mental rotation system and to propose four hypotheses on the key processes taking place within them. The first two areas involve sensory associations. The first area of these, the parieto-occipital cortex, is proposed to perform the mental manipulation of visual representations of objects under the influence of information on planned rotation actions received from motor areas. These processes rely on forward models that allow the anticipation of the rotated image that would result from an actual rotation of a concrete object. The second area, the posterior-parietal cortex, is involved in implementing the mapping between the object images and the corresponding proprioception of the limb possibly holding it (e.g., to compute the wrist orientation corresponding to a certain orientation of the seen objects), and to combine target postures with current postures to decide the next mental rotation to perform. The third and fourth areas involve frontal preparatory motor and executive cortex. In particular, the third area, the premotor cortex, implements the preparation of possible rotation movements that are then not executed with limbs, but are used to drive internal mentally-imaged rotations. The fourth and last area, the inferior lateral prefrontal cortex, supervises the whole process by remembering the target object orientation, by monitoring the success/failure of the mental rotation process, and finally by triggering the final response of the system in concert with the premotor cortex. Note that, for simplicity, for each of these processes this description refers to particular brain areas but the model has shown how all such processes are actually

based on a close interplay of a system of brain areas. For example, the mental rotation process pivots on parietal–occipital cortex but relies on a fundamental sensorimotor loop involving also posterior-parietal and premotor areas. Similarly, the decision making process pivots on the prefrontal cortex acting in concert with the posterior-parietal and premotor areas.

The architecture we proposed and its functioning mechanisms represent a further step with respect to previous computational models (e.g., Inui & Ashizawa, 2011; Sasama et al., 2009) as these focused on mental rotation mechanisms without relating them to the other supporting processes such as matching processes and decision making processes (Lamm et al., 2007). The architecture also represents an innovation with respect to previous neuro-robotic models (Seepanomwan et al., 2013a, 2013b) that did not distinguish between the brain areas possibly performing visual and proprioceptive processes and also used abstract monitoring and decision making mechanisms.

A second innovation of the model in comparison to other neuro-robotic models (Seepanomwan et al., 2013a, 2013b), shared with other neural-network models (e.g., Sasama et al., 2009), involves a more general rotation process capable of rotating different, possibly novel objects (to the condition that these are represented in terms of edges). This resembles the generalisation capabilities of humans as shown by the classic mental rotation experiments using unusual, novel object images (Hochberg & Gellman, 1977). In this respect, the model has shown that, at least for the type of 2D images used here to test the model, the training set can be formed by very simple images (e.g., sets of dots) as these are sufficient to allow the model to capture the spatial transformations needed to perform the mental rotations. This agrees with the proposal that spatial transformations are independent of the objects involved (Terekhov & O'Regan, 2013). The model also indicates that mental rotation of novel objects is easier when these involve few distinctive features, whereas it might incur longer reaction times and higher error rates with objects having several matching features while rotated, as this causes problems for the matching and decision making processes.

A third innovation of the model with respect to previous models is represented by the mechanism used to monitor the overall mental rotation process and to make the decision about the response to produce. To this purpose, the model incorporated the mutual inhibition model (Bogacz et al., 2006; Usher & McClelland, 2001) that allows a more accurate and biologically-plausible reproduction of the decision making process of the participants of target psychological experiments. This allowed the model to reproduce the key findings of experiments on mental rotation showing increasing reaction times and error rates in relation to increasing disparities of the orientation angles of the rotated and the target objects, whereas previous robotic models on mental rotation reproduced less consistent reaction times and could not reproduce error rates (see Seepanomwan et al., 2013a, 2013b).

Lastly, the use of the iCub robot platform shows that the model scales to real-world testing scenarios and hardware, e.g. is robust with respect to noise caused by the use of real images and real noisy wrist movements, although this implies some decrease of performance. Moreover, it allowed the performance of experiments where the information from the robot proprioception (wrist angle) was added to the mentally simulated proprioception, thus allowing the proposal of a hypothesis and the reproduction of the possible mechanisms underlying psychological experiments where participants perform movements while mentally rotating objects. This led us to show that overt movements congruent with the performed mental rotation are useful only when mental rotation processes are made difficult by uncertain images and image-proprioception matching or by other sources of noise. To the best of our knowledge, this represents a novel prediction of the

Mental rotation step	PFCt Target	POCi Forward model's input	POCp predicted object	PPCc (= desired * current posture)	PMCm Planned rotation	PFCtm (target matching)	PMCd (YES, NO decision)
1							
2							
3							
4							

(a) Recog, congruent proprioception.

Mental rotation step	PFCt Target	POCi Forward model's input	POCp predicted object	PPCc (= desired * current posture)	PMCm Planned rotation	PFCtm (target matching)	PMCd (YES, NO decision)
1							
2							
3							
4							
5							
6							

(b) Recog, incongruent proprioception.

Fig. 11. Activation of key areas of the model when it is supplied proprioceptive information in the case of mental rotation of the same object used in Fig. 10(a). (a) Case where proprioception is congruent with mental rotation. (b) Case where proprioception is incongruent with mental rotation.

model that could be tested within empirical experiments, e.g. using images having different levels of noise: the prediction would be that congruent overt movements improve mental rotation with an intensity increasing with the increase of the level of noise. Notice how the model allowed the study of these phenomena as its mental rotation processes are strongly embodied, i.e. they rely on the same mechanisms underlying sensory and motor processes (Borghi, Scorolli, Caligiore, Baldassarre, & Tummolini, 2013; Clark, 1997; Wilson, 2002). This facilitates the integration of mental and sensorimotor processes and information.

Although the model solves technologically rather simple tasks, the fact that it is embodied in a real agent makes it relevant for robotics. Mental rotation can be seen as an instance of planning and as such it could help to improve a robotic performance (Baldassarre, 2003; Latombe, 1991; Lozano-Perez, 1987). Among planning

problems, mental rotation is peculiar in that it involves only two possible actions, i.e. clock-wise and anti-clockwise rotations (at least when 2D images are considered). Moreover, the transformations that it requires are independent of the objects being rotated (the same holds for translations, Terekhov & O'Regan, 2013). As shown with the model, these two features of mental rotation allow the acquisition of general forward models to support planning processes that in principle can work with any type of object. Moreover, it also allows a mechanism for action selection (i.e., the mechanism deciding where to rotate the object) based on the relation between the rotated object and the target object, similar to cue-based planning strategies (Trullier, Wiener, Berthoz, & Meyer, 1997). In our model, this mechanism relied on the abstraction and integration of information related to the rotated and target objects, processed by encoding their orientations in terms of corresponding wrist proprioception. Mechanisms as simple and general as these might be

used to inspire other planning strategies to solve manipulation problems involving a low number of actions, e.g. not only rotations (Ciancio, Zollo, Baldassarre, Caligiore, & Guglielmelli, 2013; Meola et al., in press) but also linear translations in open space and 3D mental rotations.

The model has some limitations that represent important challenges for future work. First, the information flow between the system components is in part managed by non-neural mechanisms. This involves the cyclic flow of information from the sensory/proprioceptive components to the motor components and vice versa, needed to implement sequences of mental rotation steps. Although this process is commonly used in neural systems to implement planning (e.g., see Butz et al., 2003; Grush, 2004; Ziemke et al., 2005), it is not biologically plausible as the information flows are not managed by neural-like mechanisms (Baldassarre, 2003). To our knowledge, how repeated cycles of planning could be managed by dynamical neural systems is an important, difficult, still open problem (but note that the human brain might implement planning on the basis of rather different mechanisms; see Baldassarre, Mannella et al., 2013, for a biologically-plausible model implementing "one-step planning", i.e. triggering one action based on its goal activation). Second, the mental rotation process of the model now takes place at a coarse time and space granularity: does the human brain instead perform mental rotations in a continuous space and time? If so, this is an open problem for future models on mental rotation. Third, at the moment the training of the system forward models, and the correspondences between the perceived images and the related wrist orientations, is based on simple images: this process might instead rely on the robot actually rotating objects with the wrist and on the corresponding cluttered images so produced. Fourth, we proposed a mechanism explaining how overt movements affect mental rotation that directly relies on the actual body motion and its effects on mental simulation via proprioception. However, an alternative explanation should be investigated in the future: overt movements encoded in premotor cortex might directly interfere with the simulated movements encoded in the same area. Fifth, the focus of this work was on the system-level integration of the multiple processes involved in mental rotation, so we hardwired or simplified several processes taking place within the single neural maps and components of the model, in particular those within posterior parietal cortex (PPC) combining the desired and mentally simulated wrist posture. Now that we have formulated an overall hypothesis on how the mental simulation processes work together, further work might introduce specific learning processes operating within those areas, e.g. to investigate the formation of neural receptive fields encoding input information in more interesting ways than those used here (Ognibene, Rega, & Baldassarre, 2006). Sixth, the model considered only cortical areas of the brain, whereas, as mentioned in the introduction, mental simulation also involves subcortical areas (e.g. basal-ganglia are involved in the selection of actions and goals and this can be explicitly modelled, Baldassarre, 2002; Baldassarre, Mannella et al., 2013; Mannella, Gurney, & Baldassarre, 2013). Last, the model attention processes are currently hardwired, wherein they could be guided by bottom-up mechanisms based on the features of the image and top-down mechanisms related to the goals of the agent (Ognibene & Baldassare, 2015). Notwithstanding these limitations, the model proposed gives a relevant contribution to the current state-of-the-art with respect to the formulation of an integrated set of hypotheses on the neural mechanisms and processes that may underlie mental rotation in humans.

Acknowledgements

References

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9(1), 357–381.

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, 25, 189–220.

Baldassarre, G. (2001). *Planning with neural networks and reinforcement learning*. University of Essex, Department of Coputer Science.

Baldassarre, G. (2002). A modular neural-network model of the basal ganglia's role in learning and selecting motor behaviours. *Cognitive Systems Research*, 3(1), 5–13.

Baldassarre, G. (2003). Forward and bidirectional planning based on reinforcement learning and neural networks in a simulated robot. In M. Butz, O. Sigaud, & P. Gérard (Eds.), *Anticipatory behavior in adaptive learning systems* (pp. 179–200). Berlin: Springer.

Baldassarre, G., Caligiore, D., & Mannella, F. (2013). The hierarchical organisation of cortical and basal-ganglia systems: a computationally-informed review and integrated hypothesis. In G. Baldassarre, & M. Mirolli (Eds.), *Computational and robotic models of the hierarchical organization of behavior* (pp. 237–270). Berlin, Heidelberg: Springer.

Baldassarre, G., Mannella, F., Fiore, V. G., Redgrave, P., Gurney, K., & Mirolli, M. (2013). Intrinsically motivated action-outcome learning and goal-based action recall: a system-level bio-constrained computational model. *Neural Networks*, 41, 168–187.

Barto, A. G., & Sutton, R. (1998). *Reinforcement learning: an introduction*. MIT Press.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends in Cognitive Sciences*, 11(3), 118–125.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113, 700.

Borghi, A. M., & Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia*, 48(3), 763–773.

Borghi, A. M., Scorolli, C., Caligiore, D., Baldassarre, G., & Tummolini, L. (2013). The embodied mind extended: using words as social tools. *Frontiers in Psychology*, 1(4), 214.

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception*, 30(1), 61–72.

Butz, M. V., Sigaud, O., & Gérard, P. (Eds.) (2003). *Anticipatory behavior in adaptive learning systems: foundations, theories, and systems*. Vol. 2684. Berlin: Springer.

Caligiore, D., Borghi, A. M., Parisi, D., & Baldassarre, G. (2010). TRoPICALS: A computational embodied neuroscience model of compatibility effects. *Psychological Review*, 117, 1188–1228.

Caligiore, D., Borghi, A., Parisi, D., Ellis, R., Cangelosi, A., & Baldassarre, G. (2013). How affordances associated with a distractor object affect compatibility effects: A study with the computational model TRoPICALS. *Psychological Research*, 77(1), 7–19.

Caligiore, D., Ferrauto, T., Parisi, D., Accornero, N., Capozza, M., & Baldassarre, G. (2008). Using motor babbling and Hebb rules for modeling the development of reaching with obstacles and grasping. In R. Dillmann, C. Maloney, G. Sandini, T. Asfour, G. Cheng, G. Metta, & A. Ude (Eds.), *International conference on cognitive systems (CogSys2008)* (pp. E1–E8). Karlsruhe, Germany: University of Karlsruhe.

Caligiore, D., Parisi, D., & Baldassarre, G. (2014). Integrating reinforcement learning, equilibrium points and minimum variance to understand the development of reaching: A computational model. *Psychological Review*, 121, 389–421.

Caligiore, D., Pezzulo, G., Miall, R. C., & Baldassarre, G. (2013). The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neuroscience & Biobehavioral Reviews*, 37(10), 2504–2515.

Cangelosi, A., & Schlesinger, M. (2015). *Developmental robotics*. Boston, MA: MIT Press.

Canny, F. J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., & Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience*, 11(1), 9–24.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, 140(1), 102.

Ciancio, A. L., Zollo, L., Baldassarre, G., Caligiore, D., & Guglielmelli, E. (2013). The role of learning and kinematic features in dexterous manipulation: a comparative study with two robotic hands. *International Journal of Advanced Robotic Systems*, 10, e1–21.

Clark, A. (1997). *Being there: putting brain, body, and world together again*. Boston, MA: MIT Press.

Cohen, M. S., & Bookheimer, S. Y. (1994). Localization of brain function using magnetic resonance imaging. *Trends Neurosciences*, 17, 268–276.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22(1), 319–349.

Corballis, M. C., & McLaren, R. (1982). Interaction between perceived and imagined rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 215.

Di Nuovo, A., De La Cruz, V. M., & Marocco, D. (2013). Special issue on artificial mental imagery in cognitive systems and robotics. *Adaptive Behavior*, 21(4), 217–221.

Di Nuovo, A. G., Marocco, D., Cangelosi, A., De La Cruz, V. M., & Di Nuovo, S. (2012). Mental practice and verbal instructions execution: A cognitive robotics study. In *The 2012 international joint conference on neural networks. IJCNN2012*.

Dissanayake, M. G., Newman, P., Clark, S., Durrant-Whyte, H. F., & Csorba, M. (2001). A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3), 229–241.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10(6), 732–739.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, 109, 545.

Fukumi, M., Omatu, S., & Nishikawa, Y. (1997). Rotation-invariant neural pattern recognition system estimating a rotation angle. *IEEE Transactions on Neural Networks*, 8(3), 568–581.

Fukumi, M., Omatu, S., Takeda, F., & Kosaka, T. (1992). Rotation-invariant neural pattern recognition system with application to coin recognition. *IEEE Transactions on Neural Networks*, 3(2), 272–279.

Georgopoulos, A. P., Lurito, J. T., Petrides, M., Schwartz, A. B., & Massey, J. T. (1989). Mental rotation of the neuronal population vector. *Science*, 243, 234–236.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination. *Experimental Brain Research*, 112(2), 103–111.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3), 377–396.

Harris, I. M., & Miniussi, C. (2003). Parietal lobe contribution to mental rotation demonstrated with rTMS. *Journal of Cognitive Neuroscience*, 15, 315–323.

Hochberg, J., & Gellman, L. (1977). The effect of landmark features on mental rotation times. *Memory Cognition*, 5, 23–26.

Hubel, D. H. (1988). *Eye, brain and vision*. New York, NY: Scientific American Books.

Inui, T., & Ashizawa, M. (2011). Temporo-parietal network model for 3D mental rotation. In R. Wang, & F. Gu (Eds.), *Advances in cognitive neurodynamics (II)* (pp. 91–95). Netherlands: Springer.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: The cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, 18(7), 314–320.

Johnston, S., Leek, E. C., Atherton, C., Thacker, N., & Jackson, A. (2004). Functional contribution of medial premotor cortex to visuo-spatial transformation in humans. *Neuroscience Letters*, 355, 209–212.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *International Journal of Robotics Research*, 5(1), 90–98.

Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Berlin: Springer.

Kosslyn, S. M. (1996). *Image and the brain. The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., & Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology*, 35(2), 151–161.

Kulkarni, A., Yap, A. C., & Byars, P. (1990). Neural networks for invariant object recognition. In *Proceedings of the 1990 symposium on applied computing*.

Lamm, C., Windischberger, C., Moser, E., & Bauer, H. (2007). The functional role of the dorso-lateral premotor cortex during mental rotation An event- related fMRI study separating cognitive processing steps using a novel task paradigm. *NeuroImage*, 36(4), 1374–1386.

Latombe, J.-C. (1991). *The springer international series in engineering and computer science: Vol. 124. Robot motion planning*. Berlin: Springer.

Lozano-Perez, T. (1987). A simple motion-planning algorithm for general robot manipulators. *IEEE Journal of Robotics and Automation*, 3(3), 224–238.

Mannella, F., Gurney, K., & Baldassarre, G. (2013). The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Frontiers in Behavioral Neuroscience*, 7, e1–29.

Masehian, E., & Sedighizadeh, D. (2007). Classic and heuristic approaches in robot motion planning—a chronological review. *World Academy of Science, Engineering and Technology*, 29(1), 101–106.

Meola, V.C., Caligiore, D., Sperati, V., Zollo, L., Ciancio, A.L., Taffoni, F., Guglielmelli, E., & Baldassarre, G. Interplay of discrete and rhythmic manipulation movements during development: A policy-search reinforcement-learning robot model. *IEEE Transactions on Autonomous Mental Development* (in press).

Metta, G., Sandini, G., Vernon, D., Natale, L., & Nori, F. (2008). The iCub humanoid robot: an open platform for research in embodied cognition. In *Proceedings of the 8th workshop on performance metrics for intelligent systems*.

Meyer, J.-A., & Filliat, D. (2003). Map-based navigation in mobile robots—II. A review of map-learning and path-planning strategies. *Cognitive Systems Research*, 4(4), 283–317.

Michelon, P., Vettel, J. M., & Zacks, J. M. (2006). Lateral somatotopic organization during imagined and prepared movements. *Journal of Neurophysiology*, 95, 811–822.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, 31(2–3), 236–250.

Ognibene, D., & Baldassare, G. (2015). Ecological active vision: four bio-inspired principles to integrate bottom-up and adaptive top-down attention tested with a simple camera-arm robot. *IEEE Transactions on Autonomous Mental Development*, 7(1), 3–25.

Ognibene, D., Rega, A., & Baldassarre, G. (2006). A model of reaching that integrates reinforcement learning and population encoding of postures. In S. Nolfi, G. Baldassarre, R. Calabretta, J. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, & D. Parisi (Eds.), *Lecture notes in artificial intelligence: 4095. From animals to animats 9: proceedings of the ninth international conference on the simulation of adaptive behavior*, (SAB2006), (pp. 381–393). Berlin: Springer Verlag.

Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annual Review of Neuroscience*, 26(1), 381–410.

Pouget, A., & Sejnowski, T. J. (1997). Spatial transformations in the parietal cortex using basis functions. *Journal of Cognitive Neuroscience*, 9(2), 222–237.

Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R. S., Gati, J. S., … Kim, S.-G. (2000). Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *Journal of Cognitive Neuroscience*, 12(2), 310–320.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2), 131–141.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalography and Clinical Neurophysiology*, 106, 283–296.

Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1), 23–38.

Salinas, E., & Abbott, L. (1996). A model of multiplicative neural responses in parietal cortex. *Proceedings of the National Academy of Sciences*, 93(21), 11956–11961.

Sandini, G., Metta, G., & Vernon, D. (2007). The icub cognitive humanoid robot: An open-system research platform for enactive cognition. In *50 years of artificial intelligence* (pp. 358–369). Berlin, Heidelberg: Springer.

Sasama, T., Mitsumoto, H., Yoneda, K., & Tamura, S. (2009). Mental rotation by neural network. In *Fifth international conference on intelligent information hiding and multimedia signal processing, IIH-MSP2009*.

Seepanomwan, K., Caligiore, D., Baldassarre, G., & Cangelosi, A. (2013a). A cognitive robotic model of mental rotation. In *IEEE symposium on computational intelligence, cognitive algorithms, mind, and brain. CCMB2013*.

Seepanomwan, K., Caligiore, D., Baldassarre, G., & Cangelosi, A. (2013b). Modelling mental rotation in cognitive robots. *Adaptive Behavior*, 21(4), 299–312.

Shepard, R., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171(972), 701–703.

Snyder, L. H., Grieve, K. L., Brotchie, P., & Andersen, R. A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, 394, 887–891.

Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., & Frackowiak, R. S. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology*, 73(1), 373–386.

Terekhov, A.V., & O'Regan, K.J. (2013). Space as an invention of biological organisms. arXiv Preprint no. 1308.2124, pp. 1–15.

Tikhanoff, V., Cangelosi, A., Fitzpatrick, P., Metta, G., Natale, L., & Nori, F. (2008). An open-source simulator for cognitive robotics research: the prototype of the iCub humanoid robot simulator. In *Proceedings of the 8th workshop on performance metrics for intelligent systems*.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J.-A. (1997). Biologically based artificial navigation systems: review and prospects. *Progress in Neurobiology*, 51(5), 483–544.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, 8, 769–800.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, 108, 550–592.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, 68(1), 77–94.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9(4), 625–636.

Wohlschläger, A. (2001). Mental object rotation and the planning of hand movements. *Perception & Psychophysics*, 63, 709–718.

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 24(2), 397–412.

Zacks, J. (2008). Neuroimaging studies of mental rotation: a meta-analysis and review. *Journal of Cognitive Neuroscience*, 20(1), 1–19.

Ziemke, T., Jirenhed, D.-A., & Hesslow, G. (2005). Internal simulation of perception: a minimal neuro-robotic model. *Neurocomputing*, 68(0), 85–104.