UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
An Embodied Developmental Robotic Model of Interactions between Numbers and Space

Permalink
https://escholarship.org/uc/item/1rn4b6f0

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Rucinski, Marek
Cangelosi, Angelo
Belpaeme, Tony

Publication Date
2011

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

An Embodied Developmental Robotic Model of Interactions between Numbers and Space

Marek Ruciński (marek.rucinski@plymouth.ac.uk)
Angelo Cangelosi (angelo.cangelosi@plymouth.ac.uk)
Tony Belpaeme (tony.belpaeme@plymouth.ac.uk)
Centre for Robotics and Neural Systems, University of Plymouth
Plymouth, Devon, PL48AA, UK

Abstract

In this paper we describe an embodied developmental model of the interactions between the neural representations of numbers and space in the humanoid robot iCub. We show how a simple developmental process that mimics real-world cultural biases leads to the emergence of certain properties of the number and space representation system that enable the robot to reproduce well-known experimental phenomena. We demonstrate the validity of the proposed approach by showing that it leads to the reproduction of three psychological phenomena connected with number processing, namely size and distance effects, the SNARC effect and the Posner-SNARC effect.
Keywords: mathematical cognition; developmental cognitive robotics; computational modeling; size effect; distance effect; SNARC effect; Posner-SNARC effect;

Introduction

Perceiving numbers and quantities is one of the most basic perceptual skills of humans and animals (Dehaene, 1997). Given the pure and abstract character of the number concept as perceived by humans, it is no surprise that many in cognitive science pursue a better understanding of how such a peculiar concept could have emerged, how it is represented and processed, and how it relates to other processes that take place in the brain. These efforts, which can be put together under the common label of mathematical cognition, constitute a branch of science that has been gaining more and more momentum during the past few decades (Dehaene & Brannon, 2010).

Computational modeling is an important tool used in the study of mathematical cognition to understand the principles of number processing in the brain. Based on observations from experimental psychological studies as well as hints obtained through various brain imaging techniques, computer models of number representation and processing are constructed and evaluated on the basis of how well their properties match those of the biological cognitive systems. Analysis of the computer models helps us to understand how biological systems work at the algorithmic level, which in turn is necessary to understand their neural implementations.

In this paper we present an embodied developmental cognitive robotic model of interactions between number and space. In the following paragraphs we provide a short review of previous computational models of numerosity representation and processing, focusing on those most relevant to the work presented herein.

An influential connectionist model of number representation and processing has been described by Dehaene and Changeux (1993). The system consisted of a 1-dimensional visual retina, a location and normalization cluster, a summation coding layer and a place coding layer. The output from the place coding layer was used in a “same-different” comparison and “larger-smaller” comparison tasks. The system was designed to model perception and processing of non-symbolic stimuli (e.g. a cardinality of a set of items perceived visually). One of the most interesting findings was the demonstration of how the described system can autonomously “discover” the larger-smaller relation based solely on unsupervised experimentation with addition and subtraction.

One of the first models of number representation based on recurrent artificial neural networks was proposed by Rodriguez, Wiles, and Elman (1999). Here, supervised learning techniques were used to teach a simple recurrent neural network to perform a task, in which counting was required in order to succeed. In successful networks, neurons formed a special case of a discrete-time dynamical system in which numerosity was coded in the dynamical properties of trajectories realized by hidden layer units. This complex solution, radically different from traditionally used coding methods, has been obtained despite a small amount of inductive bias in the training process.

Ahmad, Casey, and Bale (2002) presented a rather complex system aimed at modeling two manifestations of numerical abilities: subitizing and counting. Their system consisted of two networks, each specifically designed to perform one of these tasks, composed of several modules playing different roles and trained separately using various machine learning techniques. Implementation of the model delivered interesting results especially in the domain of counting (which, being a more complex task with a temporal structure, has been more rarely tackled in the literature than instant comprehension of numerosity), where counting error patterns similar to those observed in children were obtained.

A relatively consistent path of increasingly complex modeling of different aspects of human mathematical cognition can be found in a series of papers by Verguts and collaborators (Verguts & Fias, 2004; Verguts, Fias, & Stevens, 2005; Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006; Chen & Verguts, 2010). The first model focused on how simple number coding methods believed to be employed in the brain (summation and place coding) can emerge as the result of an unsupervised learning process, thus showing that such systems do not have to be innate as suggested in earlier research. Building on a place-coding system with linear scaling and constant variability as the core representation of numerosity, Verguts et al. (2005) shifted the responsibility for size and distance effects from number representation to later processing stages. It was demonstrated that this leads to results consistent with experimental data. These are characterized by symmetrical priming patterns and no size effect in naming and parity tasks, combined with the presence of both size and distance effects in the comparison task. This is allegedly not possible to obtain using numerosity representations with compressed scaling and/or increasing variability, used in earlier models. An important step has been achieved by Gevers et al. (2006), where experimental phenomenon more complex than size and distance effects, namely the SNARC effect (Spacial-Numerical Association of Response Codes, (Dehaene, Bossini, & Giraux, 1993)) were modeled. The model used a dual-route architecture to explain the phenomenon, combining findings from previous computational models and other studies aiming at explaining spatial congruency effects. The simulations were compared to experimental data, predictions were made about the shape of the SNARC effect in a certain category of tasks, and these were confirmed experimentally.

Finally, the model of Gevers et al. (2006) was further extended in a recent paper by Chen and Verguts (2010), in which a representation of space was introduced instead of an "automatic pathway" present in the previous model. Chen and Verguts (2010) added a module corresponding to a "human homologue of lateral intra-parietal area in macaque monkeys", a saliency map related to the visual field, consisting of two parts characterized by contra-lateral spatial neuronal gradients. These gradients were identified as the crucial property of the model which allowed for reproduction of a number of psychological experiments, including those involving patients suffering from certain lesions.

The model we present in this paper extends the work of Chen and Verguts (2010) by addressing two drawbacks with their model. First, as it is the case with all mathematical cognition models published to-date, the system does not take directly into account any aspects of embodiment. According to current trends in cognitive science it is not possible to understand the brain in separation from the body in which it is embedded and from the environment in which it develops. In line with this, when formulating our model we considered any relevant constraints imposed by the target body (that of a humanoid robot), and designed the developmental process accordingly. Secondly, the most important phenomenon investigated by Chen and Verguts (2010), that is associations between numbers and space, have been modeled in their paper as hand-wired connections, despite extensive evidence cited by themselves that most probably it is the "environmental correlation between symbolic numbers and physical space" that creates this association in the brain. In this paper we show how necessary patterns of connections can indeed emerge from a simple developmental process.

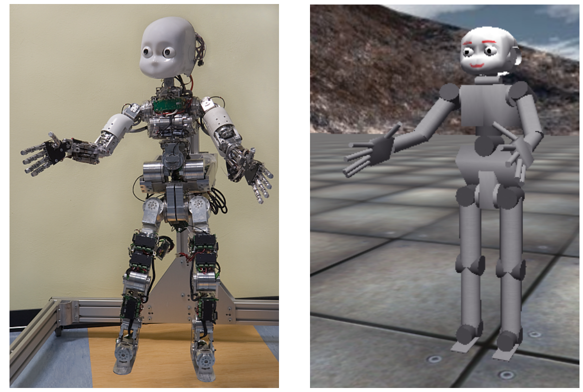

Figure 1: iCub, the humanoid robot used in modeling.

The following sections of the paper are organized as follows. First we introduce the robotic platform that has been employed in this modeling study. Then we present the architecture of our model and the process of its development. Next we demonstrate the validity of our model by showing that it is able to reproduce three phenomena in which interactions between numbers and space manifest themselves. We finish the article by drawing conclusions from the experiments and emphasizing the capability of the embodied robotic approach to be used in the modeling of mathematical cognition.

Model Description

iCub, the Humanoid Robot Platform

The model described in this paper has been designed to operate in a simulated model of the humanoid robot iCub (Metta et al., 2010). The robot itself (figure 1), is an open-source design developed recently as a benchmark platform for cognitive robotics experiments. The anatomy of the robot is intended to resemble that of a 3.5 years old human child and has a total of 53 degrees of freedom, 20 of which were used in the experiments described here (6 for head and eyes, and 7 for each of the two arms). iCub is equipped with devices which allow it for visual, auditory, tactile and proprioceptive perception. Robot software includes the iCub simulator (Tikhanoff, Cangelosi, & Metta, in press), a tool for robotic simulation experiments without the use of the physical robot. In research described in this paper only the simulated robot has been used.

Model Architecture

The architecture of the model (see figure 2) builds on results of the modeling experiments described above, as well as those of Caligiore, Borghi, Parisi, and Baldassarre (2010), where the authors formulated a general embodied model of compatibility effects focusing on motor affordances and goals. The processing of information in our model is split into two neural pathways: "ventral", responsible for processing the identity of objects as well as task-dependent decision making

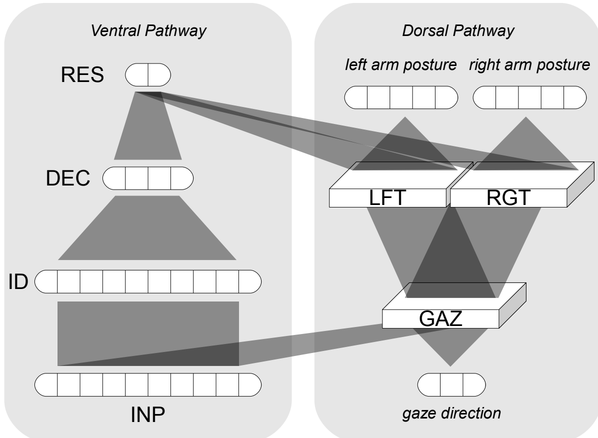

Figure 2: Architecture of the model.

and language processing, and "dorsal", involved in processing of spatial information about the locations and shapes of objects and sensorimotor transformations which provide online support for visually guided motor actions (please refer to (Caligiore et al., 2010) for an extensive discussion of motivations for such a division).

"Ventral" Pathway: Decision Making and Language Processing The "ventral" pathway is modeled in a very similar way to components of the (Chen & Verguts, 2010) model. It consists of: 1) a *symbolic input* INP which codes for the number, using place coding (same remarks about the irrelevance of the spatial arrangement of neurons as those raised in the original paper apply here); 2) a "mental number line" ID which codes for number *identity* (the meaning of the symbol) with linear scaling and constant variability; 3) a *decision layer* DEC executing each of the considered tasks, that is number comparison and parity judgment; and 4) a *response layer* RES, integrating information from both pathways and responsible for the final selection of the motor response. Similar to the practice used in (Chen & Verguts, 2010), for simplicity of implementation the actual structure of the "ventral" pathway, especially its decision layer, was adapted depending on the task to be performed, by removing components irrelevant to the task at hand. Likewise, for the number comparison task which requires more than one number to be processed at the same time, short-term memory was implemented by duplicating necessary layers (namely INP and ID). The layers were composed of the following numbers of neurons: INP and ID: 15, DEC: 4 (2 for each task), RES: 2.

"Dorsal" Pathway: Spatial Coding and Transformations The "dorsal" pathway is composed of a number of neuronal maps which code for the spatial locations of objects in the robot peripersonal working space using different frames of reference (Wang, Johnson, & Zhang, 2001): one associated with the gaze direction (GAZ), and two for each of the robot's arms: left (LFT) and right (RGT). These maps are implemented as 49-cell (7 by 7) 2-dimensional Kohonen Self-Organizing Maps (SOMs) with cells arranged in a hexagonal pattern. Input to the GAZ map comes from the 3-dimensional proprioceptive vector representing the robot gaze direction (azimuth, elevation and vergence) and input to each arm position map consists of a 7-dimensional proprioceptive vector representing the position of the relevant arm joints: shoulder pitch, roll and yaw, elbow angle and wrist pronosupination, pitch and yaw. The GAZ map is linked to both arm maps: this implements the transformation of spatial coordinates between frames of reference corresponding to these body parts (so that a position in the visual field can be translated into an arm posture corresponding to reaching to this position and vice-versa). It is important to note that this is the part of the model where the embodied approach to modeling is implemented, and where the crucial difference between our and all previous quoted models lies. This point is elaborated in the Discussion section.

Developmental Learning of the Robot

The modeling of the developmental learning process is organized around a number of sequential phases corresponding to different stages of development of a human child. First, spatial representations for sight and motor affordances have to be built and correspondences between them established. Later, the child can learn number words and their meaning. Usually in late preschool years, children learn to count. More or less at the same stage the child may be taught to perform simple numerical tasks such as number magnitude comparison or parity judgment. All these stages are reflected in our model.

Building Spatial Representations and Transformations In order to build the gaze and arm space maps, the robot performs a process equivalent to *motor babbling* (Von Hofsten, 1982), in which a child refines its internal visual and motor space representations by performing random movements with arms while observing its hands, reaching for toys in its visual field, etc. This enables the child to perform tasks such as visually guided reaching later in life. This stage of development was implemented in the robot by selecting 90 points uniformly distributed on what has been assumed to be the robot's operational space (a part of a sphere in front of the robot with 0.65m radius, centered between robot's shoulder joints, spanning $\pm 30°$ of elevation and $\pm 45°$ in azimuth). These points served as target locations for directing gaze and moving both arms of the robot using inverse kinematic modules. After a trial in which the robot reached a random position, the resulting gaze and arm postures were read from proprioceptive inputs and stored. Between each trial, the head and arms of the robot were moved to the rest position in order to eliminate any influence of the sequence in which the points have been presented on the head and arm posture at the end of the motion. These data were used to train the three SOMs using the traditional unsupervised learning algorithm. In order to reflect the asymmetry between reachable space for the left and right arm (some areas reachable by the right arm cannot be reached by the left arm and vice versa), only $2/3$ of the extreme points corresponding to an arm were used when building a spatial map for this arm (e.g. leftmost $2/3$ of all points for the left arm). Learning parameters were adjusted manually based on the observation of the learning process and analysis of how well resulting networks span target spaces.

Transformations between the visual spatial map GAZ and the maps of reachable space LFT and RGT, implemented as connections between the maps, were trained using the classical Hebbian learning rule. In a process similar to motor babbling, gaze and the appropriate arm were directed toward the same point and resulting co-activations in already developed spatial maps were used to establish links between them.

Learning Number Words and Their Meaning This stage of learning corresponds to establishing links between number words, modeled as activations in the INP layer, and number meaning, being activations in the ID layer. In the model described here links between INP and ID layers were preset manually implementing place coding with linear scaling and constant variability (as in (Chen & Verguts, 2010) and previous models). However, Verguts and Fias (2004) showed that such pattern of connections can arise from a simple supervised learning process.

Learning to Count The goal of this stage is to model the cultural biases that result in an internal association of "small" numbers with the left side of space and "large" numbers with the right side, since this is believed to be the cause of SNARC and similar effects. As an example of these biases we considered a tendency of children to count objects from left to right, which may be associated with the fact that European culture is characterized by left-to-right reading direction (Dehaene, 1997). In order to model the process of learning to count, the robot was exposed to an appropriate sequence of number words (fed to the INP layer of the model network), while at the same time robot's gaze was directed toward a specific location in space (via the input to the GAZ spatial map). These spatial locations were generated in such a way that their horizontal coordinates correlated with number magnitude (low number presented on the left, large numbers on the right) with a certain amount of Gaussian noise. Vertical coordinates were chosen to uniformly span the represented space. While the robot is exposed to this process, Hebbian learning establishes links between number word and stimuli location in the visual field.

Learning Comparison and Parity Tasks Finally, the model is trained to perform target tasks, that is number comparison and parity judgment, which corresponds to establishing appropriate links between the ID layer and neurons in the DEC layer. This process, extensively described in (Verguts et al., 2005), involves supervised learning using the Widrow-Hoff Delta learning rule after all activations in the network reach stable states. In our model we used weight values from our own reproduction of the experiments described in the cited paper.

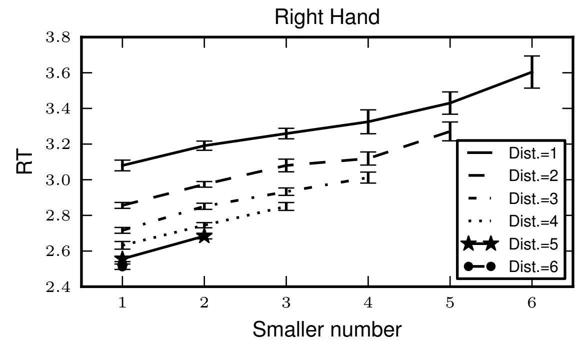

Figure 3: Simulation of the size and distance effects in the number comparison task. On all RT charts error bars show $\pm$2SEM.

Simulation Results

In order to demonstrate the validity of our model we tested it by simulating three selected tasks which have been used previously to evaluate models by other authors (Chen & Verguts, 2010). In this section we present a brief summary of the results. All of the tasks involved measuring response times (RT) of the model. These were obtained by assuming that a response is given when activity in one of the two response nodes exceeded an assumed response threshold (0.5 for experiments 1 and 2 and 0.8 for experiment 3). We report RTs aggregated over 10 independent instantiations of the model[1].

Experiment 1: Size and Distance Effects

Size and distance effects are two of the most common findings from experimental mathematical cognition studies. They are present in many tasks, but in the context of number comparison they mean that it is more difficult to compare larger numbers (size effect) and numbers which are closer to each other (distance effect). This should be evident in RTs growing with number magnitude and with decreased distance between numbers being compared. RTs obtained from simulating the experiment in our model are reported in figure 3. Response times were measured for all pairs of numbers from 1 to 7. We report results for the right hand response only (results for the left hand were similar). Clearly both size and distance effects are present in the model. Sources of the size and distance effects in our model are the same as in the model by Chen and Verguts (2010), namely monotonic and compressive patterns of weights between ID and DEC layers.

[1]In contrast to cited authors we had no access to numerical data from relevant psychological experiments, thus we were unable to perform linear regression over these data. This does not invalidate any of our results (only additional linear scaling of RTs is performed), but must be kept in mind when comparing charts from the respective papers.

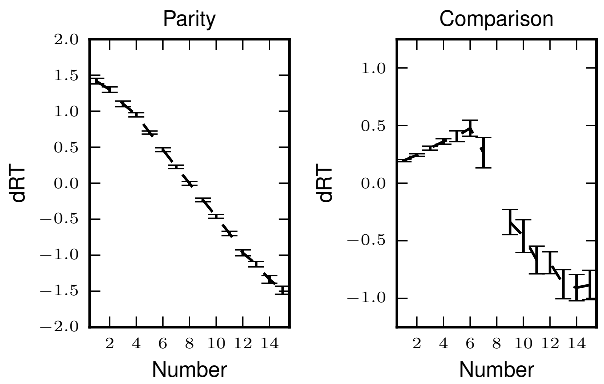

Figure 4: Simulation of the SNARC effect in parity judgment and magnitude comparison tasks.

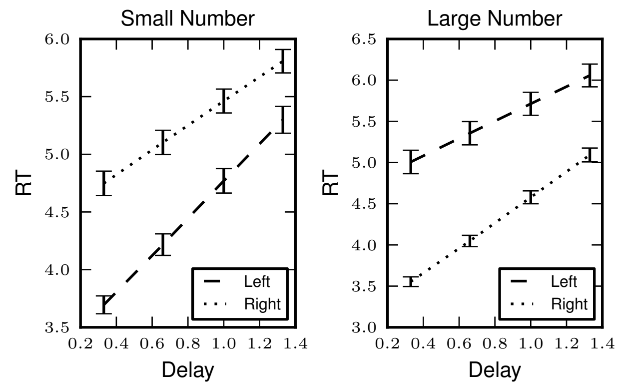

Figure 5: Simulation of the Posner-SNARC effect.

Experiment 2: SNARC effect

SNARC effect is more directly related to interactions between number and space than size and distance effects. Using a similar procedure as in (Chen & Verguts, 2010), we report RTs obtained by our model in parity judgment and number comparison tasks. Here, the difference between right hand and left hand RTs for the same number in both congruent and incongruent condition is reported. The SNARC effect should manifest itself in a negative slope on such a chart. Results of our simulations are presented in figure 4. Presence of the SNARC effect is evident in both tasks. The source of this effect in our model requires further explanation.

Quoting relevant neuroscientific research, Chen and Verguts (2010) explain sources of the SNARC effect as the result of "an initial dip toward the wrong response hand in SNARC-incongruent conditions evident in recordings of the lateralized readiness potentials in the motor cortex". Accordingly, in our model the presentation of a number word leads to an automatic activation of the relevant parts of the visual space representation, due to links established during model development (more precisely, during learning to count) – left part for small numbers, and right part for large ones. Visual space representations in turn are linked to both motor maps, although not symmetrically. As outlined above in the description of the model development, some parts of the visual space that can be reached by the right arm cannot be reached by the left arm, and vice versa. As a consequence, when transformation from the visual space map to arm maps occurs, both arm-related representations will be activated to a similar degree only for the areas in the center of the visual map. For the areas placed to the sides of the visual space, the map associated with one arm will be activated more strongly than the other, as it over-represents that side of space (this is a natural consequence of the robot morphology). Because there is a significant overlap between represented areas, the effect is not sudden, but connections between visual and motor maps form a gradient from left to right – links to the left arm map become weaker, while those to the right become stronger. Thus, for instance when a small number is presented, internal connections lead to stronger automatic activation of the representations linked with the left arm than those of the right arm, which causes the SNARC effect. In contrast to (Chen & Verguts, 2010), in our model this particular pattern of connections is not hand-wired, but emerges as a consequence of the robot morphology during the development process. We hypothesize that the presence of such neuronal gradients in the human brain referred to by (Chen & Verguts, 2010) may be ascribed to similar factors.

Experiment 3: Posner-SNARC effect

The Posner-SNARC effect is another manifestation of the connection between numbers and space, placed within the attention cuing paradigm (Fischer, Castel, Dodd, & Pratt, 2003). A small or large number presented at the fixation point acts as a cue and directs attention of the participant toward the left or right side of space, affecting the time needed to detect an object appearing in the visual field after a certain delay. The effect results in faster detection of the target on the left when a small number is presented as a cue, and on the right for large numbers, even though throughout the experiment numbers are not predictive of target locations. Simulated response times obtained from model are shown in figure 5. The effect is visible on the charts in shorter RTs for the target presented on the left for a small number as a cue, and on the right for a large number as a cue.

Discussion

In the paper we have presented an embodied developmental robotic model of interactions between numbers and space. We have described the model architecture as well as the associated developmental process. By simulating three well-known experiments we have demonstrated the validity of our approach, showing that after development is completed, our model exhibits the most important properties of the human mathematical cognition system. In this final section of the article we discuss the differences between our approach and those of authors of earlier works, thus highlighting the benefits which embodied robot simulations bring to cognitive modeling in general.

As described above, the crucial difference between our modeling approach and previous literature models is the aspect of embodiment. The robot we use in our experiments has one head and two arms, thus three separate spatial maps for each of these body parts are developed in our cognitive model. The robot proprioceptively perceives its gaze direction and arm positions using specific degrees of freedom, and as a consequence the maps in our model have to be implemented to span this specific number of dimensions. Finally, these maps are *real* spatial maps, in which activations correspond to specific positions of a material limb and vice versa. Thus such an embodied approach may greatly help to reduce arbitrariness of the model. Taking as an example the system described by Chen and Verguts (2010), "space representation" has been implemented there as an arbitrary network of connections, hand-wired in such a way so that it exhibits properties suggested by neuroscientific data. Although this allowed for a successful reproduction of a good number of experiments, the traditional connectionist approach remained inconclusive regarding how to answer the questions of *why* such a pattern of connections is present and *how* it comes into being. Supplementing the previous modeling achievements with the embodied approach and replacing previously arbitrary parts of the model with elements which have direct material interpretation allowed us to formulate hypotheses to answer these questions.

The importance of the embodied approach to cognitive modeling increases together with the level of complexity of the processes being modeled, and with the degree to which motor representations and actions are involved. In the context of mathematical cognition one may recall experiments such as physical line bisection (where the participant is asked to point at the middle of a line presented on board in front of him) or investigation of the role of finger counting habits (Fischer, 2008) or that of gesture in learning to count (Andres, Seron, & Olivier, 2007) to name just a few. While some researchers already attempted to model the former task with a purely connectionist model (Chen & Verguts, 2010), the embodied robotic approach is more suited to tackle such problems from the developmental perspective.

Results presented in this paper are part of work in progress. After connecting the model with the real iCub robot instead of its virtual equivalent, we plan to employ it to tackle issues in mathematical cognition directly involving motor representations and actions, with a special focus on the relations between gesturing and learning to count.

Acknowledgments

This research has been supported by the EU project RobotDoC (235065) from the FP7 Marie Curie Actions ITN.

References

Ahmad, K., Casey, M., & Bale, T. (2002). Connectionist simulation of quantification skills. *Connection Science*, *14*(3), 165-201.

Andres, M., Seron, X., & Olivier, E. (2007). Contribution of hand motor circuits to counting. *Journal of Cognitive Neuroscience*, *19*(4), 563-576.

Caligiore, D., Borghi, A. M., Parisi, D., & Baldassarre, G. (2010). TRoPICALS: A computational embodied neuroscience model of compatibility effects. *Psychological Review*, *117*(4), 1188 - 1228.

Chen, Q., & Verguts, T. (2010). Beyond the mental number line: A neural network model of number-space interactions. *Cognitive Psychology*, *60*(3), 218-240.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371 - 396.

Dehaene, S., & Brannon, E. M. (2010). Space, time, and number: a Kantian research program. *Trends in Cognitive Sciences*, *14*(12), 517-519.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *J. Cognitive Neuroscience*, *5*(4), 390-407.

Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex*, *44*(4), 386-392.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555-556.

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology-Human Perception and Performance*, *32*(1), 32-44.

Metta, G., Natale, L., Nori, F., Sandini, G., Vernon, D., Fadiga, L., et al. (2010). The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, *23*(8-9), 1125 - 1134.

Rodriguez, P., Wiles, J., & Elman, J. L. (1999). A recurrent neural network that learns to count. *Connection Science*, *11*(1), 5 - 40.

Tikhanoff, V., Cangelosi, A., & Metta, G. (in press). Language understanding in humanoid robots: iCub simulation experiments. *IEEE Transactions on Autonomous Mental Development*.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience*, *16*(9), 1493-1504.

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66.

Von Hofsten, C. (1982). Eye-hand coordination in the newborn. *Developmental Psychology*, *18*(3), 450-461.

Wang, H., Johnson, T. R., & Zhang, J. (2001). The mind's views of space. In *Proceedings of the Third International Conference on Cognitive Science* (pp. 191–198).