

# Information Processing in the Mirror Neuron System in Primates and Machines

Yiannis Demiris · Lisa Aziz-Zadeh · James Bonaiuto

Published online: 2 October 2013  
© Springer Science+Business Media New York 2013

**Abstract** The mirror neuron system in primates matches observations of actions with the motor representations used for their execution, and is a topic of intense research and debate in biological and computational disciplines. In robotics, models of this system have been used for enabling robots to imitate and learn how to perform tasks from human demonstrations. Yet, existing computational and robotic models of these systems are found in multiple levels of description, and although some models offer plausible explanations and testable predictions, the difference in the granularity of the experimental setups, methodologies, computational structures and selected modeled data make principled meta-analyses, common in other fields, difficult. In this paper, we adopt an interdisciplinary approach, using the BODB integrated environment in order to bring together several different but complementary computational models, by functionally decomposing them into brain operating principles (BOPs) which each capture a limited subset of the model's functionality. We then explore links from these BOPs to neuroimaging and neurophysiological data in order to pinpoint complementary and conflicting explanations and compare predictions against selected sets of neurobiological data. The results of this comparison are used to interpret mirror system neuroimaging results in terms of neural network

activity, evaluate the biological plausibility of mirror system models, and suggest new experiments that can shed light on the neural basis of mirror systems.

**Keywords** Mirror system · BODB · Computational modeling · Robotics · Imitation

## Introduction

In the past 20 years, the discovery of mirror neurons in the premotor cortex of the monkey brain has heavily influenced our understanding of how we perceive and potentially imitate the actions of others. The brain mechanisms underlying our abilities to match external observations of actions with the internal motor representations used for their execution, are central to our ability to imitate another agent's actions (Rizzolatti and Craighero 2004; Iacoboni 2005). As the mechanisms behind imitation processing are important for development, learning, and sensorimotor integration, it has become a topic of intense research and debate in biological and computational disciplines, particularly neuroimaging and robotics.

Imitation is an important topic in robotics, as the ability to imitate human demonstrations would greatly simplify the task of manually programming robots to perform various tasks. Several robotics systems have used models of the mirror system in order to imitate demonstrated tasks, however these models are highly inhomogeneous, and are typically at very different levels of analysis that neural models of this system. This makes the task of a principled meta-analysis of mirror system models extremely difficult.

In this paper, we adopt an interdisciplinary approach, using the Brain Operation DataBase (BODB) integrated environment (Arbib and Bonaiuto 2013) in order to bring together two different but complementary computational models, the MNS (Oztop and Arbib 2002; Bonaiuto et al. 2007) and HAMMER (Demiris and Hayes 2002; Demiris and Khadhour

---

Y. Demiris (✉)  
Department of Electrical and Electronic Engineering, Imperial  
College London, London, UK  
e-mail: y.demiris@imperial.ac.uk

L. Aziz-Zadeh  
Brain and Creativity Institute and Division of Occupational Science  
and Occupational Therapy, University of Southern California, Los  
Angeles, CA, USA  
e-mail: lazizzad@usc.edu

J. Bonaiuto  
Division of Biology, California Institute of Technology, Pasadena,  
CA, USA  
e-mail: bonaiuto@caltech.edu

2006) families of models. The models in the first family (FARS/MNS1/2) were designed to closely match detailed anatomical connections in the grasping process and its mirroring; they also use more biologically plausible computational representations such as recurrent neural networks for their operations. Given this emphasis these models provide close support to experimental data on the macaque brain areas related to grasping objects. The second family of models (HAMMER) was designed with the goal of enhancing the capabilities of robotic systems to observe and imitate human actions and action sequences, and although it aspires to be biologically plausible, its support basis rests more on behavioural and functional experimental data. Despite their differing starting points, and possibly surprisingly, both families of models independently draw interesting common predictions on subsets of experimental data (a notable example being the hypothesized sensitivity of mirror neurons to speed profiles of demonstrated actions), making the prospect of a meta-analysis of the models an exciting possibility.

In this paper we attempt to bring together these diverse models by functionally decomposing them into brain operating principles (BOPs) which each capture a limited subset of the model's functionality. In BODB, a BOP is a description of an "operating principle" common to a range of functions or structures (regions, circuits, neurons, etc.) of the brain. One example of a Brain Operating Principle is the Winner-Take-All (WTA) Principle, which asserts that "A Winner-Take-All network responds (after a delay) to an array of inputs of different intensity with an output array which encodes which input had the greatest intensity." BOPs provide a functional brain ontology (contrasted with structural ontologies such as brain atlas), which can then be used to analyze and compare a varied range of models. Because BOPs describe the functions performed by brain regions and not their implementation, they can also serve to describe robotics models which rely on very different implementations.

A central feature of BODB is the ability to link BOP entries to summaries of experimental data (SEDs) that provide evidence for neural circuits which perform the functionality it describes. We use this feature to explore links from the BOPs related to these two models to neuroimaging and neurophysiological data. These links are then used to generate explanations and compare predictions against neurobiological data.

## Background

### Mirror Neuron System: Experimental Data

The discovery of the mirror neuron system several decades ago was exciting in that it potentially provided a neural basis for imitation and action understanding. Researchers recording neural activity from single neurons in premotor area F5 and inferior parietal regions of macaque monkeys found that some

neurons fired both when the monkey performed an action, like reaching for a piece of food, and when the monkey watched the experimenter perform the same action (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese et al. 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). These mirror neurons were thought to match incoming visual information about another's actions to one's own motor representations, possibly allowing the observer to understand the other's action (Rizzolatti and Craighero 2004; Fadiga et al. 2009). More recent data has begun to expand upon this hypothesis. It has been found that the majority of mirror neurons are tuned to specific views of grasping actions, with a minority having view-dependent properties (Caggiano et al. 2011). Understanding another's actions in order to plan a response requires a representation of the space the actions are performed as well as the objects they are performed on. Indeed, the activity of many F5 mirror neurons is modulated by the subjective value of the grasped object (Caggiano et al. 2012), and many of them respond differently depending on the spatial relationship between the observed action and observer (Caggiano et al. 2012).

Neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), have allowed researchers to explore whether a similar mirror neuron system exists in humans. Some research supports the existence of a putative human mirror neuron system located in the inferior frontal gyrus (IFG; Brodman area 44) and the rostral inferior parietal lobule (IPL), which are the proposed human homologues to macaque mirror regions in area F5 and IPL (Rizzolatti and Craighero 2004; Gallese et al. 2004). This putative human MNS is active both during the execution of an action and the observation of the same or a similar action, comparable to that seen in macaques. As this system is involved in both action observation and execution, it is thought to be highly involved in imitation (Iacoboni 2005). In fact, there is evidence that interfering with processing in the IFG disrupts imitation (Heiser et al. 2003). Because observing others activates one's own motor systems, it is thought that the MNS is important for action and social understanding. That is, we may understand other people's actions and intentions by simulating their actions on our own motor representations (Gallese et al. 2004).

In relation to understanding the underlying mechanisms of mirror neurons, one question that is important to understand is how visual input can modulate activity in the MNS. That is, if an action is outside the motor repertoire of the observer, does it still activate the MNS? Or can visual input without a corresponding motor plan activate the MNS? There are several experiments that have explored this question. In a recent study, it was found that individuals born without hands and feet not only engaged their own MNS for actions that they could perform with a different effector (e.g., writing with a pen using their mouth), but also for actions that were impossible for them (e.g., using scissors, Aziz-Zadeh et al. 2012).

Furthermore, in a study by Cross et al. (2006), it was found that visual training alone could enhance activity in the MNS. A similar result was found from single cell recordings in the macaque, showing that repeated observation of tool use that the monkey herself was not trained to use still positively modulated activity in mirror neurons in F5 (Ferrari et al. 2005). Taken together, these studies indicate that visual input, even without a matching motor plan, can activate the MNS as long as the observed action can be associated with a known action in terms of goal or intention.

#### Mirror Neuron System: Computational Models

Models of the mirror system tend to fall into two main groups. Neural models typically focus on the response of mirror neurons during action observation, with F5 canonical neurons providing some sort of training signal to shape their activity. Most robotics systems link mirror neurons to some sort of internal model of actions such as a forward or inverse model, which is used for action observation and generation (Demiris and Hayes 2002). Various computational models of the mirror system have been reviewed in detail previously (Oztop et al. 2006; Oztop et al. 2012); the focus here is to examine the differences between these classes of models using BODB, and relate them to empirical data or predictions for future experiments.

To date there have been very few neural models of the mirror system. The Mirror Neuron System (MNS) family of models (Oztop and Arbib 2002; Bonaiuto et al. 2007) shows how the response properties of mirror neurons can be shaped by experience. In both MNS and MNS2, F5 mirror neurons acquire their response properties by learning to associate the pattern of F5 canonical activity with the visual signals generated during self-observation while producing grasps. In this sense, the F5 canonical neurons provide a training signal for F5 mirror neurons, which also receive visual input in the form of highly-preprocessed hand-object relations. Once trained, F5 mirror neurons can reproduce this pattern of activity given only the visual signals received during observation of another agent performing a grasp. Although the MNS and MNS2 models use back propagation and back-propagation-through-time, respectively, as learning algorithms, they are compatible with the Hebbian mirror neuron system (Hebbian-MNS) conceptual model (Keysers and Perrett 2004).

While the MNS models do not utilize the output of F5 mirror neurons for any motor purposes, they do allow the visual representation of an action to be mapped on to a simple motor representation. The Augmented Competitive Queuing (ACQ) model (Bonaiuto and Arbib 2010) introduced the novel hypothesis that a mirror system may contribute both to monitoring the success of a self-action and to recognition of one's own apparent actions if they deviate from one's intended actions. The model selects actions to perform by choosing the

most desirable of the currently executable actions. When the model begins to perform an action within a certain context, mirror neurons create an expectation of the sensory effects of reaching the goal of that action. If the expectation is not satisfied, then the model decreases its estimate of the action's executability—of how likely it is to succeed in the given context. But if the model fails to execute one action, it may nonetheless, in some cases, succeed in completing a movement and achieving a desirable goal (or taking a step towards such a goal). If so, the mirror system may “recognize” that the action looks like an action already in the repertoire. As a result, learning processes can increase the neural estimate of the desirability of carrying out that action when the animal attempts to achieve the goal in the given context. This model thus showed how this “what did I just do?” function of mirror neurons can contribute to the learning of both executability and desirability, and how in certain cases this can support rapid reorganization of motor programs in the face of disruptions.

While much of the latest data on mirror neuron modulation by spatial relationship and subjective object value has not been taken into account by current models, there is a model that address the data on view-dependent F5 mirror neurons (Fleischer et al. 2012). In this model, action recognition is performed using only visual mechanisms. It contains multiple networks tuned to several different visual angles. Each network is hierarchical, with neurons at the lowest level tuned to local orientation and simple shape fragments, and neurons at the highest level tuned to trajectories in hand-object position space. View-independence is achieved in this model by pooling the responses from the view-dependent output of each network.

The Dynamic Neural Field Mirror Neuron System (DNF-MNS) model combines dynamic neural fields with Hebbian learning to provide an account of goal-directed matching of action observation and action execution (Erlhagen et al. 2006). The model is directly inspired by the Hebbian-MNS model and implements the emergence of neurons with mirror properties by Hebbian learning on connections in the superior temporal sulcus (STS, Perrett et al. 1989), PF, and F5 circuit. This model does not distinguish between F5 canonical and mirror neurons. Activation of F5 neurons gives rise to immediate imitation, but the details of the performed action can be influenced by representations of the current task and goal in the prefrontal cortex.

Sausser and Billard (2006) test two models of imitation that also utilize dynamic neural fields. In the single route model spatial and motion cues as well as mirror neuron activity interact together in a single decision layer before influencing the response. The direct matching model utilizes two pathways and selection processes: one for spatial visual cue information and one comprising the mirror circuit. Both models are similar to the MNS models in that F5 mirror neurons receive

as input both an efference copy of the motor plan and visual information from STS.

Hourdakis et al. (2011) present a cortical model of imitation which is similar in architecture to the MNS models, but where F5 mirror neurons provide the training signal to F5 canonical neurons rather than vice versa. Here, F5 mirror neurons acquire their visual properties through Hebbian association of various visual input signals during action observation. In this model, F5 mirror neurons project to F5 canonical neurons, and these connections are modified via reinforcement learning upon successful imitation. Action observation thus initially gives rise to the performance of random actions (due to the random initial weights of the F5 mirror—canonical projections), which gradually come to resemble the demonstrated action as learning progresses.

Internal forward and inverse models have been hypothesized to underlie the organization of motor systems, as put forward by the MOSAIC family of models (Haruno et al. 2001). Demiris and Hayes (1999, 2002), inspired by these motor control models, proposed and implemented a robotic model that uses them in a dual role, both for recognizing and for imitating actions. Their model uses a battery of behaviors paired with forward models. Each “behavior” is a control system which takes current state and goal and provides the motor output (it is thus what an “inverse model” models); while its forward model transforms the code for the motor output (and the goal) to generate what the next state would be if the corresponding behavior were executed. The task of the architecture is to observe a particular trajectory of states of a system and determine which behavior in its own repertoire best matches it. Specifically, when the demonstrator executes a behavior, the perceived states are fed into the imitator’s available behavior modules in parallel which generate motor commands that are sent to the forward models. The forward models predict the next state based on the incoming motor commands, which are then compared with the actual demonstrator’s state at the next time step. The error signal resulting from this comparison is used to derive a confidence value for each behavior (module). The behavior with the highest confidence value (i.e. the one that best matches the demonstrator’s behavior) is selected for imitation. When an observed behavior is not in the existing repertoire, none of the existing behaviors reach a high confidence value, thus indicating that a new behavior should be added to the existing behavior set. This is achieved by extracting representative postures while the unknown behavior is demonstrated, and constructing a behavior module to go through the representative postures extracted. The architecture can be related to mirror neurons because the behavior modules are active during both movement generation and observation. However all the modules are run in parallel in the proposed architecture, so it is more reasonable to take the confidence values as representing the mirror neuron responses in this model. Demiris and Hayes

(2002) predicted that “mirror neurons that remain active for a period of time after the end of the demonstration are encoding more complex sequences that incorporate the demonstration as their first part.” One may thus view their model as a dual route model of imitation, using forward models to recognize and imitate known actions and a direct visuo-motor transformation to imitate novel actions, and add them to the repertoire as new inverse models (Demiris and Hayes 2002).

## Experimental Methodology

### Overview

To date there have been few attempts to compare mirror system models from multiple disciplines. Given their differing goals, brain models and robotics systems are typically described in very different formats. Brain models attempt to capture the workings of a brain region or subset of brain regions and therefore are typically described as a network of modules with each module corresponding to a particular region or subregion. Regions which are the main subject of investigation are typically modeled in a more-or-less biologically plausible manner, with clear relationships between some signals in the model and experimental data at some level of analysis. Robotics systems more pragmatically seek to show that a given computational architecture can capably and efficiently perform some task or set of tasks at a desired level of performance. They are also typically described in a modular fashion, reflecting the style in which they are developed, but these modules may link to one or more brain regions, may overlap multiple brain regions, or may implement functionality not thought to be present in the brain (such as computation using laser-range finding, although what may be analogous echolocation functionality exists in the brains of some species). Moreover, robotics systems are usually not concerned with implementation in a biologically plausible manner, making linkages between their internal signals and experimental measures of neural activity.

The first step in comparison between models of these two types is therefore an architectural description in a common format. From there, modules of each model’s architecture can be described according to a functional ontology of computational principles. This allows linkages to be created from these modules to summaries of experimental data that address or illustrate these same computational principles. The result is a functional decomposition of each model using a common set of building blocks and linked to a common set of experimental data. Models from varying disciplines can then be compared in this common format on the basis of the experimental data used to support them or experimental finds that are explained by or contradicted by each model.



In this paper, we describe how we use BODB to decompose the MNS family of models (FARS, MNS and MNS2) and the HAMMER model into a hierarchy of submodules, linking each submodule with the brain operating principles it implements, the experimental data used to build it, and the experimental data and corresponding simulation results used to test it. Linked BOPs provide an interface for comparing neural and robotics models in terms of functionality, while SEDs allow a comparison in terms of experimental data covered by each model.

#### Brain Operation DataBase (BODB): Operation and Principles

The Brain Operation Database (BODB) is described in more detail elsewhere (Arbib and Bonaiuto 2013), but here we give a brief overview of its principles. While BODB was designed as a database for describing computational models, the Model entries are not for storing source code. Rather, they are designed to function as a repository for structural descriptions of computational models of brain mechanisms, with fields available to provide linkages to actual implementations, simulations, documentation and descriptions. BODB is centered on the idea that a brain model should be characterized not only by a structural ontology (the brain regions or finer structures to which it corresponds) but also by a functional ontology (the Brain Operating Principles, BOPs), which it exemplifies. BOPs therefore set forth functional principles that can structure both models and observed neural function.

In addition to the brain operating principles (BOPs), in order to aid the description of complex models in a principled manner, BODB requires users to break the complex models into a hierarchy of modules with well-defined inputs and outputs. The hierarchical composition and modularity of models aids model development because one can easily reuse existing models as a part of a new model. One does not have to know the exact details underlying a module—only the *function* of such models is required (by function we mean a model behavior which maps a certain set of inputs—perhaps via dynamic internal states—to a certain set of outputs in a well-specified manner). This process is analogous to implementing a software system in an object-oriented programming manner, where we select relevant objects each with a set of inputs, outputs and its function, fitting them to the other peer modules in that system. In general we view a model as comprising a single “Top level module” hierarchically decomposed into a number of interconnected submodules, which themselves may or may not be further decomposable. If a module is decomposable, we say this “parent module” is decomposed into submodules known as its “child modules”. Otherwise, the module is a “leaf module”.

In addition to the BOPs that describe their functionality, model entries are also linked to summaries of experimental data (SEDs). SEDs are like the separate elements of review

articles, structured as entries in a database rather than as parts of a lengthy narrative. They address such high level data as assertions, summaries, hypotheses, tables, and figures that encapsulate the “state of knowledge” in a particular domain. Summaries of certain datatypes such as connectivity and brain imaging data are linked to external databases for viewing more detailed information. Connectivity SEDs are derived through federation with the CoCoMac database (Kotter 2004), which contains detailed information from tract tracing experiments done in macaques.

There are two possible relationships between models and SEDs (see Arbib et al. 2013 for further discussion of the relationships that can exist between models and SEDs). SEDs used to build the model include those that place the model in a larger context (scene setting) or are directly used to design some aspect of the model (support). SEDs used to test the model are linked to summaries of simulation results (SSRs) that describe model simulations at a level of detail relevant for comparison with experimental data. Models can link SEDs to SSRs that explain or contradict their results. Predictions made by the model, for which no experimental data exist, can be documented by linking SSRs along with a description of the experiment that could be used to verify or disprove the prediction.

BODB provides users with functionality to search for models, BOPs, and SEDs according to several criteria including related entities. This allows complex searches for models that explain a certain set of experimental data according to a specified set of brain operating principles, for example. BODB also provides model benchmarking tools that allow several models to be compared in terms of the SEDs that they link to, allowing users to identify commonalities and highlight areas for future model development. This paper is the result of using BODB’s Collaborative Workspace feature (Arbib and Bonaiuto 2013), which allows multiple users to form workspaces centered around a common interest. Workspaces allow users to share selected models and SEDs that are addressed by them and those that challenge further model development. Analysis of models in the workspace can reveal challenges for building new models by comparing them through benchmarking.

#### BODB Formalization of FARS/MNS/MNS2 and HAMMER

##### FARS

The basic planning and control system for grasping is modeled by the FARS model (named for Fagg, Arbib, Rizzolatti and Sakata; Fagg and Arbib 1998) which embeds F5 canonical neurons in a larger system. In this model, the parietal area AIP processes visual information to implement

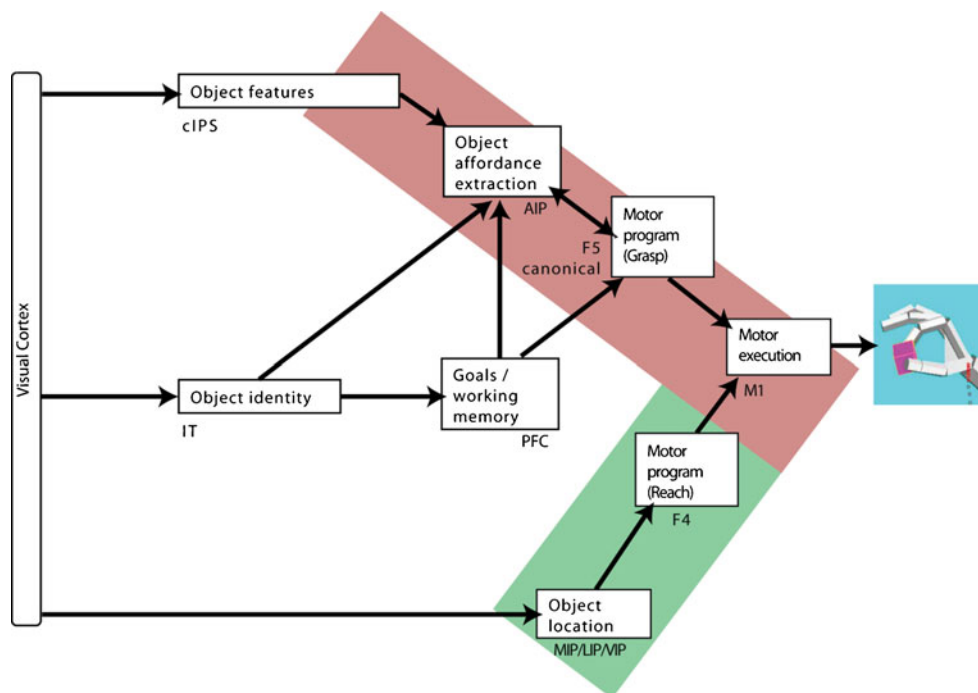
perceptual schemas for extracting grasp parameters (affordances) relevant to the control of hand movements and is reciprocally connected with the canonical neurons of F5. Primary motor cortex (F1) formulates the neural instructions for lower motor areas and motor neurons. The dorsal visual stream (i.e., which passes through AIP) does not know “what” the object is, it can only see the object as a set of possible affordances; whereas the ventral stream (from primary visual cortex to inferotemporal cortex, IT) is able to recognize what the object is. This information is passed to prefrontal cortex (PFC) which can then, on the basis of the current goals of the organism and the recognition of the nature of the object, bias the affordance appropriate to the task at hand. AIP may represent several affordances initially, but only one of these is selected to influence F5. This affordance then activates the F5 neurons to command the appropriate grip once it receives a “go signal” from another region, F6 (pre-SMA), of prefrontal cortex.

### MNS

The MNS model explains how mirror neurons are developed through learning. The main hypothesis of the model is that the activity of F5 canonical neurons which control the grasp provides a training signal for mirror neurons to associate with the temporal profile of the visual features experienced during

self-executed grasps. Thus, developmentally grasp learning precedes initial mirror neuron formation and the model shows how the mirror system may learn to recognize actions already in the repertoire of the F5 canonical neurons. Because these actions are encoded by the model in an object-centered reference frame, it can recognize another agent’s actions since they elicit similar input to the system as self-observation while performing the same actions.

In order to generate grasps for mirror system training, MNS uses a simplified version of a portion of the FARS model (Fig. 1). In this simplified version, the reach and grasp movements are generated using techniques from robotics (inverse kinematics, gradient descent, and joint angle interpolation), the F5 canonical neurons only encode the type of grasp (power, precision or side) and not the phase, the ventral visual stream is not included (except for the superior temporal sulcus) and there are no reciprocal connections between AIP and F5. MNS therefore uses FARS as more of a background organization for its architecture and cannot capture the possible influence of the ventral stream or grasp phase on mirror neuron activity. This simplifies the modeling effort to more easily achieve the goals of the MNS model—to show how mirror neuron activity can be shaped by canonical neural activity, but leaves open these avenues for future model development.



**Fig. 1** The FARS diagram in a format that allows more direct comparison with the MNS model. The portions retained by MNS are shown in the upper and lower diagonal boxes. The reach circuit (*bottom green rectangle*) is comprised of the posterior parietal cortex which extracts the object’s location, and passes it to the ventral premotor region F4 for programming the reach. The grasp circuit begins in the caudal

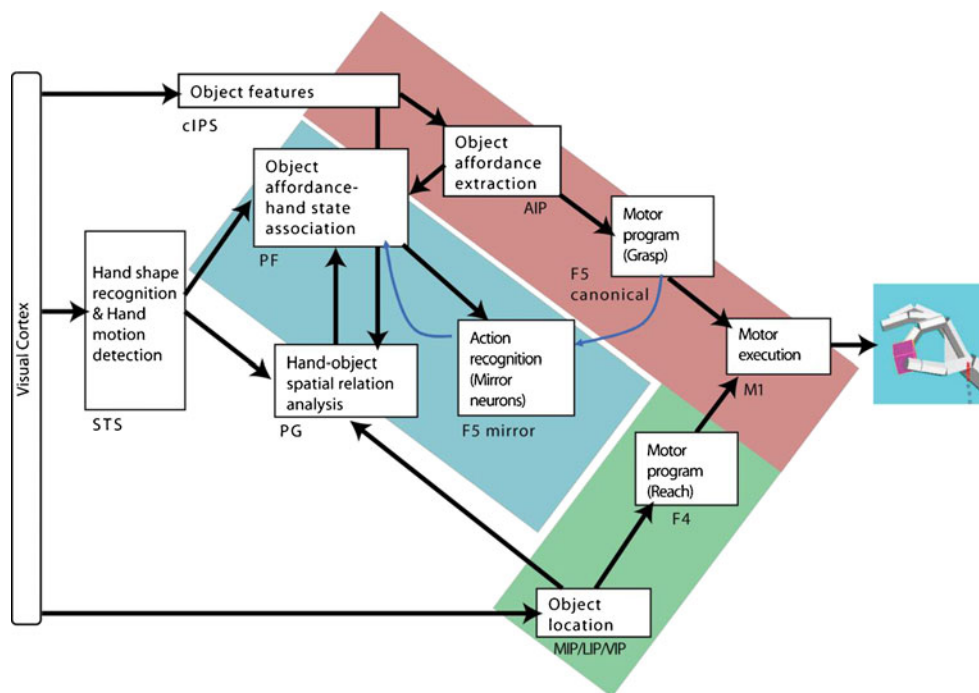
intraparietal sulcus (cIPS) which extracts object shape features for grasp affordance extraction (opportunities for grasping) in the anterior parietal sulcus (AIP). The ventral premotor region F5 plans a grasp motor program based on these affordances. The inferior temporal cortex (IT) recognizes the identity of the object and along with the prefrontal cortex (PFC) can influence the grasp

MNS is divided into three submodules which are further comprised of networks of modules representing brain regions (Fig. 2). The Reach & Grasp module controls the movement and as described above is based on a simplified portion of the FARS model. The Visual Analysis of Hand State module recovers the three dimensional configuration of the hand from the two-dimensional visual input image and outputs a representation of the shape and motion of the hand relative to an object affordance. The Core Mirror Circuit module recognizes the grasp being observed based on the trajectory of hand-object relations up to the current point in time. During observation of self-performed grasps an efferent copy of the F5 canonical neuron activity provides a training signal for learning these hand-object relation trajectory—grasp representation mappings. The inputs to the model are therefore a visual image of the hand configuration at each time step, object-centered affordance features such as orientation, shape and size, and the object location in three-dimensional space. The model outputs are the firing rate of F5 mirror neurons indicating which grasp was recognized, and a vector of motor commands (grasp type and reach location) used for performing reach and grasp movements.

In this model, the vector input of hand—object affordance relations is referred to as the *hand state*. Oztop and Arbib

(2002) showed that an artificial neural network corresponding to PF and F5 mirror neurons could be trained to recognize the grasp type from the *hand state trajectory*, with correct classification often being achieved well before the hand reached the object. During training, the activity of F5 canonical neurons encodes the grasp currently being executed by the monkey, and serves as a training signal for F5 mirror neurons. This allows F5 mirror neurons to respond to hand-object trajectories associated with different types of grasps. After training, the appropriate mirror neurons begin to fire in response to viewing the appropriate trajectories even when unaccompanied by F5 canonical firing. Crucially, this training allows F5 mirror neurons to respond to observation of hand-object relational trajectories even during observation of another agent because the hand state is based on the view of a hand moving relative to the object.

F5 canonical neurons were modeled as an array of neurons whose activity determined the type of grasp executed. The learning mechanism used was backpropagation with one hidden layer. The major drawback of the MNS model was its treatment of the hand state trajectory. At each time point, the initial segment of the trajectory up to that time was fitted by a cubic spline, and then sampled at 30 times spanning the segment to produce a 210 dimensional input vector to the



**Fig. 2** The MNS model. (i) Top diagonal: a portion of the FARS model. Object features are processed by AIP to extract grasp affordances, these are sent on to the canonical neurons of F5 that choose a particular grasp. (ii) Bottom right. Recognizing the location of the object provides parameters to the motor programming area F4 which computes the reach. The information about the reach and the grasp is taken by the motor cortex M1 to control the hand and the arm. (iii) New elements of the MNS model: At the bottom left are two schemas, one for recognizing the shape of the

hand, and the other for determining the hand's motion. Just to the right of these is the schema for hand-object spatial relation analysis. It takes information about object features, the motion of the hand and the location of the object to infer the relation between hand and object. (iv) The center two regions marked by the bottom-left rectangle form the core mirror circuit. These regions associate the visual input (hand state) with the motor program input from region F5 canonical neurons during the learning process

network. In this way, the temporal representation of hand state was pre-processed such that it could be encoded in a spatial representation for input into the feedforward network. The network was trained on a set of “self-performed” grasps using the activity of F5 canonical neurons as the training signal for the output layer of the network, representing F5 mirror neurons.

The MNS model used a connectionist architecture, but was designed based on connectivity and functional properties of relevant parts of the primate brain. The network diagram generated by BODB showing the connectivity SEDs used to build MNS is shown in Fig. 3. This figure demonstrates the difficulty in building a model based on experimental data which refers to different, sometimes conflicting, brain nomenclatures. Each connectivity SED is represented in this diagram by an arrow connecting two nodes, where the nodes represent brain regions and the direction of the arrow indicates the direction of the projection. Each node contains the abbreviation for the brain region it represents (i.e. MIP stands for medial intraparietal area), followed by the abbreviation for the nomenclature it is defined in (i.e., SP stands for Seltzer & Pandya). One problem is immediately apparent and that is that regions with the same name are defined in different nomenclatures. Efforts have been made to evaluate the degree of overlap between two regions in different nomenclatures (Bota and Arbib 2004; Kotter 2004), but currently BODB groups regions with the same name together. Another issue is that of nomenclatures that use different names for the same region. For example, Fig. 3 shows connectivity for regions 7a and 7b, which are widely agreed to correspond to areas PG and PF, respectively.

The remaining SEDs used to build the MNS model are listed in Table 1. The first SED, “Separation of dorsal and ventral visual streams” is linked to the MNS model with the *scene setting* relationship, meaning that it was not used for any design decisions, but places the focus of the model (the dorsal visual stream) in a larger context. The remaining SEDs describe neurophysiological properties of brain regions included in the model and the relevance narrative for each describes how these properties influence the functionality of the related MNS modules. Because BODB currently does not allow users to enter connectivity SEDs, some entries such as cIPS

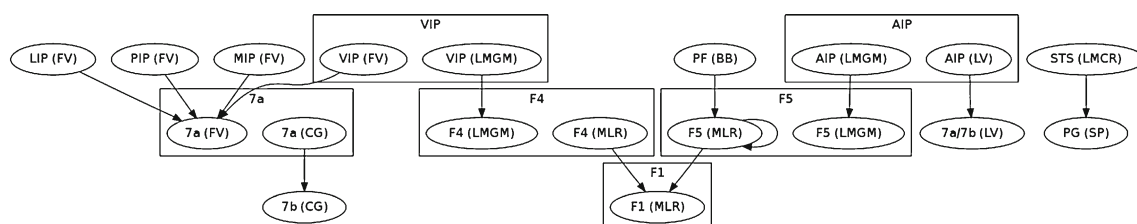
projection to AIP describe projections that are not yet included in CoCoMac, the database from which BODB connectivity SEDs are derived.

SEDs used to test the MNS model are shown in Table 2. Each test SED is linked to a summary of simulation results (SSR) from the model. The relevance narrative for each SED-SSR pair describes how the data described in the SED is explained by or contradicts the results described in the SSR. In this case, F5 mirror properties and transitive action selectivity are explained by corresponding simulation results from the MNS model. Some SSRs cannot be linked to existing SEDs and therefore make predictions by the model for future experiments. Predictions made by the MNS model and their associated SSRs are listed in Table 3. Currently, BODB allows predictions to describe what should be found if an experiment analogous to the linked SSRs is performed. In order to facilitate rapid interchange between experimentalists and modelers, future neuroinformatics tools must make explicit the link between experimental protocol and the simulation protocol used to approximate it (see Discussion).

The brain operating principles (BOPs) associated with the MNS model are listed in Table 4. MNS makes use of a supervised learning technique, back propagation, in order to train the mirror neuron response to visual inputs using F5 canonical neuron activity as a training signal. In this way the visual sensory information is coupled with motor representations for grasping. The learning rule is non-biological, but in more realistic implementations, F5 canonical neuron activity could still provide a training signal or eligibility trace for a more biologically plausible learning rule.

## MNS2

The MNS2 model extended the MNS model of the macaque mirror system to address data on audio-visual mirror neurons and mirror neurons that respond to the observation of partially hidden grasps. MNS2 used a recurrent architecture that is biologically more plausible than that of the original model. The MNS model performed a temporal-to-spatial transformation on the hand state trajectory in order to input it to a feed-forward neural network. In contrast, the MNS2 model used a recurrent network that allowed the raw hand state to be input



**Fig. 3** A connectivity SED diagram generated by BODB for the MNS model (nomenclatures: BB: von Bonin & Bailey, 1947; CG: Cavada & Goldman-Rakic, 1989; FV: Felleman & Van Essen, 1991; LMCR:

Luppino et al., 1993; LMGM: Luppino et al., 1999; LV: Lewis & Van Essen, 2000; MLR: Matelli et al., 1985; SP: Seltzer & Pandya, 1986)



**Table 1** Summaries of Experimental Data (SEDs) used to build the MNS model. Each SED lists some property of the brain or brain region(s) and can be either scene setting or supporting of the model's design. The relevance narrative describes the relationship between the model and data in further detail

Name	Description	Relationship	Relevance narrative
Separation of dorsal and ventral visual streams	The dorsal visual path extends into the parietal cortex and represents information crucial for object interaction while the ventral stream extends into the inferotemporal cortex and recognizes the identity of an object (Ungerleider & Mishkin, 1982; Goodale & Milner, 1992; Castiello & Jeannerod, 1991)	scene setting	The module focuses on the dorsal stream
AIP Grasp Selectivity	Some neurons in AIP are selective for different types of grasps	support	The Object affordance extraction module encodes grasp affordance information
cIPS projection to AIP	The caudal intraparietal sulcus projects to AIP (Sakata et al., 1997)	support	The module corresponding to cIPS projects to the module corresponding to area AIP
F4 reach selectivity	F4 represents the arm goal position and sets up the initial reach program	support	The Reach motor program module programs the reaching movement
PIP object selectivity	The posterior (or caudal) intraparietal sulcus (PIP or cIPS) codes object-centered information such as object shape and different parameters for different shapes such as cylinder width, surface orientation, and axes orientation (Sakata et al., 1997)	support	The Object features module encodes object shape information
VIP location selectivity	Area VIP represents the location and orientation of target objects in a broadly tuned population code in a peripersonal coordinate system (Colby et al., 1993)	support	The Object location module encodes the location of the center of the object
Area 7 is reciprocally connected with STS	Areas 7a, 7ip, and 7b, to a lesser extent, are reciprocally connected with the cortex of the STS (Cavada, Goldman-Rakic, 1989)	support	The module corresponding to STS projects to the modules corresponding to areas 7a and 7b
STS shape-selective cells	Shape-selective cells respond selectively to hand-object interactions and respond similarly to the sight of the monkeys own movements as well as those of others.	support	The Hand shape recognition schema responds to observation of hand-object interactions regardless of who is performing them
STS visual response	STS contains neurons that respond to the observation of actions such as walking, limb movements, and grasping	support	The Hand motion detection schema responds to the observation of hand-object actions
F5 canonical visual properties	Some F5 cells respond to the observation of graspable objects (Rizzolatti et al., 1988; Murata et al., 1997)	support	Neurons in the Grasp motor program module area activated by visual responses in the AIP module
F5 grasp selectivity	Different classes of F5 neurons discharge during different hand movements (grasping, holding, tearing, manipulating) and can be selective for either precision grip, finger prehension, or whole hand prehension (Rizzolatti et al., 1988)	support	Neurons in the Grasp motor program module are selective for different grasp types
F5	Some F5 mirror neurons the congruence between preferred observed and performed actions	scene setting	The model focuses on strictly congruent mirror neurons

**Table 1** (continued)

Name	Description	Relationship	Relevance narrative
mirror—broadly congruent	is quite loose. For example, a mirror neuron may fire during performance of precision pinches and observation of power grasps		
F5 mirror—strictly congruent	For some mirror neurons, not only must the general action (e.g., grasping) match, but also the way the action is executed (e.g., power grasp) must match as well.	support	After training, neurons in the Action recognition module behave like strictly congruent mirror neurons

at each time step. Moreover, MNS2 extended the capacity of the model to address data on audio-visual mirror neurons and on the response of mirror neurons when the target object was recently visible but is currently hidden. Another recurrent network was used to recognize sounds given by patterns of

auditory nerve firings. The output of this network was associated with the output of the main network so that mirror neurons could respond to the characteristic sound of an action even if the action itself was not currently visible. A working memory representation of the arm and hand along with

**Table 2** Summaries of Experimental Data (SEDs) used to test the MNS model, linked to an SSR generated by the model which either explains the data or is contradicted by it. The relevance narrative describes the relationship between the simulation results and experimental data in further detail

Name	Relationship	Relevance narrative
F5 mirror properties from object-centered representation	explanation	F5 mirror neurons derive their mirror properties from associating an object-centered action representation with efference copies of F5 canonical activity
<b>SED</b>		
<b>Name</b>		<b>Description</b>
F5 mirror properties		Mirror F5 neurons respond during observation of a grasping movement (di Pellegrino et al., 1992; Rizzolatti et al., 1996; Gallese et al. 1996). Like other F5 neurons, mirror neurons are active when the monkey performs a particular class of actions, such as grasping, manipulating, and placing. However, in addition the mirror neurons become active when the monkey observes the experimenter or another monkey performing an action.
<b>SSR</b>		
<b>Name</b>		<b>Description</b>
F5 action recognition		Simulated F5 mirror neurons are trained to recognize hand state trajectories using F5 canonical activity as a training signal. They are then able to recognize a grasp before the hand contacts the object.
F5 mirror response to intransitive actions	explanation	F5 mirror training during observation of self-performed grasps leads to an insensitivity to grasps that do not contact an object
<b>SED</b>		
<b>Name</b>		<b>Description</b>
F5 mirror—transitive action selectivity		To be triggered, the mirror neurons require an interaction between the hand motion and the object. The vision of the hand motion or the object alone does not trigger mirror activity (Gallese et al. 1996; Umiltà et al., 2001).
<b>SSR</b>		
<b>Name</b>		<b>Description</b>
F5 spatial perturbation		Simulated F5 mirror neurons show a reduced response during observation of a grasp that does not contact an object

**Table 3** Predictions made by the MNS model. Each prediction can be associated with one or more SSRs

Name	Description				
Effect of explicit affordance coding on mirror neuron response to hand state trajectory-affordance mismatch	<p>If object affordance information is explicitly available to mirror neurons, it can in some cases overwhelm the influence of the hand state trajectory when it is not appropriate for the object affordance</p> <p><b>SSRs</b></p> <table> <tr> <th>Name</th><th>Description</th></tr> <tr> <td>F5 mirror response to altered kinematics</td><td>Simulated F5 mirror neurons do not respond to observation of reach-and-grasp movements that have constant arm-joint velocities, disrupting the time course (but not the path) of the wrist trajectory.</td></tr> </table>	Name	Description	F5 mirror response to altered kinematics	Simulated F5 mirror neurons do not respond to observation of reach-and-grasp movements that have constant arm-joint velocities, disrupting the time course (but not the path) of the wrist trajectory.
Name	Description				
F5 mirror response to altered kinematics	Simulated F5 mirror neurons do not respond to observation of reach-and-grasp movements that have constant arm-joint velocities, disrupting the time course (but not the path) of the wrist trajectory.				
Mirror neuron response to ambiguous grasps	<p>A grasp with an ambiguous prefix may drive the mirror neurons in such a way that the system will, in certain circumstances, at first give weight to the wrong classification, with only the late stages of the trajectory sufficing for the incorrect mirror neuron to be vanquished.</p> <p><b>SSRs</b></p> <table> <tr> <th>Name</th><th>Description</th></tr> <tr> <td>F5 mirror response to ambiguous grasps</td><td>If the beginning of a grasp is ambiguous, mirror neurons selective for other grasps may become active until the grasp can be resolved.</td></tr> </table>	Name	Description	F5 mirror response to ambiguous grasps	If the beginning of a grasp is ambiguous, mirror neurons selective for other grasps may become active until the grasp can be resolved.
Name	Description				
F5 mirror response to ambiguous grasps	If the beginning of a grasp is ambiguous, mirror neurons selective for other grasps may become active until the grasp can be resolved.				
Mirror neuron response to artificial movements	<p>Mirror neurons should not respond to the observation of actions with artificial kinematics (i.e. without bell-shaped velocity profiles)</p> <p><b>SSRs</b></p> <table> <tr> <th>Name</th><th>Description</th></tr> <tr> <td>F5 mirror response to different sized objects with same hand state trajectory</td><td>When objects of different sizes are grasped with the same precision pinch hand state trajectory the mirror neurons can recognize it as a pinch. However, with explicit affordance encoding, if the object is big enough the influence of the affordance information can overwhelm that of the hand state and cause the power grasp mirror neurons to be activated</td></tr> </table>	Name	Description	F5 mirror response to different sized objects with same hand state trajectory	When objects of different sizes are grasped with the same precision pinch hand state trajectory the mirror neurons can recognize it as a pinch. However, with explicit affordance encoding, if the object is big enough the influence of the affordance information can overwhelm that of the hand state and cause the power grasp mirror neurons to be activated
Name	Description				
F5 mirror response to different sized objects with same hand state trajectory	When objects of different sizes are grasped with the same precision pinch hand state trajectory the mirror neurons can recognize it as a pinch. However, with explicit affordance encoding, if the object is big enough the influence of the affordance information can overwhelm that of the hand state and cause the power grasp mirror neurons to be activated				
Mirror neuron response to grasp-object axis mismatches	<p>Mirror neurons should not respond when the orientation of the grasp does not match the axis of the object</p> <p><b>SSRs</b></p> <table> <tr> <th>Name</th><th>Description</th></tr> <tr> <td>F5 mirror response to grasp-object axis mismatch</td><td>Simulated F5 mirror neurons do not respond when the orientation of the observed grasp does not match the axis of the object</td></tr> </table>	Name	Description	F5 mirror response to grasp-object axis mismatch	Simulated F5 mirror neurons do not respond when the orientation of the observed grasp does not match the axis of the object
Name	Description				
F5 mirror response to grasp-object axis mismatch	Simulated F5 mirror neurons do not respond when the orientation of the observed grasp does not match the axis of the object				
Temporal effect of explicit affordance coding on mirror neuron response	<p>If object affordance information is explicitly available to mirror neurons, they should respond faster when the grasp matches the observed affordance.</p> <p><b>SSRs</b></p> <table> <tr> <th>Name</th><th>Description</th></tr> <tr> <td>F5 mirror response to precision pinch with explicit affordance encoding</td><td>With explicit affordance encoding the mirror neuron selective for precision pinches was activated earlier when the observed pinch was directed toward smaller objects.</td></tr> </table>	Name	Description	F5 mirror response to precision pinch with explicit affordance encoding	With explicit affordance encoding the mirror neuron selective for precision pinches was activated earlier when the observed pinch was directed toward smaller objects.
Name	Description				
F5 mirror response to precision pinch with explicit affordance encoding	With explicit affordance encoding the mirror neuron selective for precision pinches was activated earlier when the observed pinch was directed toward smaller objects.				

**Table 4** Brain Operating Principles (BOPs) related to the MNS model

Name	Description	Relationship
Sensorimotor Coupling	The brain has representations that bridge between perceptual contents and action plans. Such common codes integrate action and perception. Perceptual stimuli are interpreted in part on the motor representations which are commonly associated with the corresponding sensory experience.	Mirror neurons couple visual sensory information with motor representations for grasping
Supervised learning	Training based on a teaching signal indicating the error on the last performance	The F5 mirror neurons are trained with backpropagation using the activity of the F5 canonical neurons as a training signal
Corollary Discharge	This principle involves the sending of a copy (known as the efferent copy) of a motor output.	The F5 mirror neurons are trained using a copy of the F5 canonical neuron output (motor signal).

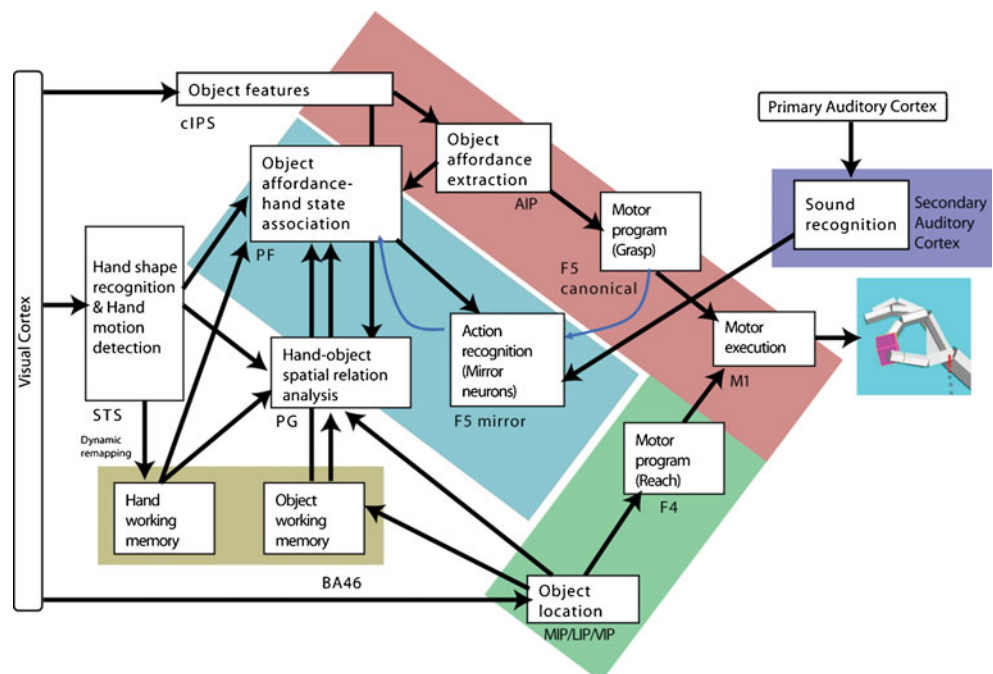
dynamic remapping of the hand representation was used to recognize partially hidden grasps (Fig. 4).

The main components of MNS2 are the audio and visual recurrent neural networks (RNNs), cochlear model, hand and object working memory, and the hand/object association module. The hand/object association schema supplies input to the visual RNN in the form of hand/object relations such as axis disparities, joint angles, distance, and velocity. The hand and object working memory schemas relay the current visual hand and object signals to the hand/object association schema, but if either information is not available because the hand or object is invisible, a stored copy of the last observed values is passed. In the case of the hand working memory, this copy

can be remapped using visual information about the movement of the arm. As in the MNS model, F5 canonical neurons encode the type of grasp being performed and their activity is used to train neurons in the output layer of visual RNN (corresponding to mirror neurons) to classify grasps based on visual information. The training algorithm used was back propagation through time (Werbos 1990), allowing the network to learn to associate hand state trajectories with the appropriate motor signal used to generate the observed grasp.

A combination of the type of object being grasped and the type of grasp being performed determined the sound that was generated in the MNS2 simulation experiments. Audio information was preprocessed by the Lyon passive ear model

**Fig. 4** The MNS2 model builds upon the MNS model to address additional experimental data. The addition of working memory networks for the hand and object allow mirror neurons to respond to grasps whose final state is occluded from view. A model of the primary and secondary auditory cortices allows the model to recognize sounds and associated them through Hebbian learning with the actions that produce them





(Lyon 1982) which generates an array of auditory nerve firing probabilities along the length of the cochlea. The normalized output of this model was associated with a localist representation of sound identity using a separate auditory recurrent neural network also trained using back propagation through time. The auditory network output layer projected to the output layer of the visual network. These connection weights were modified using Hebbian learning.

In this model, visual information about the hand is provided to working memory through the STS and the working memory relays this information to areas 7a and 7b when the hand is not visible. Visual information about the object reaches object working memory from AIP and the medial, lateral, and ventral intraparietal areas (MIP/LIP/VIP). Dynamic remapping was used to extrapolate the observed grasp trajectory once the hand disappears behind the screen. Dynamic remapping is a process whereby perceptual representations are updated based on generated motor commands, or related perceptual information. In the present model, at each time step that the screen obscures the hand, the representation of the movement of the still-visible forearm in STS is used to update the working memory representation in area 46 of the hand position. In this way, if the model observes an object that is then hidden by a screen, and then observes a grasp that disappears behind that screen, the hand trajectory will be extrapolated—and if it appears to end at the remembered object location then the grasp will be recognized.

The connectivity SED diagram generated by BODB for the MNS2 model is shown in Fig. 5. The modeled network is basically the same as that as MNS (Fig. 3), with the addition of area 46, which is connected with STS and inferior parietal areas and implements working memory in the model. The auditory projections used to extend MNS to address audiovisual mirror neurons are largely absent from CoCoMac, and are thus listed in Table 5 along with other generic SEDs used to build MNS2. Because MNS2 extends the original MNS model, it inherits the same SEDs used to build MNS and therefore Table 5 only lists additional SEDs which are unique to MNS2.

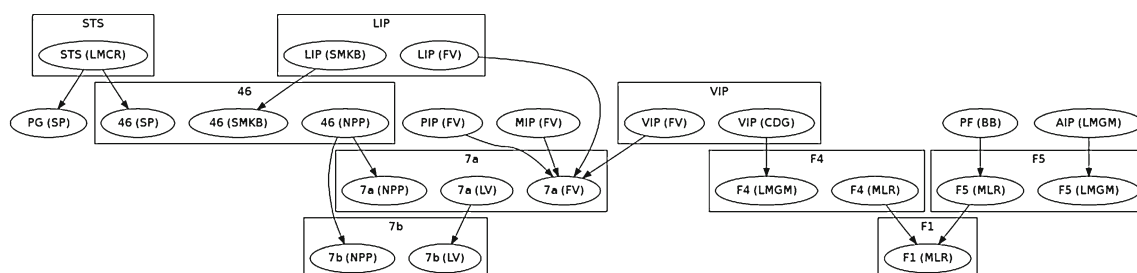
The SEDs used to test the MNS2 model are listed in Table 6. Again, because MNS2 extends MNS, it addresses some of the same SEDs such as “F5 mirror properties from

object-centered representation” and “F5 mirror response to intransitive actions”, but links them to its own SSRs which demonstrate that this new version of the MNS model can also explain these data. Additional SEDs are listed in Table 6 that test the novel features of the MNS2 model such as audiovisual mirror properties and mirror responses to partially hidden grasps. Predictions made by the MNS2 model are listed in Table 7. MNS2 predicts that if dynamic remapping is used to update the working memory representation of the hand during observation of partially hidden grasps, mirror neurons should cease firing once the hand over- or undershoots the hidden target object.

The BOPs associated with MNS2 are listed in Table 8. In addition to sensorimotor coupling, and supervised learning, MNS2 makes use of dynamic remapping, Hebbian learning, sensor fusion, temporal pattern processing, and working memory. The addition of the temporal pattern processing BOP is due to the change from MNS to MNS2 to the use of a recurrent neural network to process raw hand state trajectories rather than the temporal-to-spatial transformation that MNS required on its input. The other BOPs come from the MNS2 model’s emphasis on the role of mirror neurons within a larger network whose other regions influence their firing properties. Processed auditory information projects to mirror neurons allowing them to fuse visual and auditory signals via Hebbian learning. Working memory representations of the hand and object and dynamic remapping of hand state working memory allows mirror neurons to maintain their response and predictive capabilities during partially hidden grasps.

## HAMMER

The human ability to learn new skills by imitation has inspired roboticists to replicate such abilities on robots with the ultimate task of programming robots by demonstration (for extensive reviews of these efforts, see (Schaal 1999; Argall et al. 2009)). A subset of these approaches, has additionally attempted to relate the functional characteristics of the computational architectures to biological data. Among them, the HAMMER (Hierarchical Attentive Multiple Models for



**Fig. 5** A connectivity SED diagram generated by BODB for the MNS2 model (nomenclatures: BB: von Bonin & Bailey, 1947; CDG: Colby et al., 1993; CG: Cavada & Goldman-Rakic, 1989; FV: Felleman & Van

Essen, 1991; LMCR: Luppino et al., 1993; LMGM: Luppino et al., 1999; LV: Lewis & Van Essen, 2000; MLR: Matelli et al., 1985; NPP: Neal et al., 1990; SMKB: Schall et al., 1995; SP: Seltzer & Pandya, 1986)

**Table 5** Summaries of Experimental Data (SEDs) used to build the MNS2 model

Name	Description	Relationship	Relevance narrative
Auditory cortex projection to F5	The auditory cortex projects to the caudal inferior arcuate cortex (Deacon, 1992)	support	The module corresponding to the auditory cortex projects to the module corresponding to area F5
BA 46 role in working memory	Brodman area 46 is typically implicated in working memory (Courtney et al., 1998; McCarthy et al., 1994)	scene setting	The Working Memory schema corresponds to area 46
Indirect auditory cortex projection to premotor cortex	The auditory cortex projects to area 6 via area 8 (Arikuni et al., 1988; Romanski et al., 1999)	support	The module corresponding to the auditory cortex projects to the module corresponding to area F5
Nonprimary auditory cortex encoding	Macaque nonprimary cortex responds to complex sounds (Rauschecker et al., 1995)	support	The output of the Sound recognition schema uniquely identifies different sounds
Primary auditory cortex organization	The macaque primary cortex has a tonotopic organization (Morel et al., 1993)	support	The input to the Sounds recognition schema is a tonotopic representation of the sound
Working memory activity in area 46 and inferior parietal	Friedman and Goldman-Rakic (1994) have found activity in both the dorsolateral prefrontal cortex and inferior parietal cortex (areas 7A, 7B, 7IP, and 7M) of the macaque monkey during spatial working memory tasks.	support	The Working Memory schema is reciprocally connected with the modules representing areas 7a and 7b

Execution and Recognition) architecture (Demiris and Hayes 2002; Demiris and Khadhoury 2006), which attempts to draw functional parallels between the normal and lesioned operation of the model with data from monkey neurophysiology (mirror neurons, Gallese et al. 1996), human imaging (mirror and action observation network) and human pathological conditions (visuoimitative apraxia, Goldenberg and Hagmann 1997), and advocates multiple routes to understanding observed actions (Demiris and Hayes 1999) (Fig. 6).

#### *HAMMER's Multiple Routes to Understanding Actions*

The first design of the basic architecture was based around the need to imitate and learn basic movements, in order to allow robots to imitate basic human actions. The first iteration of this architecture (Demiris et al. 1997) used a posture estimation algorithm, storing representative postures of the demonstration in working memory and reproducing them using closed-loop control. Although this allowed a robotic head to imitate and learn basic human head movements by observation, as well as other tasks relying on body postural configurations, the discovery of mirror neurons shed light to the tight coupling between action perception and action generation, requiring that more active (generative) mechanisms are used by the

robot during observation of movements. Demiris and Hayes (1999, 2002) introduced a second route to the visual processing of actions (explained below), using multiple inverse and forward model pairs to generate (online) multiple hypotheses as to the nature of the observed action, and increase dynamically the levels of confidence (representing their activation levels) on whether they can match the ongoing demonstration. The interesting aspect of this route was that the inverse models used to control the robot during execution were also used in order to understand actions when performed by others (Fig. 7).

#### *Model's Operation—the Generative Route*

The generative route consists of multiple pairs of inverse and forward models that operate in parallel (J. Demiris and Hayes 2002). When the demonstrator agent executes a particular action the perceived states are fed into all of the observer's available inverse models. This generates multiple motor commands (representing multiple hypotheses as to what action is being demonstrated) that are sent to the forward models. The forward models generate predictions about the demonstrator's next state: these are compared with the actual demonstrator's state at the next time step, and the error signal resulting from this comparison affects the confidence values of the inverse

**Table 6** Summaries of Experimental Data (SEDs) used to test the MNS2 model

Name	Relationship	Relevance narrative
F5 mirror properties from object-centered representation	explanation	F5 mirror neurons derive their mirror properties from associating an object-centered action representation with efference copies of F5 canonical activity
	<b>SED</b>	
	<b>Name</b>	<b>Description</b>
	F5 mirror properties	Mirror F5 neurons respond during observation of a grasping movement (di Pellegrino et al., 1992; Rizzolatti et al., 1996; Gallese et al. 1996). Like other F5 neurons, mirror neurons are active when the monkey performs a particular class of actions, such as grasping, manipulating, and placing. However, in addition the mirror neurons become active when the monkey observes the experimenter or another monkey performing an action.
F5 mirror response to intransitive actions	<b>SSR</b>	
	<b>Name</b>	<b>Description</b>
	F5 action recognition	Simulated F5 mirror neurons are trained to recognize hand state trajectories using F5 canonical activity as a training signal. They are then able to recognize a grasp before the hand contacts the object.
	explanation	F5 mirror training during observation of self-performed grasps leads to an insensitivity to grasps that do not contact an object
F5 audiovisual mirror properties from Hebbian association	<b>SED</b>	
	<b>Name</b>	<b>Description</b>
	F5 mirror—transitive action selectivity	To be triggered, the mirror neurons require an interaction between the hand motion and the object. The vision of the hand motion or the object alone does not trigger mirror activity (Gallese et al. 1996; Umiltà et al., 2001).
	<b>SSR</b>	
F5 audiovisual mirror properties from Hebbian association	<b>Name</b>	<b>Description</b>
	F5 pantomimed grasp	Simulated F5 mirror neurons do not respond during observation of a pantomimed grasp.
	explanation	F5 mirror neurons that are active during observation of an action can come to respond to characteristics sounds of the action through Hebbian association with the output of sound recognition areas
	<b>SED</b>	
F5 audiovisual mirror properties from Hebbian association	<b>Name</b>	<b>Description</b>
	F5 audiovisual mirror neurons	Kohler et al. (2002) found that some of the mirror neurons in area F5 of the macaque premotor cortex that are responsive for the observation of actions associated with characteristic noises are just as responsive for the sounds of these actions.
	<b>SSR</b>	
	<b>Name</b>	<b>Description</b>
	F5 audiovisual action recognition	Hebbian association of object recognition input with F5 mirror activity allows F5 mirror neurons to respond to invisible actions when the sound associated with them is perceived.

**Table 6** (continued)

Name	Relationship	Relevance narrative
F5 mirror response to partially hidden grasps	explanation	A working memory representation of the hand and object as well as dynamic remapping of the hand position using observed arm movement allows mirror neurons to respond to observed grasps that are partially hidden by a screen.
<b>SED</b>		
<b>Name</b>	<b>Description</b>	
F5 mirror—partially hidden grasps	Mirror neurons selective for grasping do not respond to pantomimed grasps but will respond to grasps that are partially hidden by a screen if an object was recently seen behind it. These neurons will not respond to a grasp directed behind the screen if no object is known to be behind it.	
<b>SSR</b>		
<b>Name</b>	<b>Description</b>	
F5 hidden grasp	Simulated F5 mirror neurons respond to visible and partially hidden grasps, but do not respond to visible or partially hidden pantomimed grasps	
IPL visual properties through training	explanation	The model shows how IPL neurons can selectively respond to the observation of different types of grasps through training.
<b>SED</b>		
<b>Name</b>	<b>Description</b>	
IPL visual properties	The IPL contains neurons that respond to grasp observation, but not production	
<b>SSR</b>		
<b>Name</b>	<b>Description</b>	
Discriminative grasp hidden units	After training, neurons in the hidden layer corresponding to the inferior parietal lobule (IPL) selectively respond to the observation of specific grasps.	

models. At the end of the demonstration (or earlier if required) the inverse model with the highest confidence value, i.e. the one that is the closest match to the demonstrator's action is selected. Coordinate transformations between observed

and executed actions are handled through a perspective taking process (e.g. Johnson and Demiris 2005). This architecture has been implemented in real-dynamics robot simulations (Demiris and Hayes 2002) and real robots (Demiris and Johnson 2003; Demiris and Khadhour 2006). The inverse models can be arranged in a hierarchical manner where primitive inverse models are combined (either sequentially or in parallel) to form higher more complex sequences (Figs. 6 and 8).

**Table 7** Predictions made by the MNS2 model

Name	Description
F5 mirror response to partially hidden overshoot grasps	Mirror neurons should be able to detect when a partially hidden grasp will not actually contact the object and cease responding.
<b>SSRs</b>	
<b>Name</b>	<b>Description</b>
F5 mirror response to partially hidden grasp	With dynamic remapping, a partially hidden grasp that overshoots the object causes mirror neurons to cease responding.

### Model's Operation—the Passive Route

Although a large portion of the published literature on the HAMMER architecture revolved around the generative route, the imitation capabilities of this route are limited by what inverse models the architecture already has in its repertoire. If a demonstrated action or its consequences cannot be replicated by one of the inverse models already in the observer's repertoire, control is



**Table 8** Brain Operating Principles (BOPs) related to the MNS2 model

Name	Description	Relationship
Sensorimotor Coupling	The brain has representations that bridge between perceptual contents and action plans. Such common codes integrate action and perception. Perceptual stimuli are interpreted in part on the motor representations which are commonly associated with the corresponding sensory experience.	Mirror neurons couple visual and auditory sensory information with motor representations for grasping
Supervised learning	Training based on a teaching signal indicating the error on the last performance	The F5 mirror neurons are trained with backpropagation using the activity of the F5 canonical neurons as a training signal
Corollary Discharge	This principle involves the sending of a copy (known as the efferent copy) of a motor output.	The F5 mirror neurons are trained using a copy of the F5 canonical neuron output (motor signal).
Dynamic Remapping	Dynamic remapping refers to the updating of a neural map to reflect the result of an executed action.	The working memory representation of the hand location is dynamically remapped using the detection of arm motion.
Hebbian Learning	Hebbian Learning is a type of learning for neurons in which a synapse is strengthened if its activity coincides with the firing of the postsynaptic neuron -	Connections between the output of the auditory circuit and F5 mirror neurons are modified using Hebbian learning.
Sensor Fusion	Sensor fusion typically refers to the integration of data from different types of sensors, as in combining visual, auditory and tactile cues, to more accurately extract relevant information from the environment.	F5 mirror neurons combine visual and auditory information to recognize grasps
Temporal Pattern Processing	Temporal pattern processing is important for various intelligent behaviors, including hearing, vision, speech, music, and motor control. Because we live in an ever-changing environment, an intelligent system must encode patterns over time, recognizing and generating temporal patterns.	Hand state input comes as a temporal trajectory.
Working Memory	Working memory consists of contextual information that is stored for the duration of the execution span of the current task.	Working memory representations of hand and object location are used to recognize partially hidden grasps

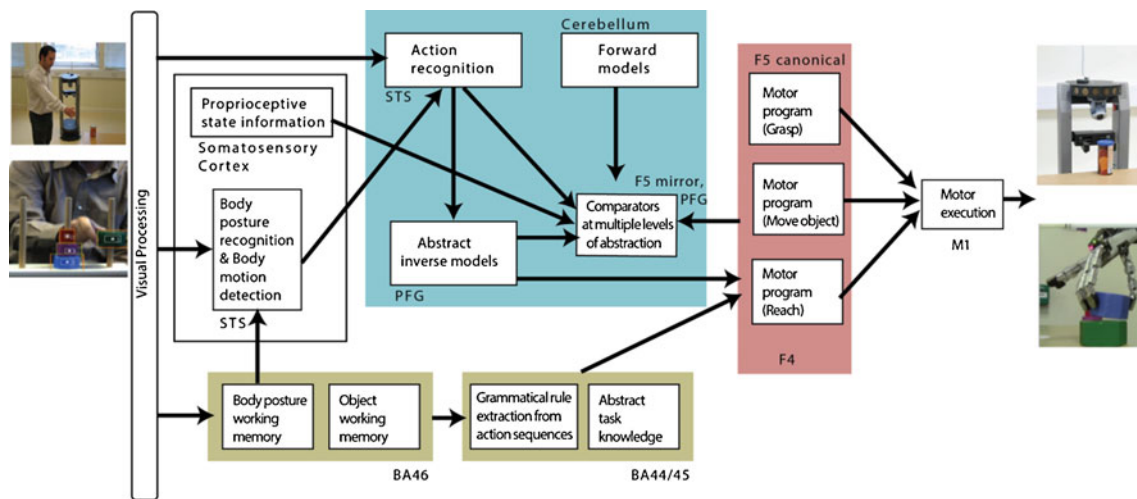
passed to the passive (or learning) route, which extracts salient features of the demonstration and attempts to replicate them, while storing successful results as new inverse models, for later use by the generative route.

In early experiments (Demiris and Hayes 2002), the extracted features were sequences of postural configurations, stored in a buffer (working memory), which were replicated using closed-loop control with proprioceptive information provided the current state and the stored postural configurations providing the target states. In more recent experiments (Lee et al. 2012) further processing (classification) of the visual information allows grammatical rules that represent the hierarchical structure of the demonstrated action sequences to be learned, and used to replicate the task, or reason about a demonstration that differs from what is expected. This has been applied to learning to reason about complex task

sequences such as the Towers of Hanoi, and complex dance sequences (Lee et al. 2012) (Figs. 9 and 10).

#### *Establishing Mappings Between Computational and Brain Structures*

A key point that becomes evident from the description of the HAMMER model above is that an exact mapping between structures of the computational models and the corresponding brain areas of humans and macaque monkeys is challenging. BODB's novel approach of structuring the description of the model around Brain Operating Principles (BOPs) allows equivalences to be drawn between robotic models and brain models at the level of functionality, while Summaries of Experimental Data (SEDs) allow for the predictions and explanations of different models to be clustered together to aid analysis and guide further model development. Utilizing



**Fig. 6** The HAMMER architecture and some of the scenarios it has been applied to (object grasping, task learning (e.g. Towers of Hanoi)) on multiple robotic systems. Tentative brain areas that model's operations can be mapped to are indicated (PFG: inferior parietal convexity, STS (Superior Temporal Sulcus), areas F5, BA 44/45/46). Key elements include comparators at multiple levels of abstraction that compare the

output of the forward models against the visual information (the upper half of the architecture), and a learning route (bottom half) that learns new inverse models by extracting visual information from the human demonstration. The architecture has been used also in postural sequence learning scenarios, for example dance learning (Lee et al. 2012)

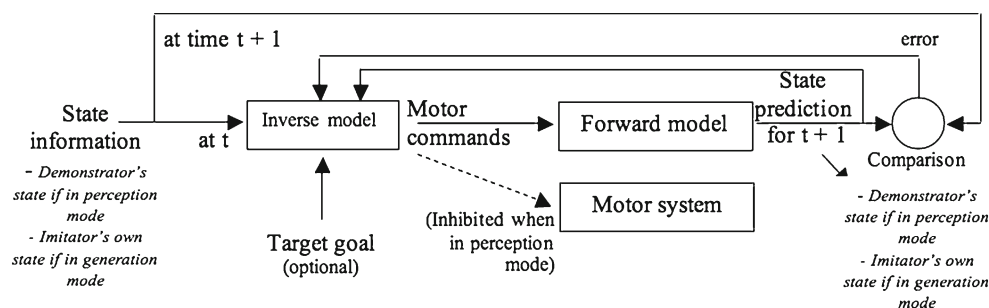
BODB, Tables 9 and 10 display the current list of summaries of experimental data (SEDs) used to build HAMMER and the current list of SEDs and Simulation Results (SSRs) used to test HAMMER, respectively. These entries, when clicked in the BODB integrated environment, further expand to reveal detailed descriptions.

Table 11 lists the HAMMER Brain Operating Principles as captured by its BODB description. Among them, the most important ones are the existence and utilization of internal (inverse and forward) models, a winner-take-all mechanism for selecting among the alternative models based on the confidence (a result of their prediction accuracy), and the combination of top-down and bottom-up control of attentional resources.

## Discussion

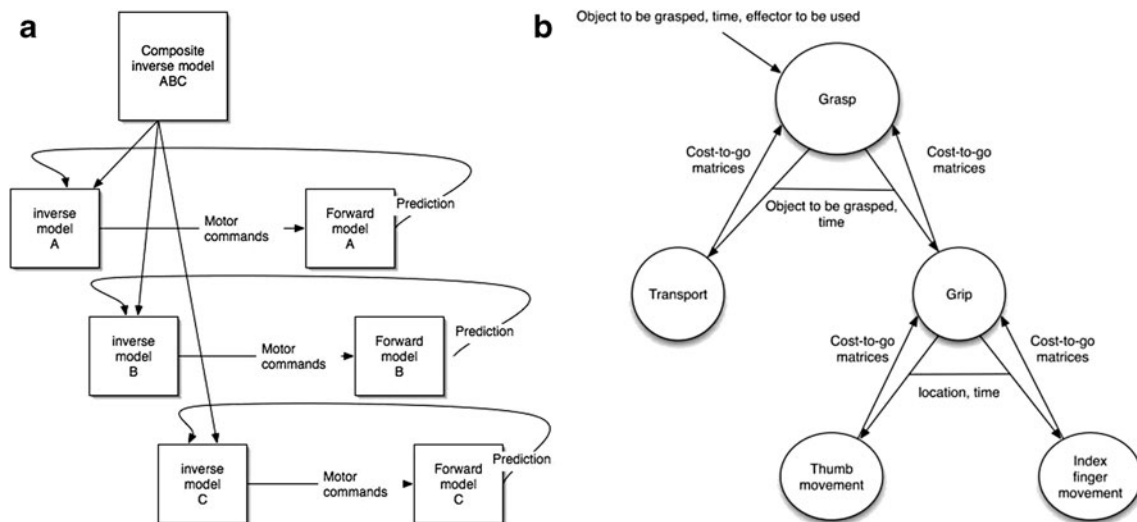
The working hypothesis of the paper is that using a neuroinformatics database tool, the Brain Operation Database (BODB), will allow us to compare diverse models of a brain process such as the mirror system. The four models included in the comparison cover a diverse spectrum, from a pure grasping model (FARS) to the two MNS models, to the HAMMER imitation robot model.

BODB allows us comparison at three levels: first, we can directly compare the BOPs used to create each model (Table 12). Both the MNS and HAMMER models advocate close sensorimotor coupling, with large overlapping areas between action perception and generation, and due to the nature of the



**Fig. 7** The architecture's basic building block, an inverse model (akin to the concepts of controller, schema, behaviour, and motor plan) paired with a forward model (akin to the concepts of predictor, and critic (in actor-critic architectures)). Multiple of these pairs operate in parallel in the generative route of the HAMMER architecture, providing multiple

hypotheses with respect to the nature of the ongoing action demonstration). The predictions of these hypotheses are compared against the observed states, and the inverse/forward model pairs that predict accurately receive reinforcement signals



**Fig. 8** (left) HAMMER can build hierarchies of composite inverse & forward models by arranging them sequentially or in parallel (Demiris and Johnson 2003). (Right) An example of a hierarchical organization for

a grasp movement (Demiris and Simmons 2006) with the exact nature of hierarchical organizations of movements in the brain still an open research issue

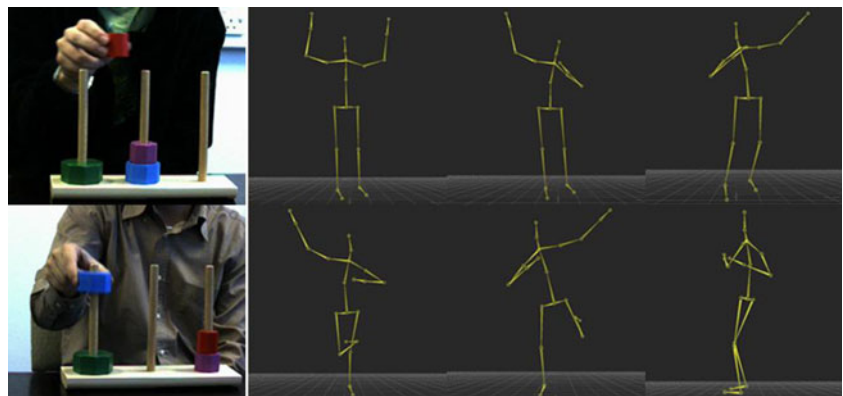
process they are modeling, both engage in temporal pattern processing. In both models, working memory is used to store hand and object positions, and recall them when such information is no longer available (for example with reach movements to hidden (but previously shown) objects).

Secondly, we can directly compare the experimental data (SEDs) that the models have been mapped to. In Figs. 10 and 11, we can see BODB's model benchmarking tool display, with both the SEDs used for building the model (direct "support" or the more general "scene setting"), as well as the biological SEDs that have an explanation offered by the models.

The BODB model benchmarking results naturally reflects the models' different starting points, and their different *raison d'être*. The FARS/MNS models were designed to closely match detailed anatomical connections in the grasping

process and its mirroring, so have direct support from biological SEDs on connectivity (area projection entries for example between AIP, F5, and STS), and should be viewed as a neuroanatomically accurate monkey mirror neuron model; on the other side, HAMMER was designed with (and can directly support) the goal of enhancing the capabilities of robotic systems to observe and imitate human actions and action sequences, and as such has direct support from behavioural and functional SEDs (for example, congruency data on F5 mirror neurons, or the hierarchical organization of behavior in primates). From the modelling perspective, HAMMER should be seen as a more general model that can be instantiated as a human or monkey model by constraining the inverse and forward models that are allowed, but at the cost of reduced neuroanatomical specificity. MNS2 is unique among the models in addressing the audiovisual properties of the mirror

**Fig. 9** Inputs to the passive route in various scenarios: (Left) A demonstration of a solution to the 3-disk Tower of Hanoi is shown (left-up), learned, and used to reason about the demonstration of a 4-disk version (left-down) (Right): Sequences of dance movements with recursive components can be learned (human data acquired with an Optitrack motion capture system (Lee et al. 2012))



Symbol	Actions	
L	Lift a disk	S→AABCC (6.99) [0.60]
D	Drop a disk	ASC (2.66) [0.23]
A	Move between 1 and 2	AASCC (0.93) [0.08]
B	Move between 1 and 3	CS (0.64) [0.05]
C	Move between 2 and 3	AABAC (0.46) [0.04]

**Fig. 10** Example of primitive inverse models (*left*) and grammatical structure learned, along with associated probabilities for these sequences to occur during the solution of a task (*right*) (from Lee et al. 2012)

neurons, while HAMMER's dual route approach draws interesting parallels to human visuomotor apraxia data (Demiris and Hayes 1999). While each individual model offers some unique correlations directly derivable from its design, interesting discussion points directly arise from entries that are common among all models.

- **F5 mirror properties:** in all models, certain nodes are active both during observation and execution of an action (the fundamental property of mirror systems). However, only the HAMMER model ascribes a functional role to mirror signals, an indication of how confident the inverse-forward pairs are that they can predict the ongoing action, although MNS was developed within a conceptual framework that assumes they are used for feedback-based control.
- **F5 Transitive action selectivity:** in all models, grasp pantomime operations (i.e. hand motion without the presence of a target object) do not activate grasp mirror neurons, either due to the nature of the design of the underlying inverse models (HAMMER), or the training that the underlying neural networks have received.
- **F5 grasp selectivity:** in all models, different hand movements (grasping, manipulating, among others) activate different F5 mirror neurons. This is due to the design of each model, i.e. localist coding in MNS and controller design in HAMMER.
- **F5 canonical neurons visual properties:** in all models, observation of graspable objects activates motor representations that would be active if the agent performed the grasping action.
- **STS shape-selective cells:** in all models, STS visual cells respond similarly to the sight of the agent's own movements as well as those of others. This highlights the fundamental difference between these cells and F5 mirror neurons and that is that mirror neurons respond during performance of grasps in the dark (without visual feedback from the hand). This makes sense given that the STS is not traditionally considered a motor region, directly feeding into motor pathways.

BODB allows a third method for comparing models through prediction entries. We have seen examples of the

former in the common entries list earlier; both models offer a range of (potentially testable) predictions, which we will examine below:

- **Response to different speed profiles of the demonstration:** both models predict that the mirror neurons will be sensitive to the speed profiles of the demonstrated act, although the two models derive this from different processes. In MNS2 the progression of the observed reach is encoded in the recurrent connection weights so reaches performed at speeds the model has not seen before disrupt this encoding. In HAMMER this is due to an inability of the inverse and forward models to accurately predict the sensory consequences of an observed fast reach. This property of MNS2 was not previously investigated, but was spurred by this comparison. In this case, the model comparison tools provided by BODB inspired simulations on an existing model in order to determine if it could make the same predictions as another model. This prediction that both families of models offer can be tested by demonstrating (to human or animal subjects) actions at artificially manipulated varying speed profiles (including unattainable ones), and recording the neuronal response.
- **Trainability of new mirror neurons:** both models predict that new mirror neurons can be learned (adding new inverse models through the passive route (HAMMER) and training the recurrent neural network with new sequences (MNS2). A particular area of interest is intransitive movements since although monkeys cannot perform them, humans do. In its current implementation, MNS assumes that all the inputs to the model come from regions that analyze the hand state in relation to object affordances. One possibility might be that the human STS has more advanced representations of observed actions in multiple reference frames, and in humans there is a direct connection between the STS and the IFG (arcuate fasciculus) whereas in monkeys there is not. Alternatively, it might be more conceptual—in humans, we can understand abstract goals (intransitive actions) whereas monkeys may not. This might be a connection between mentalizing regions (Van Overwalle and Baetens 2009) figuring out why someone might be making an intransitive action and the MNS modeling the action plan. However monkeys do not have comparable mentalizing capacities, therefore there is no top down connection to the MNS.
- **Altering the visual properties of existing mirror neurons:** in the HAMMER model, visual states are predicted through internal execution of the corresponding inverse model; embodiment constraints are imposed on their execution and constraint the range of allowable visual properties (an example of this was given earlier, that of reduced response to unachievable speed profile). In order to modify the visual properties, execution for the purposes of



**Table 9** SEDs used to build HAMMER as captured by its BODB-based description

Name	Description	Relationship	Relevance narrative
Action goal encoding at distinct levels of abstraction	Grasping Neurons of Monkey Parietal and Premotor Cortices Encode Action Goals at Distinct Levels of Abstraction during Complex Action Sequences	support	Inverse models are built hierarchically with higher inverse models encoding more distal goals while constituent inverse models encoding more proximal goals.
Area 46 projection to F5	Area 46 is reciprocally connected with F5 (Luppino et al., 1990)	scene setting	Inverse models in F5 make use of working memory in area 46
Area 46 working memory	Area 46 is implicated in working memory in tasks requiring that information be held during a delay period (Quintana & Fuster, 1993)	scene setting	Model requires that information (e.g. location of objects and hands) is held in working memory and retrieved if actual measurements are no longer available
BA 46 role in working memory	Brodmann area 46 is typically implicated in working memory (Courtney et al., 1998; McCarthy et al., 1994)	scene setting	Model contains a working memory holding object and limb positions
Broca's area involved in the hierarchical organisation of behaviour	Broca's area process hierarchically structured behaviours regardless of their temporal organization (fMRI data)	scene setting	The model extracts and retains a hierarchical representation of actions and their primitive components
F5 canonical visual properties	Some F5 cells respond to the observation of graspable objects (Rizzolatti et al., 1988; Murata et al., 1997)	support	Observation of objects will trigger the eligibility flag of inverse models that can act on these objects
F5 grasp selectivity	Different classes of F5 neurons discharge during different hand movements (grasping, holding, tearing, manipulating) and can be selective for either precision grip, finger prehension, or whole hand prehension (Rizzolatti et al., 1988)	support	Different types of grasps are encoded in different inverse models, whose execution is compared against the observed data, increasing their confidence levels
F5 mirror—broadly congruent	Some F5 mirror neurons the congruence between preferred observed and performed actions is quite loose. For example, a mirror neuron may fire during performance of precision pinches and observation of power grasps	support	The model contains multiple inverse models, each operating on its own type of states (e.g. next hand position, or relation between hand and object position, or whether an object property has changed (e.g. has been lifted).
F5 mirror—partially hidden grasps	Mirror neurons selective for grasping do not respond to pantomimed grasps but will respond to grasps that are partially hidden by a screen if an object was recently seen behind it. These neurons will not respond to a grasp directed behind the screen if no object is known to be behind it.	support	The model holds a short term working memory holding the locations of different objects. It can operate from memory instead of needing to saccade there again.
F5 mirror—strictly congruent	For some mirror neurons, not only must the general action (e.g., grasping) match, but also the way the action is executed (e.g., power grasp) must match as well.	support	Different types of grasps are encoded in different inverse models, whose execution is compared against the observed data, increasing their confidence levels.
F5 mirror properties	Mirror F5 neurons respond during observation of a grasping	support	Inverse models are active both during generation and

**Table 9** (continued)

Name	Description	Relationship	Relevance narrative
	movement (di Pellegrino et al., 1992; Rizzolatti et al., 1996; Gallese et al. 1996). Like other F5 neurons, mirror neurons are active when the monkey performs a particular class of actions, such as grasping, manipulating, and placing. However, in addition the mirror neurons become active when the monkey observes the experimenter or another monkey performing an action.		perception of actions; the level of activation of the inverse models is dependent on the confidence of these inverse models that they can explain the demonstrated actions, i.e. is dependent on the quality of predictions (i.e. the computation of the error between the predicted and actual next states) they generate through the forward models.
IPL/PFG mirror neurons	Neurons in the inferior parietal lobule (area PFG) show mirror properties for hand/mouth motor acts (e.g. monkey reaches and grasps food and eats it)	support	Higher inverse models encode combinations of primitive inverse models to form a complete complex act; they are active both when executing and when observing such acts
Mountain Gorilla foraging behavior is hierarchically organized.	Wild mountain gorillas of central Africa are shown to execute hierarchically organized behavioral programs during complex nettle processing and foraging. These 'programs' contain recursive and optional subroutines that execute adaptively depending on local circumstances and skill (Byrne & Russon, 1998).	scene setting	The stochastic context free grammar representations used in the grammatical learning of the task can support recursion and optionality
Mountain Gorillas can imitate hierarchical organization of behavior	Mountain Gorillas of central Africa are shown to be able to imitate the hierarchical organization of nettle processing behavior from observation of adult demonstrators (Byrne & Russon, 1998).	support	The model can extract the hierarchical structure of the observed task and imitate it
PF/PFG goal-selective neurons	Area PF/PFG in the inferior parietal lobule contains neurons selective for the final goal of an action (eating vs. placing).	scene setting	Higher inverse models encode combinations of primitive inverse models to form a complete complex act
STS projection to area 46	There are multiple connections from different areas of STS to area 46 (Seltzer & Pandya, 1989)	support	The information held in working memory are obtained through the STS
STS shape-selective cells	Shape-selective cells respond selectively to hand-object interactions and respond similarly to the sight of the monkeys own movements as well as those of others.	support	Object-hand relationships are computed and fed to the inverse models irrespective of who the agent is
STS visual response	STS contains neurons that respond to the observation of actions such as walking, limb movements, and grasping	support	The model relies on visual information extracted from the demonstration on demonstrator's postural states and body-object relations.
VIP location selectivity	Area VIP represents the location and orientation of target objects in a broadly tuned population code in a peripersonal	scene setting	The locations of objects are calculated from visual information and used by the

**Table 9** (continued)

Name	Description	Relationship	Relevance narrative
Visuoimitative apraxia data	coordinate system (Colby et al., 1993)	support	model to perform the reaching movements
	Goldenberg and Hagmann 1997 demonstrate that, following brain damage, certain patients lose their ability to imitate novel meaningless gestures		Damage to the model's passive route will hinder the observer's ability to imitate novel gestures while allowing imitation of known ones through the active predictive route.

**Table 10** SEDs and SSRs used to test HAMMER

Name	Relationship	Relevance narrative
Building a motor simulation de novo: observation of dance by dancers	explanation	The model can acquire new postural sequences by demonstration and incorporate them in its repertoire as new inverse models; these are subsequently available to the mirror system for perceiving demonstrated sequences
<b>SED</b>		
<b>Name</b>	<b>Description</b>	<b>Description</b>
Building a motor simulation de novo: observation of dance by dancers.		Complex motor resonance (to complex dance sequences) can be built de novo over 5 weeks of rehearsal
<b>SSR</b>		
<b>Name</b>	<b>Description</b>	<b>Description</b>
Incorporation of new postural sequences		The model can learn new postural sequences, incorporate them in its inverse model repertoire, and make them available to the mirror system for recognition of the new sequences
F5 mirror response to intransitive actions	explanation	Grasp inverse models require an object in order to generate the motor commands to grasp it and require the distance between the arm and the object to monitor its progress.
<b>SED</b>		
<b>Name</b>	<b>Description</b>	<b>Description</b>
F5 mirror—transitive action selectivity		To be triggered, the mirror neurons require an interaction between the hand motion and the object. The vision of the hand motion or the object alone does not trigger mirror activity (Gallese et al. 1996; Umiltà et al., 2001).
<b>SSR</b>		
<b>Name</b>	<b>Description</b>	<b>Description</b>
F5 pantomimed grasp		The grasp inverse model requires the distance between the object and the hand; without an object this cannot be computed and the corresponding model's confidence levels will not increase.

**Table 10** (continued)

Name	Relationship	Relevance narrative
Impaired response to novel gestures	explanation	Damage to the model's learning (passive) route will impair the model's ability to imitate actions that are not part of its repertoire
<i>SED</i>		
<b>Name</b>		<b>Description</b>
Visuoimitative apraxia data		Goldenberg and Hagmann 1997 demonstrate that, following brain damage, certain patients lose their ability to imitate novel meaningless gestures
<i>SSR</i>		
<b>Name</b>		<b>Description</b>
Damage to passive route		Damage to the non-predictive route does not allow the model to learn and incorporate new actions in its repertoire, without harming the model's ability to recognise and imitate known actions.

internal simulation will need to be decoupled from any embodiment constraints. The same holds true for the MNS/MNS2 models since they use efference copies of motor commands as training signals for mirror neurons.

Finally, BODB can be used to discover and document both data and modeling features not covered by the MNS and HAMMER models, which suggest directions for future modeling efforts.

- **Suppressing automatic imitative tendencies:** Recently a population of pyramidal tract neurons has been found in F5 which fire during performance of a grasp, but whose activity is suppressed during observation of the same grasp (Kraskov et al. 2009), suggesting that they may be involved in the suppression of automatic imitative responses. However, monkeys have shown little, if no, imitative ability. While humans exhibit disorders such as echopraxia (compulsive imitation) that may indicate inhibition of automatic imitation in the normally functioning brain, the role of these pyramidal tract mirror neurons in the monkey is unclear. Models that link mirror neuron activity to motor output in a biologically realistic way might suggest a functional role for neurons with these properties.
- **Planning a nonimitative response to an observed action:** As mentioned in the Background section, recent data on the monkey mirror system suggest the role it may play in understanding the goal of an observed action and planning a response. Early research on macaque mirror neurons focused on the representation of the action in relation to the object's affordances, however later studies have examined the effect of the subjective value of grasped

object (Caggiano et al. 2012) and the metric and functional workspace that the observed action takes place in (Caggiano et al. 2009). Future models that take these data into account could provide an explanation for these results in terms of understanding the context in which an observed action takes place, allowing a meaningful response to be planned.

- **Understanding the intentions behind observed actions:** Recent data suggests that the human mirror system may interact with a mentalizing system for observing mental states (Spunt et al. 2011). The Mental State Inference Model (Oztop et al. 2005) suggests a mechanism where a mirror system model interacts with a simple mentalizing model to infer the intention behind an observed action. Future mirror system models which are extended to account for human data must include the interaction between the mirror system and other networks such as mentalizing and emotional systems (Spunt and Lieberman 2012).

## Conclusions

In this paper we utilized a neuroinformatics database tool, BODB, to examine the commonalities and differences between diverse models of the mirror system. Such comparisons given the different starting points of the examined models are rare, but useful if we are to determine which biological aspects require further modeling, as well as which biological principles can be effectively utilized in robotic systems. Gaps and predictions in computational models could in principle also drive further neuroscience experiments.



**Table 11** HAMMER's seven Brain Operation Principles (BOPs) as captured by its BODB description

Name	Description	Relationship
Sensorimotor Coupling	The brain has representations that bridge between perceptual contents and action plans. Such common codes integrate action and perception. Perceptual stimuli are interpreted in part on the motor representations which are commonly associated with the corresponding sensory experience.	Visual perceptual stimuli are coupled with the motor programs that can generate them
Temporal Pattern Processing	Temporal pattern processing is important for various intelligent behaviors, including hearing, vision, speech, music, and motor control. Because we live in an ever-changing environment, an intelligent system must encode patterns over time, recognizing and generating temporal patterns.	The input for the model, an action demonstration, is a temporal pattern processed in real time
Attention	In general, sensory systems deliver far more data than is relevant to current behavior and memory storage of the organism. Attention is the process of selecting, more or less successfully, of focusing on currently relevant data.	The model uses a top-down attention mechanism to direct the sensorimotor resources of the observer to the relevant parts of the demonstration
Extraction of Abstract Structure	One aspect of cognitive sequence processing is the extraction of temporal, hierarchical and other abstract structure from, patterns. This is essential because some of the most useful information will be in the hierarchical structuring.	The model extracts abstract structure as hierarchical tasks decompositions (for example using stochastic context free grammar rules). The result can be used to generalise the input description to different tasks with similar abstract structure
Internal Models	Internal models are neural mechanisms that can mimic properties of the external world as a basis for recognizing objects, and planning actions and interactions. In motor control, the focus is on how a system will respond to commands or actions. Forward models capture the forward or causal relationship from inputs to the system to the resultant outputs (which may depend on the internal state of the modeled system. They can be used to stabilize sensory feedback oscillations which are caused by delays. Inverse model represent the inverse transformation from the desired movement of the controlled object to motor commands serving to attain these movement goals.	The dual route model's predictive route is composed by multiple inverse-forward model pairs, running in parallel, competing for which model will explain best the sensory information contained in a demonstrated action.
Top-Down/Bottom-Up Hybridization	Many cognitive functions depend on the interaction between top-down knowledge driven processes and bottom-up sensor-driven processing.	The attention mechanism of the model operates in concurrent top-down (which information do different models require) and bottom-up (salient information extraction) modes, to determine in what aspects of the input will the model focus the sensorimotor and information processing resources of the observer on.
Winner-Take-All (WTA)	A Winner-Take-All network responds (after a delay) to an array of inputs of different intensity with an output pattern which under normal circumstances encodes the input which had the greatest intensity.	The multiple inverse and forward models compete to explain the ongoing demonstration, and the winner (the pair with highest confidence) wins, and is selected to be imitated

**Table 12** Common brain operation operations across the MNS and HAMMER models

Name	Description
Sensorimotor Coupling	The brain has representations that bridge between perceptual contents and action plans. Such common codes integrate action and perception. Perceptual stimuli are interpreted in part on the motor representations which are commonly associated with the corresponding sensory experience.
Temporal Pattern Processing	Temporal pattern processing is important for various intelligent behaviors, including hearing, vision, speech, music, and motor control. Because we live in an ever-changing environment, an intelligent system must encode patterns over time, recognizing and generating temporal patterns.
Working Memory	Working memory consists of contextual information that is stored for the duration of the execution span of the current task.

Our analysis revealed that interesting new directions for modeling (as well for neuroscience experiments) revolve around the issue of how certain mirror neuron properties can be learned or altered. Our analysis brought into light a number of further desiderata from experiments and neuroinformatics databases that modelers could utilize in their mirror system modeling efforts, as well as a number of potential extensions for using this methodology as a blueprint for advancing other areas of neuroinformatics.

With respect to the mirror systems specifically, the following desiderata become evident:

- Response data of F5 mirror neurons (and other mirror areas) to actions performed in variable speed profiles; these will provide the validation of the speed-sensitivity predictions that both MNS and HAMMER have put forward.
- Kinematic recordings of observed and performed actions in the mirroring scenarios of current electrophysiological recordings; these will allow the exact correlation and automatic tuning of inverse/forward model activity in the HAMMER family of models, and the automatic training of the recurrent neural networks in the MNS family of models.
- Multi-electrode recordings from mirror and related areas: useful information from these sets would be the extraction of timing information on the propagation of neural activity in these areas, as well as determination of the hierarchical organization (if any) of these areas.
- Response data of F5 mirror neurons to multiple concurrent demonstrations of different actions; these will illuminate

whether or not there is a parallel activation of multiple hypotheses, as the HAMMER model suggests.

- F5 recording experiments in which the monkey observes an action and is then required to make some response indicating that they understand either a) the effects of the action in relation to themselves, and/or b) the actor's intention.

The BODB approach we utilized in this paper to compare mirror system models can also serve as a blueprint for a new approach to advance neuroinformatics; from the experience gained in the reported work, the following desiderata were discovered as important for future extensions of BODB:

- Inverse functionality, i.e. supplying a description of desired operation (for example, coordinate transformation, or responses to certain type of stimuli) and receiving entries of brain areas that are known to perform such operations. This would add a new route to constructing brain model outlines from functional requirements.
- Model versioning should capture the way in which new models build upon (parts of) one or more older models. A computational model of the brain, in general, is not a static and isolated component. In fact, models generally extend from precursor models with knowledge generalized from new experimental data or make use of many smaller models (namely *modules*) to explain brain function more explicitly, if the information makes a new model more realistically fitting to that brain mechanism. Consider, for example, the FARS model developed by Fagg and Arbib (1998). With support from new experimental data, this model was extended to implement many successors; for example, the MNS model of the mirror system for monkey grasping by Oztop and Arbib (2002). Both a more realistic learning model and the need to address data on audiovisual mirror neurons and the grasping of hidden objects then led to the improved MNS2 model (2007).
- Methods for verifying which connectivity data from multiple nomenclatures refer to projections to and from the same underlying brain region.
- Methods for formalizing the simulation protocol used to test a model and the corresponding experimental protocol that could be used to test its predictions. SED versioning (analogous to model versioning mentioned earlier) would also allow the faithful replication of model results even if new data
- Methods for relating and contrasting BOPs between different models. For example, the MNS2 model utilizes the following BOPs that HAMMER does not: dynamic remapping, Hebbian learning, sensor fusion, and supervised learning. Some of these such as supervised learning are more implementation details, while others such as working memory and sensor fusion allow the core mirror system model to interact with other systems. Similarly,

SEDs	Dual-route imitation model (Demiris)	FARS (Fagg - Arbib)	Mirror Neuron System (MNS) (Oztop - Arbib)	Mirror Neuron System 2 (MNS2) (Bonaiuto - Rosta - Arbib)
AIP Motor Properties		Explanation		
AIP Grasp Selectivity		Explanation	Support	Support
Separation of dorsal and ventral visual streams		Scene Setting	Scene Setting	Scene Setting
AIP Object Selectivity		Explanation		
F5 mirror properties	Support	Scene Setting	Explanation	Explanation
F5 mirror - strictly congruent	Support		Support	Support
F5 mirror - broadly congruent	Support		Scene Setting	Scene Setting
STS visual response	Support		Support	Support
F5 audiovisual mirror neurons				Explanation
F5 mirror - transitive action selectivity	Explanation		Explanation	Explanation
cIPS projection to AIP		Support	Support	Support
Projection from AIP to F5		Support	Support	Support
Projection from F4 to F1		Scene Setting	Support	Support
Projection from F5 to F1		Scene Setting	Support	Support
VIP location selectivity	Scene Setting	Scene Setting	Support	Support
F4 reach selectivity		Scene Setting	Support	Support
Projection from F5 to F5		Support	Support	
Projection from VIP to F4		Scene Setting	Support	
F5 grasp selectivity	Support	Support	Support	Support
F5 mirror - partially hidden grasps	Support			Explanation
STS shape-selective cells	Support		Support	Support
STS projection to area 46	Support			
F5 canonical visual properties	Support	Support	Support	Support
IT object selectivity		Scene Setting		
Projection from 46 to F6		Support		
Projection from AIP to F4		Support		
Projection from F5 to AIP		Support		
Projection from F5 to V46		Support		
Projection from F6 to F5		Support		
Projection from V46 to F5		Scene Setting		
SMA role in complex movements				
Projection from LIP to 7a			Support	Support
Projection from MIP to 7a			Support	Support
Projection from PF to F5			Support	Support
Projection from PIP to 7a			Support	Support
Projection from STS to PG			Support	Support
Projection from VIP to 7a			Support	Support
PIP object selectivity		Support	Support	Support
Projection from 7a to 7b			Support	
Projection from AIP to 7a/7b			Support	
Area 7b connection with AIP				Support
Auditory cortex projection to F5				Support
Indirect auditory cortex projection to premotor cortex				Support
Projection from 46 to 7a				Support
Projection from 46 to 7b				Support
Projection from 7a to 7b				Support
Projection from LIP to 46				Support
Projection from STS to 46				Support
Projection from VIP to F4				Support
F2 arbitrary stimulus association		Support		
F5 grasp phase selectivity		Support		
AIP Phasic Activity		Explanation		
IPL visual properties				Explanation
BA 46 role in working memory	Scene Setting			Scene Setting
PF/PFG goal-selective neurons	Scene Setting			
Visuoimitative apraxia data	Explanation			
Area 46 working memory	Scene Setting	Support		
Building a motor simulation de novo: observation of dance by dancers.	Explanation			
Nonprimary auditory cortex encoding				Support
Primary auditory cortex organization				Support
Working memory activity in area 46 and inferior parietal				Support
BG action selection		Support		
F6 go signal detection		Support		
Pre-SMA role in complex sequences		Scene Setting		
Mountain Gorillas can imitate hierarchical organization of behavior	Support			
Area 46 projection to F5	Scene Setting			
Action goal encoding at distinct levels of abstraction	Support			
Broca's area involved in the hierarchical organisation of behaviour	Scene Setting			
IPL/PFG mirror neurons	Support			
Mountain Gorilla foraging behavior is hierarchically organized.	Scene Setting			

**Fig. 11** Model benchmarking display allows for rapid comparison of differences and complementarities of models

HAMMER is based on BOPs not used by the MNS2 model such as attention, extraction of abstract structure, internal models, top-down/bottom-up hybridization, and winner-take-all which allow it to model the interaction of the mirror system with the rest of the motor system for the purposes of imitation.

- Methods for constructing comparative schematics that visualise similarities and differences (for example, overlaps such as common connectivity between brain areas) between models; this will also allow the construction of larger scale models through the merging of component ones, and allow the identification of missing blocks for further research and modelling).

Regardless of these additions, the BODB approach of categorizing and clustering models and data along the axes of Brain Operating Principles, Summaries of Experimental Data and Summaries of Simulation Results and predictions, makes BODB a valuable neuroinformatics database tool for the comparison of diverse neuroinspired computational and robotic models.

### Information Sharing Statement

The following resources utilized in this research are freely available to the general public:

- The Brain Operations Database (BODB) infrastructure is available at <http://bodb.usc.edu>.
- Robot middleware (with example implementations for the NAO and icub humanoids) based on the HAMMER architecture is available at the Personal Robotics Laboratory website <http://www.imperial.ac.uk/personalrobotics>.

**Acknowledgements** YD's contributions were partially supported through EU FP7 grants EFPA (grant FP7-ICT-270490) and ALIZE (grant FP7-ICT-248116).

### References

Arbib, M., Bonaiuto, J. (2013). The Neuroinformatics of Brain Modeling, Part II: The Brain Operation Database, BODB. Neuroinformatics in press.

Arbib, M., Plangprasopchok, A., Bonaiuto, J., Schuler, R. (2013). The neuroinformatics of brain modeling, part I: from empirical data to brain operating principles. Neuroinformatics in press.

Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57, 469–483.

Aziz-Zadeh, L., Sheng, T., Liew, S. L., & Damasio, H. (2012). Understanding otherness: the neural bases of action comprehension and pain empathy in a congenital amputee. *Cerebral Cortex*, 22, 811–819.

Bonaiuto, J., & Arbib, M. (2010). Extending the mirror neuron system model, II: what did I just do? A new role for mirror neurons. *Biological Cybernetics*, 102, 1–19.

Bonaiuto, J., Rosta, E., & Arbib, M. (2007). Extending the mirror neuron system model, I—audible actions and invisible grasps. *Biological Cybernetics*, 96, 9–38.

Bota, M., Arbib, M. A. (2004). Integrating databases and expert systems for the analysis of brain structures: connections, similarities, and homologies. In *Neuroinformatics*, (vol. 2, pp 19–58). United States.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: a hierarchical approach. *Behavioral and brain sciences*, 21(5), 667–684.

Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of. *Science*, 324, 403–406.

Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., & Casile, A. (2011). View-based encoding of actions in mirror neurons of area f5 in macaque premotor. *Current Biology*, 21, 144–148.

Caggiano, V., Fogassi, L., Rizzolatti, G., Casile, A., Giese, M. A., & Thier, P. (2012). Mirror neurons encode the subjective value of an observed action. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 11848–11853.

Castiello, U., & Jeannerod, M. (1991). Measuring time to awareness. *Neuroreport*, 2(12), 797–800.

Cavada, C., & Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *Journal of Comparative Neurology*, 287(4), 422–445.

Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of neurophysiology*, 69(3), 902–914.

Cross, E. S., Hamilton, A., & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *NeuroImage*, 31, 1257–1267.

Demiris, J., Hayes, G. (1999). Active and Passive Routes to Imitation, Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts. K. Dautenhahn and C. Nehaniv (eds.), (pp. 81–87), Edinburgh, UK, April 6–9, 1999.

Demiris, J., Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components; a biologically plausible computational model. In *Imitation in animals and artifacts*. MIT Press.

Demiris, Y., & Johnson, M. (2003). Distributed, predictive perception of actions: a biologically inspired robotics architecture for imitation and learning. *Connection Science*, 15, 231–243.

Demiris, Y., & Khadhour, B. (2006). Hierarchical attentive multiple models for execution and recognition of actions. *Robotics and Autonomous Systems*, 54, 361–369.

Demiris, Y., & Simmons, G. (2006). Perceiving the unusual: temporal properties of hierarchical motor representations for action perception. *Neural Networks*, 19, 272–284.

Demiris, J., Rougeaux, S., Hayes, G., Berthouze, L., Kuniyoshi, Y. (1997). Deferred imitation of human head movements by an active stereo vision head. IEEE. pp 88–93.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp Brain Res*, 91(1), 176–80.

Erlhagen, W., Mukovski, A., & Bicho, E. (2006). A dynamic model for action understanding and goal-directed imitation. *Brain Research*, 1083, 174–188.

Fadiga, L., Craighero, L., & Ausilio, D. (2009). Broca's area in language, action and music. *Annals of the New York Academy of Sciences*, 1169, 448–458.

Fagg, A. H., & Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks*, 11, 1277–1303.

Ferrari, P. F., Rozzi, S., & Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience*, 17, 212–226.



- Fleischer, F., Christensen, A., Caggiano, V., Thier, P., & Giese, M. A. (2012). Neural theory for the perception of causal actions. *Psychological Research*, 76, 476–493.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain: A Journal of Neurology*, 119(Pt 2), 593–609.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Science*, 8, 396–403.
- Goldenberg, G., & Hagmann, S. (1997). The meaning of meaningless gestures: a study of visuo-imitative apraxia. *Neuropsychologia*, 35, 333–341.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1), 20–25.
- Haruno, M., Wolpert, D. M., & Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation*, 13, 2201–2220.
- Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience*, 17, 1123–1128.
- Hourdakis, E., Savaki, H. E., & Trahanias, P. (2011). Computational modeling of cortical pathways involved in action execution and action observation. *Neurocomputing*, 74, 1135–1155.
- Iacoboni, M. (2005). Neural mechanisms of imitation. *Curr Opin Neurobiol.*
- Johnson, M., Demiris, Y. (2005). Perceptual Perspective Taking and Action Recognition. *International Journal of Advanced Robotic Systems*, 2:4, pp 301–308.
- Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Science*, 8, 501–507.
- Kotter, R. (2004). Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac database. *Neuroinformatics*, 2, 127–144.
- Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron*, 64, 922–930.
- Lee, K., Kim, T.-K., Demiris, Y. (2012). Learning reusable task components using hierarchical activity grammars with uncertainties. In: *Robotics and Automation (ICRA)*, 2012 I.E. International Conference on, pp 1994–1999.
- Lyon, R. (1982). A computational model of filtering, detection, and compression in the cochlea. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*. 7:1282–1285.
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of neurophysiology*, 78(4), 2226–2230.
- Oztop, E., & Arbib, M. A. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87, 116–140.
- Oztop, E., Wolpert, D., & Kawato, M. (2005). Mental state inference using visual control parameters. *Brain Research. Cognitive Brain Research*, 22, 129–151.
- Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: a computationally guided review. *Neural Networks: The Official Journal of the International Neural Network Society*, 19, 254–271.
- Oztop, E., Kawato, M., Arbib, M. A. (2012). *Mirror Neurons: Functions, Mechanisms and Models*. Neuroscience Letters.
- Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, 146, 87–113.
- 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. *Brain Res Cogn Brain Res*, 3(2), 131–41.
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., & Tanaka, Y. (1997). The TINS Lecture The parietal association cortex in depth perception and visual control of hand action. *Trends in neurosciences*, 20(8), 350–357.
- Sausser, E. L., & Billard, A. G. (2006). Parallel and distributed neural models of the ideomotor principle: an investigation of imitative cortical pathways. *Neural Networks*, 19, 285–298.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Science*, 3, 233–242.
- Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology*, 281(1), 97–113.
- Spunt, R. P., & Lieberman, M. D. (2012). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *NeuroImage*, 59, 3050–3059.
- Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2011). Identifying the what, why, and how of an observed action: an fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, 23, 63–74.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 31(1), 155–165.
- Ungerleider, L. G., & Mishkin, M. (1982). Analysis of Visual Behavior. In D. J. Ingle, M. A. Goodale, R. J. W. Mansfield (Eds.).
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, 48, 564–584.
- Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78, 1550–1560.