The Role of Selective Attention and Action Selection in the Development of Multiple Action Capabilities

Luca Simione & Stefano Nolfi
Institute of Cognitive Sciences and Technologies, CNR
Viale S. Martino della Battaglia 44, Roma, Italy
luca.simione@, stefano.nolfi@istc.cnr.it

Abstract

In this paper we illustrate how the capacity to select the most appropriate actions when handling contexts affording multiple conflicting actions can be solved either through a selective attention strategy (in which the stimuli affording alternative actions are filtered out at the perceptual level through top-down regulation) or at later processing stages through an action selection strategy (through the suppression of the premotor information eliciting alternative actions). By carrying out a series of experiments in which a neuro-robot develops an ability to choose between conflicting actions, we were able to identify the conditions that lead to the development of solutions based on one strategy or another. Overall, the results indicate that the selective attention strategy constitutes the most simple and straightforward mechanism enabling the acquisition of such capacities. Moreover, the characteristics of the adaptive/learning process influence whether the adaptive robot converges towards a selective attention and/or action selection strategy.

1. Introduction

At every moment the environment presents animals with many opportunities and demands for actions. From a short-term behavioural perspective, this implies that the agents should be able to choose among the multiple conflicting actions available within their behavioural repertoire. From a long-term developmental perspective, this implies that they should overcome the difficulties that arise from the need to concurrently develop two interdependent abilities: the ability to identify the appropriate set of complementary actions and the ability to identify the contexts affording these alternative actions. In this paper we investigate this issue from both a short and long term perspective through a synthetic approach, with the design of a neuro-robot that develops multiple action capacities during a long-term adaptive process and selects among them while it interacts with the environment in which is situated. The rationale behind the attempt to study both how these capacities are acquired and how they are realized is that the mechanisms and processes that enable the development of actions and action selection capabilities can strongly constrain the way in which such capacities are realized in "mature" individuals.

To realize the importance of handling the conflicts that originate from the need to develop and execute competing actions we should consider that perceptual stimuli can evoke actions afforded by them automatically with little or no intention to act (Duncan-Johnson & Kopell, 1981; Gibson, 1979; Goldberg & Segraves, 1987; Miller & Hackley, 1992). The direct association between stimuli and actions, however, implies that contexts affording multiple conflicting actions might lead to the production of chaotic un-effective behaviours (Neumann, 1987). This can only be avoided through a selection process that blocks the execution of conflicting actions (Humphreys et al., 2010; Tucker & Ellis, 1998).

The identification of the mechanisms that enable natural organisms to appropriately select actions still represents an open research challenge (Seth, Prescott, & Bryson, 2011). However, the experimental evidences collected to date seem to support two types of mechanisms/strategies, that we termed 'action selection' and 'selective attention'. The action selection strategy consists in concurrently processing the perceptual information affording multiple conflicting actions in parallel by later selecting an action through a biased competitive process that operates at the pre-motor level. The selective attention strategy, instead, consists in blocking the sensory information that can elicit alternative actions already at the perceptual level through top-down feedback (i.e. through efferent connections from motor areas to perceptual areas). For this reason, in the experiments reported in this paper we provided our robot with the prerequisites necessary to develop and realize one or the other strategy and we analysed whether and in which cases the adapted robot relies on the former, the latter, or both strategies.

The study of how natural or artificial organisms can concurrently develop multiple actions skills as well as an ability to select among them represents an almost unexplored research territory. In that respect, therefore, the experiments presented in this paper constitute a rather original contribution that extends the results obtained in other pioneering studies (Petrosino, Parisi, & Nolfi, 2013; Seth et al., 2011). More specifically, as we will see, our results elucidate the relation between the characteristics and the dynamics of the learning process through which the agents develop their skills and the mechanisms/strategies that they use to master their task.

The paper is organized as follows. In the next section we discuss the neuro-psychological evidences in support of the action selection and selective attention strategies. In section 2 and 3 we describe the experimental scenario and the obtained results. In section 4 we discuss the implication of our results and we draw our conclusions.

1.1 On the role of action selection and selective attention

Cisek (2007) proposed that the tendency to execute multiple conflicting actions is solved through a biased competition operating at the pre-motor level. According to his affordance competition hypothesis, the brain processes sensory information to specify, in parallel, several potential actions that are concurrently available. These potential actions compete against each other while information is collected to bias this competition toward the selection of the most appropriate response. More specifically, the author suggests that the dorsal visual system specifies actions which compete against each other within the fronto-parietal cortex, while a variety of biasing influences are provided by prefrontal regions and the basal ganglia.

Evidences in support of this hypothesis have been collected in a series of experiments in which primate subjects were performing an instructed-delay reaching task (Cisek & Kalaska, 2005) organized into three phases. The monkeys located in front of a monitor were exposed to: (1) two circles, one red and one blue, located around a fixation point, (2) a non-spatial cue indicating the colour of the target circle, and (3) an arrow pointer. After the presentation of the arrow pointer, that acted as a ‘go’ signal, the subjects were asked to move the pointer toward the location of the target circle. By recording the state of directionally tuned cells of the premotor cortex the authors observed two simultaneous sustained signals corresponding to the two reach options during the first phase. Moreover they observed how the neural activity associated with the target and non-target circles increased and decreased respectively, during the second phase, after the presentation of the cue encoding the colour of the target circle. Overall the collected data indicate that alternative reach actions can be activated in parallel at the level of the premotor cortex and can be later enhanced or suppressed in order to execute only one of the conflicting actions. Additional evidences have been reported in a successive study (Pastor-Bernier & Cisek, 2011).

The hypothesis that action selection is realized through a selective attention process operating at the perceptual level has been proposed by Castiello (1999), according to which overt and covert attention mediate the elaboration of sensory information into motor states, i.e. filter out information affording alternative conflicting actions already at the perceptual level. This theory is supported by a series of evidences collected in reach-to-grasp experiments in which human subjects had to reach a target object in a 3D space in the presence of similar distracting objects (Bonfiglioli & Castiello, 1998; Castiello, 1996; Chieffi, Gentilucci, Allport, Sasso, & Rizzolatti, 1993; Howard & Tipper, 1997; Jackson, Jackson, & Rosicky, 1995; Tipper, Howard, & Houghton, 1998; Tipper, Howard, & Jackson, 1997). The analysis of these experiments indicates that when attention is raised on the target object, the perception of the distractors does not affect the trajectory with which the target is reached, irrespectively of the hand-distractor distance (Jackson et al., 1995) and irrespectively of the similarity between features of the target and of the distractor (Bonfiglioli & Castiello, 1998; see Castiello, 1999). On the contrary, when the attention is also raised on a distractor, its properties interfere with those of the target causing a deterioration of the action (Castiello, 1999).

Overall these results suggest that the filtering out of the perceptual information affording conflicting actions realized through selective attention at early processing levels constitutes a radical solution that eliminate the verification of conflicts in later processing stages. Additional experimental evidences that support this hypothesis has been collected by several other authors. Indeed, in a control experiment analogous to that illustrated at the beginning of this section but in which the colour of the target circle was known in advance, Cisek did not find any evidence of activity related to the relative direction of the non-relevant object in the directionally tuned neurons of the premotor cortex (Cisek, 2007). Moreover, an active inhibition of the representations of distractor/irrelevant stimuli in perceptual areas has been documented in several studies (Most, Scholl, Clifford, & Simons, 2005; Munneke, Heslenfeld, Usrey, Theeuwes, & Mangun, 2011; Payne & Allen, 2011; Seidl, Peelen, & Kastner, 2012; Triesch, Ballard, Hayhoe, & Sullivan, 2003).

The apparent conflict between the two body of evidences reviewed above can be explained by considering the characteristic of the tasks. Indeed, the evidences supporting the action selection hypothesis have been collected in experimental scenario in which the cue indicating the relevant object was available only at a certain stage of the process. On the contrary, the evidences supporting the selective attention strategy has been collected in experimental scenarios in which the participants knew the location of the target object from the beginning. Moreover, this might indicate that tasks that admit selective attention solutions are potentially compatible with action selection solutions as well.

To investigate this theme we carried out a series of experiments in which a humanoid robot controlled by a neural network is trained for the ability to look at an object of a given colour by possibly ignoring a second object with a different colour. To verify the conditions in which the robot relies on action selection and/or selective attention strategies, we provided the robot with neural network controllers that support the development of the former, the latter, or both strategies. Moreover, to disentangle the effect of the training process, we compared the results obtained in alternative experiments in which the robot was trained through a trial and error or through a learning by demonstration method.

2. Experimental Scenario

The task consisted in visually following an object of a specific colour by possibly ignoring a second object of a different colour. The colour of the object to be selected was explicitly indicated by the experimenter to the robot from the beginning of each trial. The study has been replicated six times by varying two experimental conditions: the architecture of the neural controller of the robot and the training procedure. More specifically, the neural controller could include or not a set of modulatory output units that could reduce and/or suppress the activity of selected sensory information. Moreover the robot was trained through a trial and error learning procedure or through a learning by demonstration procedure. In the former case it was rewarded on the basis of its overall ability to perform the task (i.e. on the basis of the percentage of time spent by looking at the selected object). In the latter case it received a teaching signal that specified, for each step, the state that the actuators should have to direct the sight of the robot toward the target object. In the following sub-sections we describe the method in detail.

2.1. The robot

Our robot platform was the iCub (Fig. 1) (Sandini, Metta, & Vernon, 2004), a humanoid robot provided with a rich set of sensors and actuators. In our experiments, however, only the camera located in the right eye and the two actuators that control the lateral and vertical movement of the neck were used. The experiments have been carried out in simulation by using FARSA (Massera, Ferrauto, Gigliotta, & Nolfi, 2013), a freely available developmental robotic tool that provides a realistic simulator of different robotic platforms and other tools for carrying research in adaptive robotics.

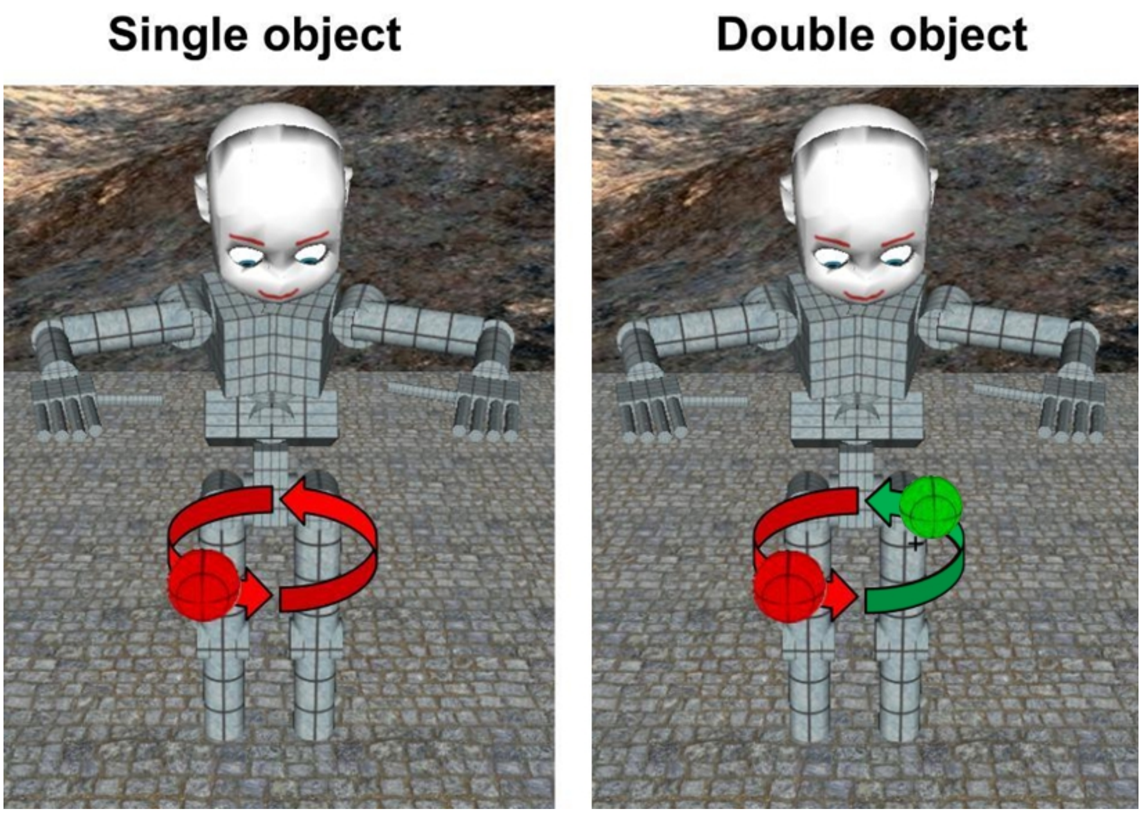

Figure 1. The simulated iCub robot. The environment of the robot might contain either a single red or green sphere (left panel), or two spheres, one green and one red (right panel). The spheres move along a circular trajectory at a fixed speed clockwise or counter clockwise.

2.2. The neural network of the robot

The robot was provided with a simple neural network including three layers. As mentioned above, the experiments have been repeated by using three different variations of the architecture (see Fig. 2). In all cases, the perceptual layer included eight visual neurons, that encoded the offset of the red and green objects along the horizontal and vertical dimension with respect to the centre of the visual field of the right eye, and two neurons that encoded the colour of the target object (i.e. the object that should be observed by the robot). The state of the visual neurons was calculated on the basis of pre-elaborated information extracted from the right camera through a simple colour blob identification software routine. The activity of the visual neurons varied in the range [0.0, 1.0] and assumed a null value [0.0] when the object of a specific colour was located over the centre of the visual field, with respect to the horizontal or vertical dimension, or when an object of a specific colour was absent from the visual field of the robot. The state of the last two neurons was set to [1.0, 0.0] when the robot was asked to look at the red object and to ignore the green object, if present, and to [0.0, 1.0] in the opposite case.

The internal layer included six neurons.

The output layer included two motor neurons that set the velocity and direction of the two motors that controlled the flexion and rotation of the neck of the robot (i.e. the direction was determined by whether the activation state was over or below 0.5 and the speed varied proportionally to the offset with respect to this neural value). The robot should bring the selected object toward the centre of its visual field by moving its head (the position of the eyes was kept fixed).

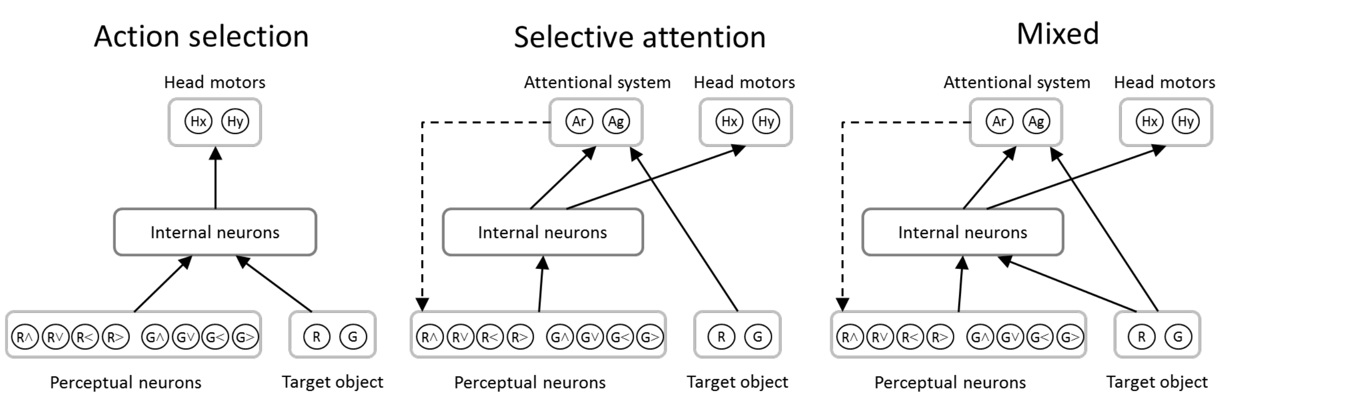

Figure 2. The three neural network architectures used in the corresponding experimental conditions. The perceptual, internal and output layers are displayed from bottom to top. The full

arrows indicate the connections among neurons. The dashed lines indicate modulatory connections.

The neurons of the internal and of the output layer were updated on the basis of a standard logistic function. The state of the neural network was updated every 100ms.

In the first experimental condition, shown on the left panel of Fig. 2, the internal neurons received connections from both the visual and target object neurons and projected connections to the motor neurons. This neural architecture included the neural circuits that could enable the development of action selection solution in which the impact of irrelevant visual information was reduced and possibly eliminated within the first and/or the second processing layer.

In the second experimental condition, shown in the central panel of Fig. 2, the network included two additional output units whose activation state at time t was used to set the gain of the red and green visual input at time $t+1$ (see the dashed line in Fig.2). Since the output of these visual neurons were calculated by multiplying their activity and gain, these additional output neurons could reduce or suppress the impact of the activity of the perceptual neurons that encoded the relative position of the red and/or of the green object. A similar technique has been used by Kruschke (1992) to study the role of selective attention in category learning. The two additional output units received connections directly from the target object units. The internal neurons received connections from the visual perceptual neurons and projected connections to the motor neurons. This neural architecture includes the neural circuits that could enable the development of a selective attention solution in which the irrelevant perceptual information could be filtered out at the perceptual level by means of the top-down regulatory connections from the two additional output units to the visual neurons. Notice also how the lack of connections from the target object units to the internal neurons implied that this second architecture could not support the development of action selection strategies.

The architecture used in the third experimental condition, shown in the right of Fig. 2, was identical to that described above but also included connections from the target object neurons to the internal neurons. This last architecture, therefore, could support the development of either selective attention or action selection strategies as well as mixed solutions.

The proposed neural architectures do not aim to model the anatomy of the brain but only the minimal set of characteristics that need to be taken into account to study the role of selective attention and/or action selection.

2.3. The adaptive process

The connection weights and biases that determined how the robot reacted to stimuli were initially set randomly and then varied during the adaptive/learning process. Thus the same robot was instantiated with different neural network controllers, initialized with randomly extracted parameters for each replication, and then trained. We refer to these different instantiations as individuals.

During the adaptation process the robot was allowed to interact with the environment for a series of trials, each lasting 15s, during which it was exposed to two blocks of four trials, of which two involving a single red or green object and the other two involving both a red and a green object (see Fig. 1). The single object trials were introduced to ensure that, as in natural ecological settings, the robot could face both simple and more complex conditions. Objects consisted of spheres with a diameter of 3.5cm that were moved along a circular trajectory clockwise or counter-clockwise with an angular speed of 24°/s. The circular trajectory had a radius of 10cm, and was centred at a distance of 32.5±2.5cm from the robot with a lateral offset of ±2.5cm with respect to the centre of the robot. At the beginning of each trial, the position of the head was reset as depicted in Fig.1, the centre of the spherical trajectory was randomly chosen among the four possible points indicated above, and the initial position of the object was set on the uppermost point of the corresponding circular trajectory in half of the trials and on the lowermost point in the other half. When two objects were presented simultaneously, they were placed one of the uppermost point and the other on the lowermost point of the corresponding circular trajectory and then moved either clockwise or counter-clockwise.

The experiments were realized by using two alternative training methods: a trial and error and a learning by demonstration method. In the trial and error condition, the free parameters were varied randomly and variations were retained or discarded on the basis of their overall effect on the ability of the individual to bring and maintain the target object over the centre of the view field. This was realized by using an evolutionary method (Nolfi & Floreano, 2000). The reason behind the choice of this algorithm is that it is one of the simplest yet most effective ways to train an embodied neural network through a trial and error process based on a distal reward (Schlesinger, 2004). The initial population consisted of 20 randomly generated genotypes encoding the connection weights and biases of 20 corresponding neural controllers (each parameter was encoded by eight bits and normalized in the interval [-5.0, 5.0]). During each generation, each individual was allowed to produce an offspring, that was a varied copy of itself generated by mutating each bit with a probability of 2%. Offspring were then evaluated and used to replace the worst genotype of the population or discarded on the basis of whether they achieved or not a better fitness. The fitness for a trial was computed on the basis to the following equation:

$$\mathrm{fit} = \mathrm{e}^{(-\mathrm{dist}*10)} \tag{1}$$

where *e* is the exponential function, *dist* is the mean Cartesian distance between the centre of the view field and the centre of the target object in cm. The evolutionary process has been conducted for 1500 generations. The experiment has been replicated 10 times with different seeds for the random number generator.

In the learning by demonstration experimental condition the robot was trained through back-propagation on the basis of a series of demonstration generated through an automated procedure. The connection weights and biases were initialized randomly in the interval [-0.5, 0.5] and the robot was exposed to 1000 blocks of four trials (of which two involving either a single red or green object and two involving both a red and a green object). The teaching input (i.e. the desired state that the two motor neurons should assume for each time step) was computed online by a software routine that calculated the direction and the speed that the two actuators of the head should assume to ensure that the visual perception of the target object moved and/or remained toward the centre of the robot's view field. The learning rate has been set to 0.2. The experiment has been replicated 10 times with different seeds for the random number generator.

In the case of the experimental conditions in which the robot neural network also included the attentional system, the sum of the delta errors back-propagated on the first and the second block of four perceptual visual neurons were used as delta errors for the first and the second attentional output neurons. In other words, the teaching signal for the attentional output neurons that regulated the relative impact of the red and green perceptual neurons at time $t+1$ was calculated indirectly on the basis of whether a larger or lower activation of the red and green perceptual neurons would have reduced the difference between the desired and the actual state of the motor neurons.

Overall the combination of three experimental conditions varying with respect to the architecture of the neural network and two experimental conditions varying with respect to the training method led to a total of six experimental conditions. For each condition, 10 replications were ran.

3. Results

We describe the results obtained by using the trial-and-error and the learning by demonstration method respectively in section 3.1 and 3.2. In each section, the results obtained by using the three different neural architectures described above are reported.

To compare the performance obtained in different experimental conditions we calculated an error measure that consisted in the mean distance between the centre of the visual field and the target object scaled by the size of the visual field. These data were collected by post-evaluating the best individual (i.e., that with the lowest average error across all trials) of the last generation (in the case of the trial-and-error experiments) or the individuals at the end of the training process with their connection weights fixed (in the case of learning by demonstration experiment). In all cases, the error measures were calculated by evaluating the individuals for 100 blocks of 4 trials.

Moreover, for each experimental condition, we analysed the extent to which trained individuals were able to ignore irrelevant visual information.

3.1. Trial-and-error learning condition

As show by the results reported in Fig. 3, the experiments conducted with the Selective Attention architecture led to better results (i.e. lower error) than the experiments performed with the Action Selection architecture. The Selective Attention architecture also led to better results with respect to the Mixed architecture, although the difference was less marked with respect to the Action Selection architecture.

An ANOVA analysis performed by using the neural architecture as independent factor showed a significant main effect of this factor on the average errors, $F(2, 27)=69.77$, $p<0.01$. Post-hoc analyses conducted with t-tests revealed that the individuals with the Action Selection architecture obtained higher error (0.077) than both individuals with the Selective Attention architecture (0.053), $p<0.001$, and individuals with the Mixed architecture (0.063), $p<0.001$. Individuals with the Mixed architecture obtained higher error than individuals with the Selective Attention architecture, $p<0.01$.

The analysis of the error obtained in different type of trials (see Fig. 3) indicated that the individuals trained with the Selective Attention and Mixed architecture displayed close to optimal performance in all cases while the individuals trained with the Action Selection architecture performed well only in three out of four cases. More specifically, they displayed close to optimal performance during the trials in which they experienced a single red or green object and during the trials in which they experienced both objects and should visually follow the object of one colour. The best individual with the Action Selection architecture, for example, displayed a relatively low error during the trials in which it was asked to foveate the red object (Red only and Red conditions) and in which it was asked to foveate the green object and the red object was missing (Green only condition) but displayed a higher error when it should foveate the green object when also the red object was present (Green condition) (Figure 3, left panel). A qualitatively similar outcome was observed in the other replications, although the colour of the object that could be correctly foveated during double object trials varied in different replications (result not shown for reason of space).

On the contrary, by analysing the post-evaluated errors obtained in the other experimental conditions, in which the evolving robot was provided with the additional output units that modulated the impact of visual information, we can see how evolved individuals reached similar close to optimal performance in each type of trial (see Fig.3, middle and right panel).

Overall this indicated that the addition of output units that can modulate the impact of different type of perceptual information facilitated the development of good solutions and enabled evolving robot to reach close to optimal performance. The inspection of the state of the two additional output units indicated that the robot used them to completely suppress irrelevant information during double object trials (i.e. the output units tended to assume a [1.0, 0.0] and

[0.0, 1.0] states during double object trials in which the robot was asked to look respectively at the red or at the green object).

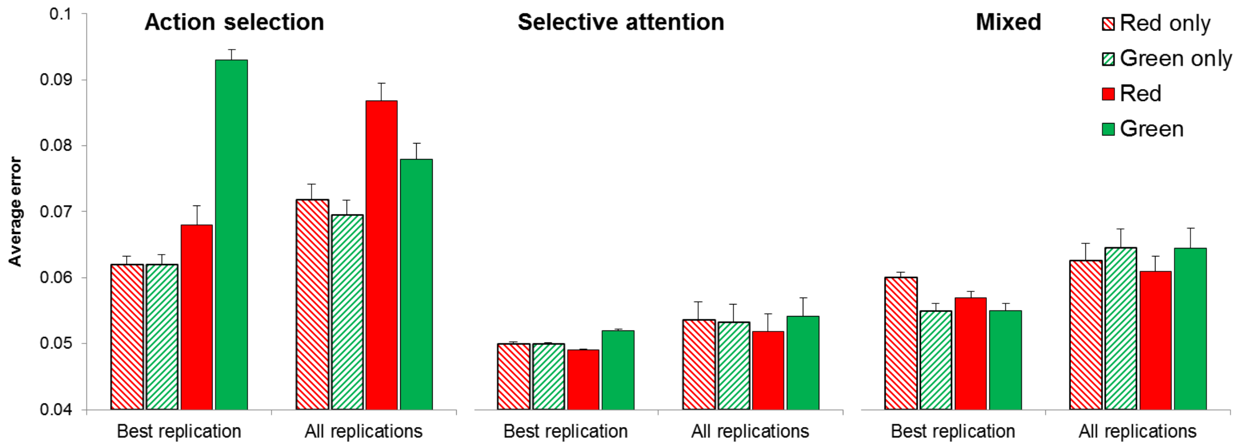

Figure 3. Average errors obtained by post-evaluating the robots trained in the trial and error condition. For each neural architecture (Action Selection, Selective Attention and Mixed) the error obtained in the case of the best replication and the average error obtained over the ten performed replications are shown. Each histogram indicates the average error and the standard error obtained in trials in which the robot was asked to pay attention to red objects and was exposed to a single red or to both a red and green object (Red only and Red condition) and in which the robot was asked to pay attention to green objects and was exposed to a single green or to both a green and red object (Green only and Green condition).

The analysis of the course of the adaptive process provided some indication that could explain why the individuals provided with the Action Selection architecture trained with the trial and error method, unlike the individual provided with the other architectures, remained stacked in local minima in which only three out of four types of trials were mastered correctly. Figure 4 shows how the error varied throughout generations, in the case of the best replication, in the three experimental conditions. As can be seen, the robot provided with the Selective Attention or Mixed architecture concurrently improved their ability to master all type of trials and reached close to optimal performance already after about 300 generations. The robot provided with the Action Selection architecture, instead, first developed an ability to master the single object trials (after about 300 generations) and then an ability to master one of the two double objects trials (after about 1300 generations).

Overall this indicates that in the Action Selection condition the adaptive process first converged on a suboptimal solution that consists in moving the head toward either the red or the green objects, irrespectively of the state of the target input units (that enables the robots to master the single object trials) and then on a second sub-optimal solution that consisted in responding more strongly to the position of the object of one colour than to the object of the other colour (that enable the robot to master relatively well also one of the two double object trials). On the other hand, the convergence toward such a strategy brought the adaptive process to a local minimum in which small variations could not lead to further improvements and in which progress could only be achieved through a significant restructuring of the strategy.

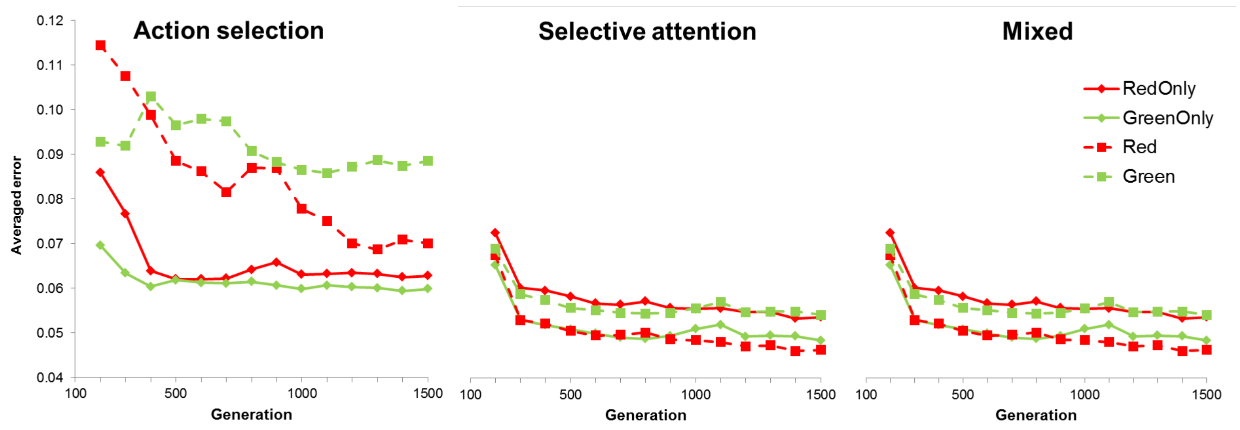

Figure 4. Average errors throughout generations in the case of the best replication of the experiment performed with the trial and error training method. Each curve displays the average error reported during the four type of trials. To avoid the need to use a much larger error scale, data are shown from generation 100 on only. Data obtained by post-evaluating the individuals for 400 trials.

This explanation was further confirmed by the analysis of the impact of the irrelevant perceptual information on the internal and motor neurons, shown in Figure 5. Such measure was calculated by storing the state assumed in each step by the neurons while the robot was exposed to a series of double and single object trials with its head blocked, and by computing the absolute difference between the state of the internal and motor neurons in the former and latter case with the target object in the same position. Differences are then averaged over all steps.

As shown in Figure 5, while in the Selective Attention and Mixed condition the impact of the irrelevant visual information was quickly reduced during the first generations, in the case of the Action Selection condition it remained rather high up the end of the evolutionary process and, overall, did not decrease throughout generations. In other words, the variations that enabled the individuals trained in the Action Selection condition to improve their ability to master first single object trials and then one of the two object trials did not favour the development of an ability to differentiate between relevant and irrelevant perceptual information, neither in part. This, however, then prevented further progresses, since the ability to conditionally filter out the irrelevant perceptual information was a pre-requisite for fully mastering the task. Moreover, the rather similar trend observed in the Selective Attention and Mixed conditions (Figure 5) suggested that also in the latter condition the irrelevant visual information was primarily filtered out through the top-down regulatory connections. Qualitatively similar results were obtained by analysing all replications of the experiments (results not shown for reasons of space).

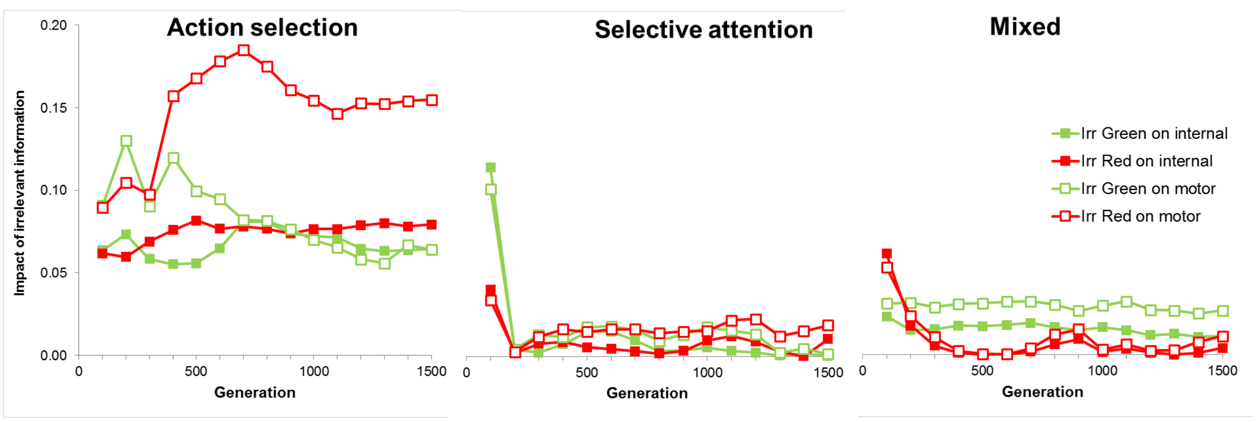

Figure 5. Impact of irrelevant perceptual information (Irr Red = impact of red position when the target is green, IrrGreen = impact of green position when the target is red) on internal and motor neurons activity throughout generations. Data computed for the best individual of each generation of the best replication, for each experimental condition. Data averaged over 100 successive generations for each data bin.

To verify that the differences in performance could not be explained by the presence of additional neurons and/or connections in the Selective Attention and Mixed conditions, we carried out two additional control experiments in which we used: (i) an Action Selection architecture with 8 internal neurons (instead of 6), and (ii) a Selective Attention architecture in which the modulatory connections were directed toward the internal neurons rather than toward the perceptual visual neurons. All other parameters were kept the same.

The fact that the results obtained in these two control experiments (see Figure 6) are fairly similar to that obtained in the experiment in which the robot were provided with the standard Action Selection architecture (Figure 3, left panel) indicated that the raw amount of neural resources as well as the possibility to modulate the neural activation were not sufficient by themselves to ensure the development of close to optimal solutions in all conditions. Comparisons conducted with t-tests indicated that the error observed with the standard Action Selection architecture (0.077) was not significantly different from both the error obtained in the control experiment with two additional internal neurons (0.074), *p=0.16*, and from the error obtained in the control experiment in which the attentional output units modulated the state of the internal neurons, (0.078), *p=0.62.* Moreover, the average errors displayed by the robot trained in the two control conditions were significantly different from the error displayed by the robot trained with the standard Selective Attention architecture, *p<0.001*, and with the Mixed architecture, *p<0.001.*

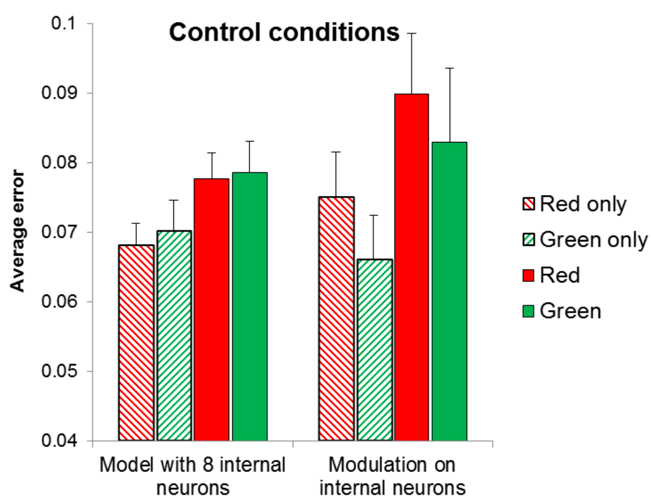

Figure 6. Average errors obtained by post-evaluating the robots trained in the two control conditions (see text).

3.3 Results obtained in the learning by demonstration condition

The analysis of the performance in the learning by demonstration condition, in which the robot was asked to produce, for each experienced input state, the action that minimized the relative distance between the barycentre of the target object and centre of the robot visual field, indicated that in this case the individuals developed a close to optimal ability in all cases, independently of the type of neural architecture (see Fig. 7). An ANOVA analysis performed by using the neural architecture as independent factor showed no significant main effect of the architecture on the average errors, $p=.49$. Indeed, t-test comparisons revealed no differences between average error obtained by individuals with the Action Selection (0.063), Selective Attention (0.059), and Mixed (0.059) architectures (i.e. $p>0.05$ in all cases).

Thus, individuals trained through the learning by demonstration method achieved close to optimal performance irrespectively from the neural architecture, with almost identical performances obtained in the case of the Selective Attention and Mixed architecture, as well as in the case of Action Selection architecture. Indeed, the visual inspection of the robot behaviour at the end of the training process during a post-evaluation test in which the connection weights were frozen showed that all replications managed to solve the problem rather well in all trials.

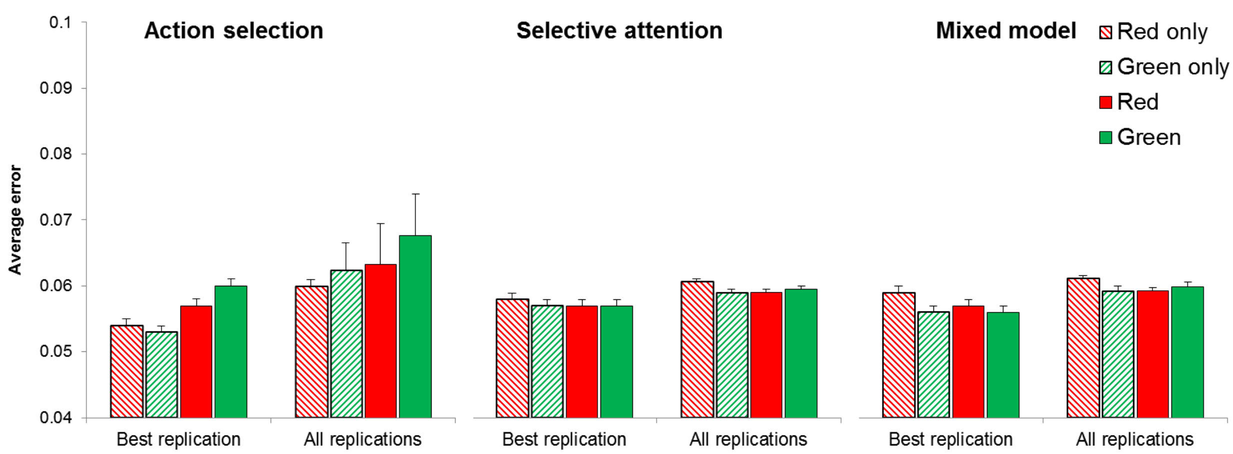

Figure 7. Average errors obtained by post-evaluating the individuals trained in the learning by demonstration condition. For each neural architecture (Action Selection, Selective Attention and Mixed) the error obtained in the case of the best replication and the average error obtained over the ten performed replications are shown. Each histogram indicates the average error and the standard error obtained in trials in which the robot was asked to pay attention to red objects and was exposed to a single red or to both a red and green object (Red only and Red condition) and

in which the robot was asked to pay attention to green objects and was exposed to a single green or to both a green and red object (Green only and Green condition).

We then conducted also for the learning by demonstration experiments the analysis of the impact of the irrelevant perceptual information on the internal and motor neurons. As shown in Figure 8, the individuals with both the Selective Attention and Mixed architectures learned how to reduce the impact of the irrelevant perceptual information already after the first hundreds of trials. This seems to indicate that, as in the trial and error training condition, the irrelevant perceptual information was primarily filtered out through the top-down regulatory connections. In the Action Selection condition, instead, the acquisition of the capability to eliminate the impact of irrelevant information on the motor states took considerably more training time. Notice also how the reduction of the impact of the irrelevant information was carried out by both the two processing layers, as demonstrated by the fact that at the end of the training process the impact of irrelevant information on the internal neurons was still considerable while the impact on the motor neurons was almost null. Qualitatively similar results were obtained by analysing all replications of the experiments (results not shown for reasons of space).

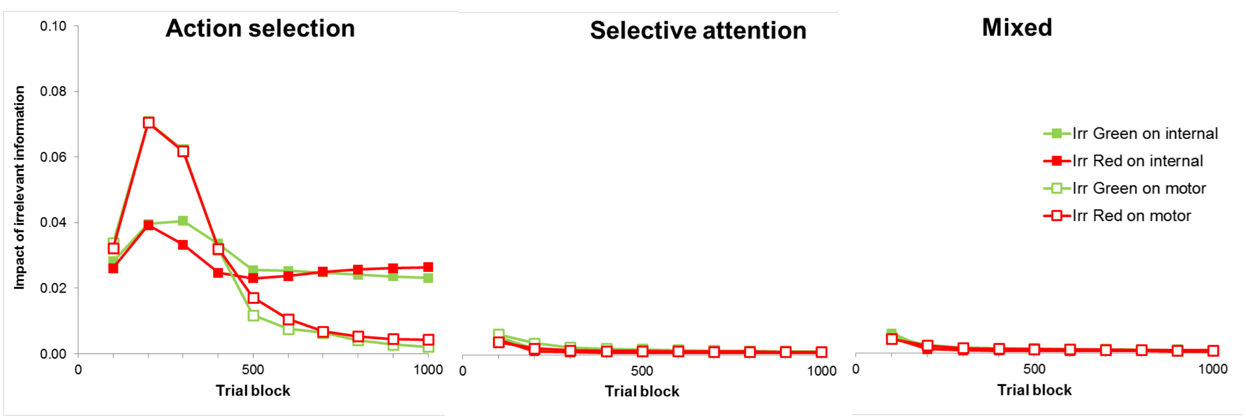

Figure 8. Impact of irrelevant perceptual information (Irr Red = impact of red position when the target is green, IrrGreen = impact of green position when the target is red) on internal and motor neurons activity throughout trials. Data computed over each block of four trials for the individual of the best replication, for each experimental condition. Data averaged over 100 successive trial blocks.

4. Discussion

As we mentioned in the introduction, the problem of selecting actions in contexts affording multiple conflicting behavioural responses can be solved either through a selective attention or through an action selection process. The former strategy consists in using top-down regulatory connections to filter out the perceptual information affording alternative actions. The latter strategy, instead, consists in processing in parallel the perceptual information affording multiple conflicting actions by later selecting an action through a biased competitive process that operates at the pre-motor level.

To verify the implications of the utilization of the former and/or of the latter strategy from a functional and developmental perspective we carried out a series of experiments in which a neuro-robot was trained for the ability to visually follow either a red or a green moving object in contexts in which the robot was rewarded for following only one of them.

Overall the obtained results indicate how a selective attention strategy, realized through the exploitation of top-down regulatory connections from pre-motor areas to perceptual areas, represents the most simple and straightforward way to select actions. Indeed, the obtained results indicate how the availability of top-down regulatory connections is always exploited to supress the perceptual information affording conflicting actions in robot learning through a trial and error process based on a distal reward as well as in robot learning through a learning by demonstration process.

The tendency to exploit a selective attention strategy can be explained by considering that it enables the synthesis of solutions that are more parsimonious from the control point of view and that are more compatible with a developmental account. In these solutions context independent situations, corresponding to single object trials, can be handled through the development of direct associations between stimuli and afforded actions, and context dependent situations, corresponding to double objects trials, can be handled by the combination of the capacity above with the capacity to filter out irrelevant perceptual information through top-down regulatory connections.

On the contrary, from the point of view of an action selection strategy, the development of direct association between stimuli and actions afforded by them does not constitute a stepping stone toward the development of better strategies and might rather leads to an organization that is incompatible with more effective solutions and that therefore needs to be completely rearranged to also master contexts affording conflicting actions. More specifically, the tendency to associate stimuli to afforded actions in context independent situations does not constitute an elementary ability that can be reused, in combination with other acquired skills, to also handle context dependent situations but rather a step toward sub-optimal local minima that can then prevent the possibility to fully master context dependent situations.

Our results also indicate that the characteristics of the training process co-determine whether the individuals converge toward a selective attention and/or an action selection solution and to optimal or sub-optimal solutions. Indeed, for the reason described above, when the architecture provide top-down regulatory connections, the adaptive process tends to converge toward solutions that exploit primarily a selective attention mechanism to filter out irrelevant perceptual information. Moreover, while agents adapting through a learning by demonstration method display an ability to acquire either optimal action selection or optimal selective attention strategies, agents adapting through a trial and error method display an ability to acquire optimal selective attention strategies but only sub-optimal action selection strategies. This last difference can be explained by considering that the training feedback indicates more explicitly that the irrelevant perceptual information should be ignored in the case of the learning by demonstration method than in the case of the trial and error method.

Overall our results suggest that selective attention should play a wide role in natural organisms especially in the case of capacities that are acquired through trial and error developmental processes based on distal rewards. Although, as we pointed out in the introduction, action selection remains the only feasible strategies in contexts in which the perception of the stimuli affording multiple conflicting action precedes temporarily the perception of the information that can be used to determine the action to be chosen (Cisek & Kalaska, 2005; Cisek, 2007).

Finally our results suggest that a primary role played by the modulatory effects of motor areas on perceptual areas is that to resolve the conflicts between multiple actions afforded by the current context. This functional role might have been overlooked to date. Indeed, the role of these top-down influences is usually associated to the capability to anticipate the forthcoming sensory states on the basis of the current planned action (see Wolpert & Miall, 1996).

References

Bonfiglioli, C., & Castiello, U. (1998). Dissociation of covert and overt spatial attention during prehension movements: selective interference effects. *Perception & Psychophysics*, *60*(8), 1426–40.

Castiello, U. (1996). Grasping a fruit: selection for action. *Journal of Experimental Psychology. Human Perception and Performance*, *22*(3), 582–603.

Castiello, U. (1999). Mechanisms of selection for the control of hand action. *Trends in Cognitive Sciences*, *3*(7), 264–271.

Chieffi, S., Gentilucci, M., Allport, A., Sasso, E., & Rizzolatti, G. (1993). Study of selective reaching and grasping in a patient with unilateral parietal lesion. Dissociated effects of residual spatial neglect. *Brain: A Journal of Neurology*, *116*(5), 1119–37.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1485), 1585–99. doi:10.1098/rstb.2007.2054

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron*, *45*(5), 801–14. doi:10.1016/j.neuron.2005.01.027

Duncan-Johnson, C. C., & Kopell, B. S. (1981). The Stroop effect: brain potentials localize the source of interference. *Science (New York, N.Y.)*, *214*(4523), 938–940. doi:10.1126/science.7302571

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. *Book* (Vol. 40, p. 332). Boston: Houghton Mifflin. doi:10.2307/989638

Goldberg, M. E., & Segraves, M. A. (1987). Visuospatial and motor attention in the monkey. *Neuropsychologia*, *25*(1A), 107–118. doi:10.1016/0028-3932(87)90047-9

Howard, L. A., & Tipper, S. P. (1997). Hand deviations away from visual cues: indirect evidence for inhibition. *Experimental Brain Research.*, *113*(1), 144–52.

Humphreys, G. W., Yoon, E. Y., Kumar, S., Lestou, V., Kitadono, K., Roberts, K. L., & Riddoch, M. J. (2010). The interaction of attention and action: from seeing action to acting on perception. *British Journal of Psychology (London, England : 1953)*, *101*(Pt 2), 185–206. doi:10.1348/000712609X458927

Jackson, S. R., Jackson, G. M., & Rosicky, J. (1995). Are non-relevant objects represented in working memory? The effect of non-target objects on reach and grasp kinematics. *Experimental Brain Research.*, *102*(3), 519–30.

Kruschke, J. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44.

Massera, G., Ferrauto, T., Gigliotta, O., & Nolfi, S. (2013). FARSA: An Open Software Tool for Embodied Cognitive Science. In *Advances in Artificial Life, ECAL 2013* (Vol. 12, pp. 538–545). MIT Press. doi:10.7551/978-0-262-31709-2-ch078

Miller, J., & Hackley, S. A. (1992). Electrophysiological evidence for temporal overlap among contingent mental processes. *Journal of Experimental Psychology. General*, *121*(2), 195–209. doi:10.1037/0096-3445.121.2.195

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: sustained inattentional blindness and the capture of awareness. *Psychological Review*, *112*(1), 217–42. doi:10.1037/0033-295X.112.1.217

Munneke, J., Heslenfeld, D. J., Usrey, W. M., Theeuwes, J., & Mangun, G. R. (2011). Preparatory effects of distractor suppression: evidence from visual cortex. *PloS One*, *6*(12), e27700. doi:10.1371/journal.pone.0027700

Neumann, O. (1987). Beyond capacity: A functional view of attention. In A. F. Heuer & H. Sanders (Eds.), *Perspectives on perception and action.* (pp. 361–394). Lawrence Erlbaum Associates, Inc.

Nolfi, S., & Floreano, D. (Eds.). (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines.* Cambridge, MA: MIT Press/Bradford Books.

Pastor-Bernier, A., & Cisek, P. (2011). Neural correlates of biased competition in premotor cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(19), 7083–8. doi:10.1523/JNEUROSCI.5681-10.2011

Payne, H. E., & Allen, H. A. (2011). Active ignoring in early visual cortex. *Journal of Cognitive Neuroscience*, *23*(8), 2046–58. doi:10.1162/jocn.2010.21562

Petrosino, G., Parisi, D., & Nolfi, S. (2013). Selective attention enables action selection: evidence from evolutionary robotics experiments. *Adaptive Behavior*, *21*(5), 356–370. doi:10.1177/1059712313487389

Sandini, G., Metta, G., & Vernon, D. (2004). RobotCub: an open framework for research in embodied cognition. *4th IEEE/RAS International Conference on Humanoid Robots, 2004.*, *1*. doi:10.1109/ICHR.2004.1442111

Schlesinger, M. (2004). Evolving agents as a metaphor for the developing child. *Developmental Science*, *7*(2), 158–64.

Seidl, K. N., Peelen, M. V, & Kastner, S. (2012). Neural evidence for distracter suppression during visual search in real-world scenes. *The Journal of Neuroscience : The Official*

Journal of the Society for Neuroscience, *32*(34), 11812–9. doi:10.1523/JNEUROSCI.1693-12.2012

Seth, A. K., Prescott, T. J., & Bryson, J. (2011). *Modelling Natural Action Selection*. (A. K. Seth, T. J. Prescott, & J. Bryson, Eds.) (p. 560). Cambridge, MA: Cambridge University Press.

Tipper, S. P., Howard, L. A., & Houghton, G. (1998). Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1373), 1385–93. doi:10.1098/rstb.1998.0292

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective Reaching to Grasp: Evidence for Distractor Interference Effects. *Visual Cognition*, *4*(1), 1–38. doi:10.1080/713756749

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, *3*, 86–94.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 830–46.

Wolpert, D. M., & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Networks : The Official Journal of the International Neural Network Society*, *9*(8), 1265–1279.