# Exploration and learning in capuchin monkeys (*Sapajus spp.*): the role of action-outcome contingencies

**Eugenia Polizzi di Sorrentino**[1]**, Gloria Sabbatini**[1]**, Valentina Truppa**[1]**, Anna Bordonali**[2]**, Fabrizio Taffoni**[3]**, Domenico Formica**[3]**, Gianluca Baldassarre**[4]**, Marco Mirolli**[4]**, Eugenio Guglielmelli**[3]**, Elisabetta Visalberghi**[1]

[1]Unit of Cognitive Primatology and Primate Center, ISTC-CNR, Rome
[2]Department of Life Sciences and Systems Biology, University of Torino, Torino
[3]Laboratory of Biomedical Robotics and Biomicrosystems, Center for Integrated Research, Università Campus Bio-Medico, Rome
[4]Laboratory of Computational Embodied Neuroscience, ISTC-CNR, Rome

Corresponding author:
Eugenia Polizzi di Sorrentino: Istituto di Scienze e Tecnologie della Cognizione, CNR, via Aldrovandi 16b, 00197 Roma, Italy. email: eugenia.polizzi@istc.cnr.it

## Abstract

Animals have a strong propensity to explore the environment. Spontaneous exploration has a great biological significance since it allows animals to discover and learn the relation between specific behaviours and their consequences. The role of the contingency between action and outcome for learning has been mainly investigated in instrumental learning settings and much less in free exploration contexts. We tested 16 capuchin monkeys (*Sapajus* spp.) with a mechatronic platform that allowed complex modules to be manipulated and to produce different outcomes. Experimental subjects could manipulate the modules and discover the contingencies between their own specific actions and the outcomes produced (i.e., the opening and lighting of a box). By contrast, Control subjects could operate on the modules but the outcomes experienced were those performed by their paired Experimental subjects ("yoked-control" paradigm). In the Exploration Phase, in which no food reward was present, Experimental subjects spent more time on the board and manipulated the modules more than Yoked subjects. Experimental subjects outperformed Yoked subjects in the following Test phase, where success required recalling the effective action so to open the box, now baited with food. These findings demonstrate that the opportunity to experience action-outcome contingencies in the absence of extrinsic rewards promotes capuchins' exploration and facilitates learning processes. Thus, this intrinsically motivated learning represents a powerful mechanism allowing the acquisition of skills and cognitive competence that the individual can later exploit for adaptive purposes.

*Keywords*: learning, intrinsic motivation, contingency, mechatronic platform

## Introduction

Animals are highly motivated to explore, and understanding the role of such motivation in promoting learning has recently attracted the interest of a broad range of disciplines from animal behaviour (Antunes and Biala 2012; Byrne 2013), to neuroscience (Li et al. 2003; Redgrave and Gurney 2006) and machine learning (Baldassarre and Mirolli 2013b; Barto et al. 2004; Oudeyer et al. 2007; Schembri et al. 2007).

Psychologists introduced the concept of intrinsic motivations, i.e., drives for which actions are performed "for their own sake", to explain what motivates animals to explore, play, or engage in other behaviours in the absence of external reinforcement (Hughes 1997). Macaques

exploring a mechanical puzzle, acquire knowledge about its functioning and eventually solve this puzzle in the absence of extrinsic (e.g., food) rewards (Harlow et al. 1950). Intrinsically motivated learning processes allow the acquisition of competence, which helps individuals to achieve information about the environment features (White 1959). Although competence may not be driven by specific problems, the learned skills can act as "building blocks" out of which animals can devise solutions as new problems arise. Thus, intrinsic motivations represent a set of important mechanisms underlying the acquisition of skills and cognitive competences later exploitable for adaptive purposes (Deci 1975; Ryan and Deci 2000a,b; Baldassarre and Mirolli 2013a, Mirolli and Baldassarre 2013).

When animals have the opportunity to interact with the environment by manipulating objects or combining objects with surfaces, they may discover and learn the contingency between one action and its outcome (for example, discovering that banging an object produces noise). The role of contingency has been mainly studied through instrumental learning paradigms in which behaviour was rewarded (e.g., Rescorla, 1968, Pearce 2008). In contrast, the role of action-outcome contingencies in unrewarded contexts has been little investigated. For example, visual and auditory stimuli are effective as reinforcing agents in operant conditioning situations showing that response-contingent stimulation promotes operant responses in mice (Kish 1955), rats (Winefield and Glow 1980) chickens (Meyer and Collins 1971) and primates (Blatter and Schultz 2006; Butler 1954, 1957). Rats learn to press a lever to cause the onset of a light in the absence of primary rewards (Reed et al. 1996) suggesting that the opportunity to discover (and control) action-outcome contingencies may be intrinsically motivating. More recently, Buchanan-Smith and Badhini (2011) adopted a yoked-control protocol to rule out the effect of action-outcome contingencies (namely, the opportunity to switch on and off light and heat) from the effect of the outcomes themselves (namely, the change in light and heat). Their study on captive marmosets (*Callithrix jacchus*) demonstrated that control over supplementary lighting and heat was more rewarding than the effects themselves. However, it did not clarify whether the knowledge acquired through these experiences could be subsequently recalled and used in a goal-directed fashion, leaving its adaptive value untested.

Our experiment aims to understand the role of action-outcome contingency in promoting intrinsically motivated learning processes. For this purpose, tufted capuchin monkeys (genus *Sapajus*, Lynch Alfaro et al. 2012a,b) are particularly suited given their explorative and manipulative attitudes (Fragaszy et al. 2004). Capuchins exhibit a great variety of behaviours to explore and act on the environment especially while foraging (Fragaszy et al 2004; Perry and Manson 2008; Terborgh 1983). Both wild and captive capuchins spontaneously perform object-object and object-surface combinations (Byrne and Suomi 1996; Fragaszy and Adams-Curtis 1991; Fragaszy and Boinski 1995; Panger 1998; Visalberghi 1988), tool use (Ottoni and Mannu 2001; Visalberghi and Fragaszy 2013), and gather information about tool affordances in the absence of extrinsic rewards (Manrique et al. 2011).

The rationale of our study is to verify whether discovering and repeatedly experiencing congruent action-outcome contingencies through spontaneous exploration improves problem solving ability in a subsequent task, which requires recalling the information acquired during the previous exploration. Our experiment involves two phases. During the first exploration phase (Phase 1) subjects could explore the properties of two modules contained in a mechatronic board (see Taffoni et al. 2012 for further details) and possibly discover the relation between their actions and the outcomes produced (for example, that the rotation of a given module opens a box which, in this phase, did not contain a reward). During the following test phase (Phase 2), the box was rewarded and subjects had to recall the action that in Phase 1 produced the opening of the box to retrieve the reward. While for Experimental subjects the outcomes experienced in the first phase were contingent with their own actions, for Control subjects the outcomes mirrored those experienced by the Experimental subjects, instead of being produced by them. By means of this yoked-control paradigm, we assessed whether

experiencing congruent action-outcome contingencies allows learning whereas the mere experience of the board associated with incongruent outcomes does not.

Given the evidence reviewed thus far, we predicted that congruent action-outcome contingencies discovered in Phase 1 would promote spontaneous exploration in the Experimental group, and that their absence would diminish exploration in Yoked subjects (Prediction 1). We also predicted that the opportunity to open the box would lead Experimental subjects to perform the effective actions (i.e., those associated with the opening of the box) more frequently than ineffective ones (Prediction 2). Furthermore, we expected that an increased number of box openings in Phase 1 should lead to a shorter latency to solution in Phase 2 (Prediction 3). Assuming the above predictions would have been satisfied, we expected Experimental subjects to outperform Yoked subjects in Phase 2 (Prediction 4). Finally, as Yoked subjects could experience congruent action-outcome contingencies in Phase 2, we expected them to improve their performance during this Phase (Prediction 5).

**Methods**

*Subjects*

The subjects were 16 socially-housed adult tufted capuchin monkeys (8 females and 8 males) hosted at the Unit of Cognitive Primatology and Primate Centre, ISTC-CNR of Rome. The groups were housed in enclosures consisting of an outdoor area (group A = 106 m³, group B = 128 m³, group C = 374 m³, group D = 130 m³) and two indoor cages (overall of about 25 m³). Capuchins were tested individually in the indoor area, to which they have access through a sliding door from the adjacent outdoor enclosure. Each subject was separated from the group solely for the purpose of testing, just before each testing session. Subjects belonging to the Experimental and Yoked groups had a comparable experimental history with perceptual and cognitive tasks. Monkey chow (Altromin-A pellets, Rieper standard diet for primates), fresh fruits and vegetables were given every afternoon after testing. Water was freely available at all times.

*Apparatus*

The mechatronic board consisted of a vertical element (80x20x40 cm) attached to a semi-transparent base (80x60x20 cm) equipped with two identical modules (called *Circular Taps*) placed 50 cm from one another and 11 cm from the vertical element (Fig 1). Each *Circular tap* consisted of a 6 cm vertical metal bar capped by a 6 cm horizontal metal disc. Each *Circular tap* could be lifted 4 cm, rotated clockwise and rotated counter-clockwise (Fig 1). A control software running on a remote laptop allowed experimenters to programme the association between actions (e.g., lifting the bar, rotation of the tap of the left or right module) and specific outcomes, such as the opening of an opaque rewarding box placed at the centre of the vertical element. Each one of the four possible actions produced a different sound (four different kinds of bell sounds). The opening of the box was associated with the activation of the lights below the box and inside it. The experimenter could fill the box with a reward through an opening positioned at the back of the vertical element.Finally, a wide-angle camera fixed on the top of the board allowed video-recording of the workspace during the experiment.



**Fig 1** The mechatronic board. The modules are on the *right* and *left* sides of the platform, whereas the *black square* in the *vertical central panel* is the opaque rewarding *box*. The *circles* below the *box* could light while the ones close to the modules could produce sounds. The *grey arrow* shows the rotation of the tap and the *white arrow* the lifting of the *bar*.

*Procedure*

The experiment involved an exploration phase (Phase 1) and a testing phase (Phase 2). Phase 1 consisted of a 12-minute session in which each subject could explore the board and manipulate the modules. The duration of the session was chosen on the basis of a pilot experiment not involving the subjects of the experiment so to minimize the possible decrease of interest toward the board due to habituation, and still allow enough time for exploration. Each Experimental subject had to reach the criterion of opening the box at least 10 times, before facing Phase 2.

Each subject of the Experimental group was paired with one subject of the Yoked-control (hereafter, Yoked) group, and individuals within each pair were matched, as much as possible, in terms of sex, age, previous experience, and level of exploration. Exploration was assessed during a 5-min test during which subjects were individually presented with an apparatus equipped with a metal handle that could be rotated. The level of exploration did not differ among groups, neither in terms of number of actions directed toward the handle (unpaired *t*-test: $t_{(14)}=-1.51$, $P=0.153$) nor in time spent in contact with the apparatus (unpaired *t*-test: $t_{(14)}=-0.57$, $P=0.578$). During Phase 1, the Experimental subjects could manipulate the modules (and experience the relative action-outcome associations) and open the box by performing a specific action. This action consisted of rotating the tap of one of the two modules for at least 45 degrees (either clockwise or counter-clockwise). When the correct rotation was performed (box-opening rotation, hereafter BO rotation), the box opened along with a specific sound and a light stimulus appeared below and inside the box. The other actions (rotating the tap of the other module and performing lifting actions on both modules) did not open the box and were associated only with sounds (non box-opening actions, hereafter NBO actions). The module associated with the opening of the box was counter-balanced among subjects. Yoked subjects could operate on the modules but no outcome was directly produced. Instead, the outcomes they experienced were identical to those performed by their paired Experimental subjects (see video-clip, Online Resource 1). This was done to provide Yoked subjects with the same number of outcomes of their paired Experimental subjects, while preventing the Yoked subjects from repeatedly experiencing congruent associations between their own actions and outcomes. For both Experimental and Yoked subjects, Phase 2 consisted of 10 consecutive trials, each lasting maximum 2 minutes. For each of the 10 trials, the experimenter baited the box with one reward (one unshelled peanut kernel) while monitoring whether the subject was paying attention to the baiting. In each trial, capuchins could manipulate the modules and if the correct action (i.e., the BO rotation) was performed, the box opened so that they could retrieve the reward, and a new trial started over. If no correct action was performed within the 2 minutes, the subject was separated in the adjacent enclosure, the reward extracted from the box, and a next trial started over.

*Data collection*

All sessions were video-recorded and data were extracted *a posteriori* from the videos. E.P.D.S. scored all trials and a second rater (GS) scored a random selection of 20% of the trials and the percentage of agreement was 81.2%.

During Phase 1 the experimenters recorded the subjects' latency to approach the board, the time spent in contact with the board, and the number of manipulative actions (rotation, lift) performed on each module. During Phase 2, the experimenters recorded the time at which subjects retrieved the reward, the number and type of manipulative actions performed to get the reward, and the total number of reward items obtained.

*Analysis*

We used parametric statistics as data showed a normal distribution. We controlled for possible differences between Experimental and Yoked subjects' level of neophobia toward the board by comparing their latency to approach it with an unpaired *t*-test. In order to test whether contingency played a role in promoting subjects' exploration (Prediction 1) we compared the time spent on the board and the number of actions performed by Experimental and Yoked subjects during Phase 1 with unpaired *t*-tests. As part of Prediction 1, we also tested whether the lack of

congruent action-outcome contingencies could cause a decrease in Yoked subjects' interest toward the board. For both Experimental and Yoked group, we correlated the mean number of manipulative actions performed in each minute of Phase 1 with time (1-min time bins) and tested its significance with a Spearman rank correlation test. To test Prediction 2, two repeated measure Anovas were used to reveal whether Experimental and Yoked subjects preferred the box-opening (BO) rotation over the other non-box-opening (NBO) actions (namely, the lifting of the module whose rotation opened the box and the rotation and lifting of the other module). The dependent variable was the percentage of each action performed, and the repeated measure (within-subject effect) was the type of action in each module. *Post-hoc* analyses (Tukey's HSD test) were then performed to reveal where significant differences occurred. To test Prediction 3, for both Experimental and Yoked groups we correlated the total number of BO rotations performed by subjects in Phase 1 with the mean time to retrieve the reward in Phase 2, and tested the significance by means of Spearman rank correlation tests.

To test Prediction 4, we analysed subjects' performance during Phase 2 by comparing the percentage of reward obtained between Experimental and Yoked groups, and the mean time to retrieve the reward between Experimental and Yoked subjects who solved the task by means of unpaired *t*-tests. In these analyses we did not consider whether subjects performed wrong actions before the correct one (i.e., BO rotation). Therefore, we repeated the above analyses by considering only when rewards were obtained by performing the correct action as first action (i.e. "first-shot rewards"). Finally, to test Prediction 5 for both Experimental and Yoked group we tested whether subjects that solved the task improved their performance across trials by correlating the mean time to retrieve the reward with trial sequence (from 1 to 10) by means of Spearman rank correlation tests. All tests were run with STATA 10 (STATAcorp 2005) and alpha level was set at 0.05. All tests were two-tailed.

# Results

## Prediction 1

The latency to approach the mechatronic board did not differ between Experimental and Yoked subjects (Exp.: mean ± SE: 8.37±5.55 sec; Yoked: 14±7.04 sec; unpaired *t*-test: $t_{(14)}$=-0.62, *P*=0.541). This indicates that individuals were indeed well balanced between groups based on this parameter. Experimental subjects spent more time in contact with the board (mean ± SE: 7.65 ±0.96 min) than Yoked subjects (3.9±1.01 min, unpaired *t*-test: $t_{(14)}$=2.62, *P*=0.0198) and performed significantly more actions (mean ± SE: 45.25 ±9.38) than Yoked subjects (7.5 ±6.03, unpaired *t*-test: $t_{(14)}$=-3.38, *P*=0.004).

Experimental subjects did not decrease the number of manipulative actions across the 12 1-min blocks (Spearman rank correlation test: *r*=-0.35 *N*=12, *P*=0.254 Fig 2a), while Yoked subjects did so significantly (Spearman rank correlation test:: *r*=-0.60, *N*=12, *P*=0.0238, Fig 2b), confirming Prediction 1.

## Prediction 2

The percentage of actions performed by Experimental subjects differed significantly depending on the type of action ($F_{(3,7)}$= 6.956, *P*=0.004, Fig 3).

Specifically, post-hoc analyses revealed that the BO rotation was performed significantly more than the lifting actions (Tukey's test: lifting of BO module: *P*<0.05; lifting of NBO module: *P*<0.05), although not significantly more than the rotation of the NBO module (*P*>0.05). Therefore, Prediction 2 is only partially confirmed. By contrast, Yoked subjects did not show any preference for a given action ($F_{(3,7)}$= 0.545, *P*=0.559, Fig 3).

## Prediction 3

According to Prediction 3, an increased exploration in Phase 1 should lead to a shorter latency to solution in Phase 2. Experimental subjects that more frequently performed the BO rotation during Phase 1 were faster in retrieving the reward in Phase 2 (Spearman rank correlation: *r*=-0.71, *N*=8, *P*=0.046). By contrast, no correlation was found in Yoked subjects (Spearman rank correlation: *r*=-0.56, *N*=5,
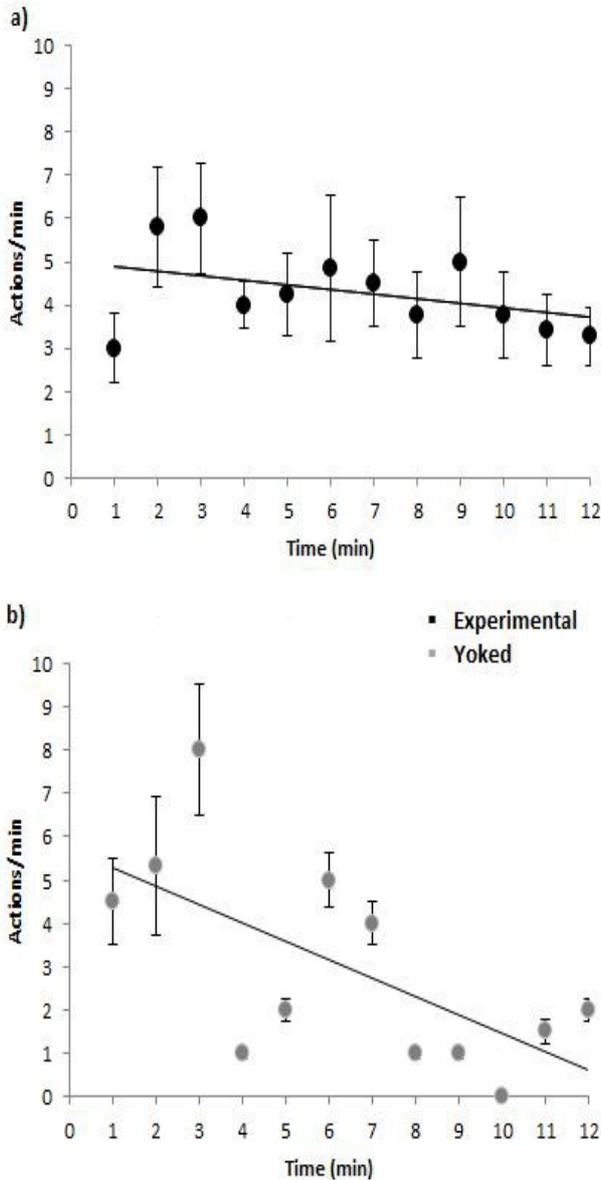
*P*=0.322).



a)



b)

**Fig 2** Correlation between the per-minute mean number of actions performed during Phase 1 and time for a) Experimental (*black*) and b) Yoked (*grey*) group.

*Prediction 4*

All Experimental subjects always retrieved the reward from trial 1 on, whereas only 2 Yoked subjects did so. In addition, 3 Yoked subjects were successful only in some trials and 3 never retrieved the reward. Overall, Experimental subjects obtained a higher percentage of rewards than Yoked subjects (unpaired *t*-test= $t_{(11)}$=3.29, *P*= 0.005, Fig 4a). Moreover, the time to reward retrieval was significantly shorter in the Experimental group than in the Yoked group

(unpaired *t*-test $t_{(11)}$=3.02, *P*= 0.012, see figure in Online Resource 2), confirming Prediction 4.

The analysis of "first-shot" rewards show that Experimental subjects outperformed significantly the Yoked subjects both in terms of percentage of rewards (mean ± SE Experimental: 60 ± 9%; Yoked: 25 ± 3%; unpaired *t*-test: $t_{(11)}$=2.79, *P*=0.027) and of time to retrieve the reward (mean ± SE Experimental: 6.6 ± 0.6 sec; Yoked: 61.8 ± 18 sec; unpaired *t*-test: $t_{(11)}$=-2.59, *P*=0.025).
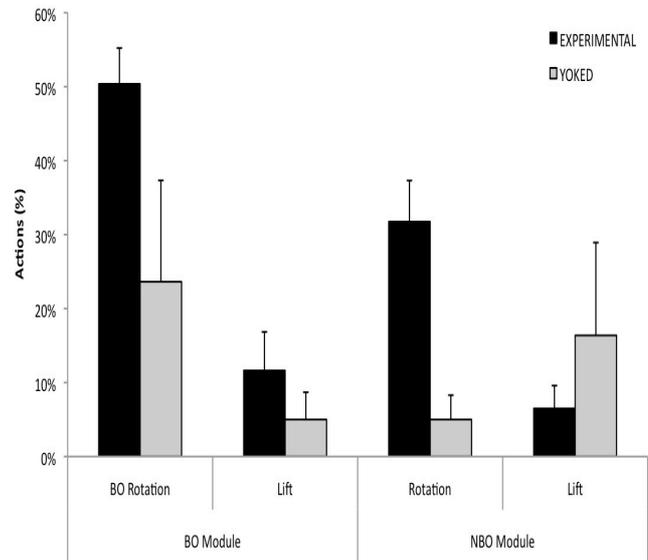


**Fig 3** Mean (+ SE) percentage of manipulatory actions performed by the Experimental (*black*) and Yoked (*grey*) subjects during Phase 1. *BO Rotation* box-opening rotation; *BO Module* module whose rotation is associated with the box opening; *NBO Module* module whose actions (rotation and lift) are not associate with the box opening.

*Prediction 5*

The time at which solvers retrieved the reward did not decrease significantly across trials in the Experimental group (Spearman rank correlation: *r*= -0.51, *N*=10, *P*=0.135), whereas it did so in the Yoked group (Spearman rank correlation: *r*= -0.64, *N*=10, *P*=0.042, see figure in Online Resource 3), thus confirming Prediction 5. The same results were obtained by taking into account the "first-shot" rewards (Experimental group: *r*= -0.40, *N*=10 *P*=0.244; Yoked group: *r*=-0.73, *N*=10, *P*=0.0162). Nevertheless, as only 2 Yoked

6

subjects retrieved the rewards in all the 10 trials and 3 subjects did so only occasionally, the robustness of the result for the Yoked group should be taken with caution due to the small sample size and the variability in performance among subjects that contributed to this analysis.

**Discussion**

Our findings demonstrated that action-outcome contingency experienced during free exploration helps capuchin monkeys to acquire skills that they later exploit for goal-directed purposes, thus supporting the idea that intrinsic motivation can sustain exploration and foster learning. Experiencing the contingency between actions and outcomes allows individuals to exercise control over their environment, and this is likely to influence many different aspects of individuals' behaviour. Indeed, there is evidence that control over the environment affects, learning, cognition, social skills and emotional responses in humans and other animals. Rhesus macaques exposed to a broader range of social/physical contingencies during their infancy, when tested later in life, showed better cognitive and social abilities than individuals with poor experience of contingencies (Capitanio and Mason 2000). Similarly, rhesus macaques that could control their environment had lower levels of fear and increased copying responses (Mineka et al. 1986). Finally, control over the environment (or perception of control) has a strong effect on children emotional, social and cognitive functioning (Gunnar 1980a,b).

By using the unrewarded exploration phase and the yoked-control paradigm, we were able to appreciate the role of spontaneous exploration (not extrinsically rewarded) and the role of experiencing congruent action-outcomes associations on learning, as never done by previous studies. The Experimental subjects were exposed to congruent action-outcomes associations throughout the experiment and this contingency played a fundamental role in sustaining their exploration. In contrast, the Yoked subjects experienced the contingencies produced by their paired Experimental subject and the outcomes were in most cases inconsistent with their own actions. As expected on the basis of Prediction 1, the Yoked subjects significantly decreased their interest and exploration of the board possibly because of the lack of control and congruence.

According to Prediction 2, the opportunity to open the box should lead subjects to perform the effective action more frequently than the ineffective ones. As expected, Experimental subjects performed the BO rotation more often than the NBO lifting actions. Since Yoked subjects did not show a preference for a specific action, the hypothesis that capuchins overall prefer rotations (regardless of the possible outcomes) over lifting actions can be rejected, and therefore Prediction 2 is confirmed. A possible explanation for the lack of preference toward the BO rotation over the NBO rotation could be that the association between action and outcome might establish before the association between location and outcome. This hypothesis seems to be confirmed by the results of a similar experiment involving a yoked-control paradigm with young children (Taffoni et al., unpublished data). When a mechatronic board equipped with three push buttons and three boxes was presented to the subjects, in Phase 1 experimental children learned the association between the effective action (i.e., pushing the buttons for at least 1 sec) and its outcome (i.e., boxes opening), whereas only in Phase 2 they learned the spatial association between the rewarded box (only one box was rewarded in each trial) and the correct button to push.

Thus, the opportunity to experience action-outcome contingencies leads animals to focus on the consequences of their actions and eventually to learn from them. In our experiment, experiencing congruent action-outcome contingencies had indeed a beneficial effect on subsequent learning. As expected on the basis of Prediction 3, subjects that more frequently discovered the rotation that opened the box in Phase 1 were also better at retrieving the reward in Phase 2. Similarly, when humans and rats freely explore a virtual environment to reach a hidden target area, the amount of exploratory movements performed positively correlated with subsequent competence in reaching the target, suggesting that exploration fosters action-learning (Stafford et al. 2012).

As expected on the basis of Prediction 4, Experimental subjects outperformed Yoked subjects in Phase 2, both in terms of percentage of rewards retrieved and time to solution. This suggests that action-outcome contingencies (and so, control over the environment) sustain exploration as well as learning in the absence of immediate benefits. Recently, research on the neural basis driving intrinsically motivated learning has focussed on the role of neuromodulator dopamine (DA) (Dayan and Balleine 2002; Mirolli et al. 2013; Schultz 1998; Fiorillo 2004; Hooks and Kalivas 1994). Redgrave and Gurney (2006) postulated that when novel and surprising stimuli are contingent with what an animal/agent does, the subsequent (phasic) release of DA in the brain can act as primary reward allowing the brain to learn action–outcome routines when the outcome has no immediate benefits. This claim fits well with the results of our experiment, in which subjects were required to recall actions learned in the absence of immediate benefits and recruit them to obtain food. As expected on the basis of Prediction 5, when congruent action-outcome contingencies were introduced to Yoked subjects in Phase 2 their performance improved across trials. By contrast, the Experimental subjects did not improve because they already learned the correct action-outcome association during Phase 1 and thus were already able to retrieve the reward in trial 1 of Phase 2. Strong inter-individual differences were present in Yoked subjects: though some of the Yoked subjects learned the task, three subjects did not even manipulate the modules. Given the small sample size, we cannot provide a strong interpretation for this result. Nevertheless, we can speculate that when outcomes occur regardless of what the subject does (as for Yoked subjects in Phase 1), then the subject's exploratory activities decrease as if experiencing a "learned helplessness" (Seligman and Maier 1967). Also *Drosophila* flies tested in a yoked condition in which they could not control the duration of heat pulse, quickly decreased their activity compared to experimental flies that could control the duration by resuming locomotion when the heat pulse occurred (Yang et al. 2013). It is worth noting that the reduction in activity of the yoked flies persisted after the removal of heat pulse, as it occurred in our experiment to the manipulation rates of a few Yoked subjects when action-outcome contingencies were restored in Phase 2 .

Interestingly, the three Yoked subjects who failed in Phase 2 were those that interacted less with the board in Phase 1, whereas those that succeeded were those that explored more during Phase 1. High levels of manipulation increase the chance of experience action-outcome contingencies. In the case of Yoked subjects, contingencies (though misleading) may provide positive motivational feedbacks and prevent inactivity (similar to the "immunization effect", Seligman et al. 1975). Thus, individuals that are more explorative to begin with, are better equipped to cope with the negative feedbacks due to the lack of control over the environment.

In conclusion, we demonstrated that intrinsically motivated exploration promotes learning and that exploration and learning are heavily impacted by the opportunity to control the environment, as suggested by previous theories (Glow et al. 1972; Glow and Winefield 1978). Future research should explore the extent to which intrinsic motivations promote skill acquisition during cumulative (i.e., sequential/hierarchical) learning tasks and evaluate the role of specific personality/temperament traits in activities where intrinsic motivations play a key role.

## Ethical standards

All procedures complied with protocols approved by the Italian Health Ministry (Licence no. 12/2011-C) and were performed in full accordance with the European law on humane care and use of Laboratory animals.

## References

Antunes M, Biala G (2012) The novel object recognition memory: neurobiology, test procedure, and its modifications. Cognitive Processing 13(2):93-110

Baldassarre G, Mirolli M (2013a) Intrinsically Motivated Learning Systems: An Overview. In: Baldassarre G, Mirolli (eds) Intrinsically Motivated Learning in Natural and Artificial Systems. Springer, pp 1-14

Baldassarre G, Mirolli M (2013b) Intrinsically Motivated Learning in Natural and Artificial Systems. Springer.

Barto AG Singh S Chentanez N (2004) Intrinsically motivated learning of hierarchical collections of skills. In: Triesch J Jebara T (eds.), International Conference on Developmental Learning (ICDL2004), Piscataway, NJ:IEEE, pp 112-119.

Blatter K, Schultz W (2006) Rewarding properties of visual stimuli. Exp Brain Res 168 (4):541-546

Buchanan-Smith HM, Badihi I (2011) The psychology of control: Effects of control over supplementary light on welfare of marmosets. Appl Anim Behav Sci 137 (3):166-174

Butler RA (1954) Incentive conditions which influence visual exploration. J Exp Psychol 48 (1):19

Butler RA (1957) Discrimination learning by rhesus monkeys to auditory incentives. J Comp Physiol Psychol 50 (3):239-241

Byrne G, Suomi S (1996) Individual differences in object manipulation in a colony of tufted capuchins. J Hum Evol 31 (3):259-267

Byrne RW (2013) Animal curiosity. Curr Biol 23 (11):469-471

Capitanio JP, Mason WA (2000) Cognitive style: problem solving by rhesus macaques (*Macaca mulatta*) reared with living or inanimate substitute mothers. J Comp Psychol 114 (2): 115-125.

Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. Neuron 36 (2):285-298

Deci EL (1975) Intrinsic motivation. Plenum. New York.

Fiorillo C (2004) The uncertain nature of dopamine. Mol Psychiatry 9 (2):122-123

Fragaszy DM, Adams-Curtis LE (1991) Generative aspects of manipulation in tufted capuchin monkeys (*Cebus apella*). J Comp Psychol 105 (4):387-397

Fragaszy DM, Boinski S (1995) Patterns of individual diet choice and efficiency of foraging in wedge-capped capuchin monkeys (*Cebus olivaceus*). J Comp Psychol 109 (4):339-348

Fragaszy D, Visalberghi E, Fedigan L (2004) The complete capuchin: the biology of the genus *Cebus*. Cambridge University Press

Glow PH, Roberts JE, Russell A. (1972). Sound and light preference behaviour in naïve adult rats. Austr J Psychol 24(2): 173-178.

Glow PH, Winefield AH (1978) Response-contingent sensory change in a causally structured environment. Anim Learn Behav 6 (1):1-18

Gunnar, M. R. (1980a) Contingent stimulation: A review of its role in early development. In S. Vevine, & H. Ursin (eds.) Coping and health. New York: Plenum, pp. 101-119.

Gunnar, M. R. (1980b) Control, warning signals, and distress in infancy. Dev. Psychol., 16, 281-289

Harlow HF, Harlow MK, Meyer DR (1950) Learning motivated by a manipulation drive. J Exp Psychol 40 (2):228-234

Hooks MS, Kalivas PW (1994) Involvement of dopamine and excitatory amino acid transmission in novelty-induced motor activity. J Pharmacol Exp Ther 269 (3):976-988

Hughes RN (1997) Intrinsic exploration in animals: motives and measurement. Behav Processes 41 (3):213-226

Kish GB (1955) Learning when the onset of illumination is used as reinforcing stimulus. J Comp Physiol Psychol 48 (4):261-264

Li S, Cullen WK, Anwyl R, Rowan MJ (2003)

Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. Nature Neurosci 6(5):526-531

Lynch Alfaro, J. W., Boubli, J. P., Olson, L. E., Di Fiore, A., Wilson, B., Gutiérrez-Espeleta, G. A., Chiou, K. L., Schulte, M., Neitzel, S., Ross, V., Schwochow, D., Nguyen, M. T. T., Farias, I., Janson C. & Alfaro, M. E. (2012a) Explosive Pleistocene range expansion leads to widespread Amazonian sympatry between robust and gracile capuchin monkeys. J Biogeography 39: 272-288.

Lynch Alfaro, J. W., De Souza, Silva, J. & Rylands, A. B. (2012b). How different are robust and gracile capuchin monkeys? An argument for the use of *Sapajus* and *Cebus*. Am J Primat 14: 1–14.

Manrique HM, Sabbatini G, Call J, Visalberghi E (2011) Tool choice on the basis of rigidity in capuchin monkeys. Anim Cogn 14 (6):775-786

Meyer, ME, Collins, MD (1971). Light deprivation and sensory reinforced behavior in chicks. Perceptual and motor skills, 32(2): 602-602.

Mineka, S., Gunnar, M., Champoux, M. (1986). Control and early socioemotional development: Infant rhesus monkeys reared in controllable versus uncontrollable environments. Child Dev., 1241-1256.

Mirolli M, Baldassarre G (2013) Functions and mechanisms of intrinsic motivations. In: Baldassarre G, Mirolli (eds) Intrinsically Motivated Learning in Natural and Artificial Systems. Springer, pp 49-72

Mirolli M, Santucci V, Baldassarre G (2013): Phasic dopamine as a prediction error signal of intrinsic and extrinsic reinforcements: A computational model. Neural Networks 39: 40-51

Ottoni EB, Mannu M (2001) Semifree-ranging tufted capuchins (*Cebus apella*) spontaneously use tools to crack open nuts. Int J Primatol 22 (3):347-358

Oudeyer PY, Kaplan F, Hafner V V (2007) Intrinsic motivation systems for autonomous mental development. IEEE Transactions on Evolutionary Computation 11 (2):265-286.

Panger MA (1998) Object use in free-ranging white faced capuchins (*Cebus capucinus*) in Costa Rica. Am J Phys Anthropol 106 (3):311-321

Pearce JM (2008) Animal learning and cognition: An introduction (3rd edition). Hove: Psychology Press

Perry S, Manson JH (2008) Manipulative Monkeys. Harvard University Press

Redgrave P, Gurney K (2006) The short-latency dopamine signal: a role in discovering novel actions? Nature Rev Neurosci 7 (12):967-975

Reed P, Mitchell C, Nokes T (1996) Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. Anim Learn Behav 24 (1):38-45

Rescorla RA (1968) Probability of shock in the presence and absence of CS in fear conditioning. J Comp Physiol Psychol 66 (1):1-5

Ryan RM, Deci EL (2000a) Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemp Educ Psychol 25 (1):54-67

Ryan RM, Deci EL (2000b) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 55 (1):68-78

Schembri M, Mirolli M, Baldassarre G (2007): Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot. Proceedings of the 6th International Conference on Development and Learning. London, Imperial College: E1-6

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80 (1):1-27

Seligman ME, Maier SF (1967) Failure to escape traumatic shock. J Exp Psychol 74 (1):1-9

Seligman ME, Rosellini RA, Kozak MJ (1975) Learned helplessness in the rat: time course, immunization, and reversibility. J Comp Physiol Psychol 88(2):542-7

Sirianni G, Visalberghi E (2013) Wild Bearded Capuchins Process Cashew Nuts Without Contacting Caustic Compounds. Am J Primatol 75 (4):387-393

Stafford, T., Thirkettle, M., Walton, T., Vautrelle,

N., Hetherington, L., Port, M., Redgrave, P. (2012). A novel task for the investigation of action acquisition. PloS one, 7(6), e37749.

Taffoni F, Vespignani M, Formica D, Cavallo G, Di Sorrentino EP, Sabbatini G, Truppa V, Mirolli M, Baldassarre G, Visalberghi E (2012) A mechatronic platform for behavioral analysis on nonhuman primates. J Integr Neurosci 11 (01):87-101

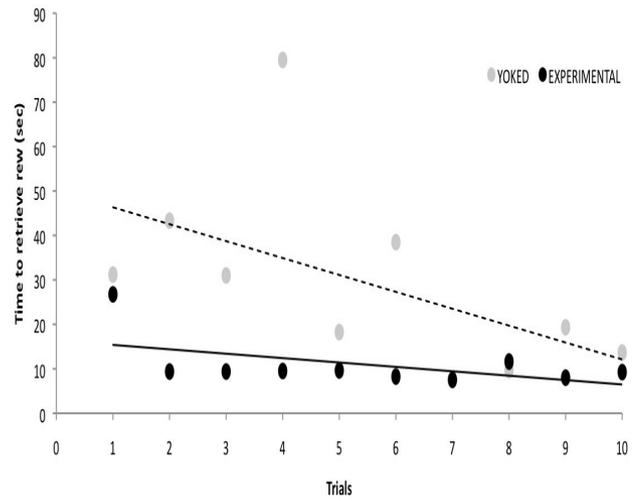Terborgh J (1983) Five New World primates: a study in comparative ecology. Princeton University Press

Visalberghi E (1988) Responsiveness to objects in two social groups of tufted capuchin monkeys (*Cebus apella*). Am J Primatol 15 (4):349-360

Visalberghi, E., Fragaszy, D. (2013). The Etho-Cebus Project: Stone-tool use by wild capuchin monkeys. In: Sanz, C., Call, J., & Boesch, C. (eds) Tool use in animals: cognition and ecology. Cambridge University Press 203-223.
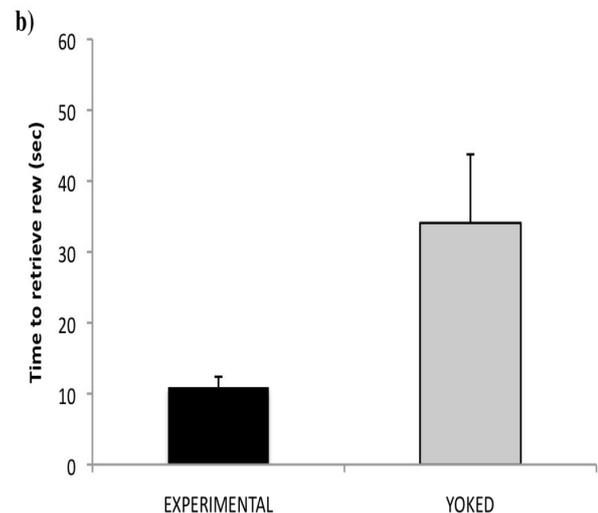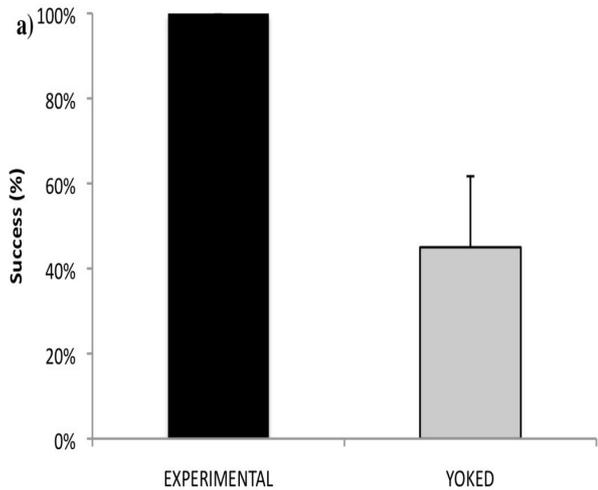
White RW (1959) Motivation reconsidered: The concept of competence. Psychol Rev 66 (5):297

Winefield AH, Glow PH (1980) Active versus passive control in sensory contingent bar pressing in rats. J Gen Psychol 103 (1):27-40

Yang, Z., Bertolucci, F., Wolf, R., Heisenberg, M. (2013). Flies Cope with Uncontrollable Stress by Learned Helplessness. Curr Biol, 23(9): 799-803.

## Supplementary material

**Exploration and learning in capuchin monkeys (*Sapajus* spp.): the role of action-outcome contingencies**

**Online Resource 3.** Correlation between the mean time to retrieve the reward and trial sequence for Experimental (*black*) and Yoked subjects (*grey*).



**Online Resource 2.** a) Mean (+ SE) percentage of rewards obtained during Phase 2 by Experimental (*black*) and Yoked (*grey*) subjects; b) Mean + SE time to retrieve reward during Phase 2 by Experimental (*black*) and Yoked (*grey*) subjects.