Paladyn, J. Behav. Robot. 2018; 9:374–390 DE GRUYTER

Research Article Open Access

Aliaa Moualla*, Sofiane Boucenna, Ali Karaouzene, Denis Vidal, and Philippe Gaussier

Is it useful for a robot to visit a museum?
The impact of cumulative learning on a robot population

https://doi.org/10.1515/pjbr-2018-0025
Received February 13, 2018; accepted October 9, 2018

Abstract: In this work, we study how learning in a special environment such as a museum can influence the behavior of robots. More specifically, we show that online learning based on interaction with people at a museum leads the robots to develop individual preferences.

We first developed a humanoid robot (Berenson) that has the ability to head toward its preferred object and to make a facial expression that corresponds to its attitude toward said object. The robot is programmed with a biologically-inspired neural network sensory-motor architecture. This architecture allows Berenson to learn and to evaluate objects. During experiments, museum visitors' emotional responses to artworks were recorded and used to build a database for training. A similar database was created in the laboratory with laboratory objects. We use those databases to train two simulated populations of robots. Each simulated robot emulates the Berenson sensory-motor architecture.

Firstly, the results show the good performance of our architecture in artwork recognition in the museum. Secondly, they demonstrate the effect of training variability on preference diversity. The response of the two populations in a new unknown environment is different; the museum population of robots shows a greater variance in preferences than the population of robots that have been trained only on laboratory objects. The obtained diversity increases the chances of success in an unknown environment and could favor an accidental discovery.

Keywords: neural networks, aesthetic preferences, pattern recognition, online learning, robot's population, chance discovery

***Corresponding Author: Aliaa Moualla:** Neurocybernetic team, ETIS, ENSEA, University of Cergy Pontoise, CNRS UMR 8051; E-mail: aliaa.moualla@ensea.fr
Sofiane Boucenna: Neurocybernetic team, ETIS, ENSEA, University of Cergy Pontoise, CNRS UMR 8051
Ali Karaouzene: Neurocybernetic team, ETIS, ENSEA, University of Cergy Pontoise, CNRS UMR 8051
Denis Vidal: Anthropologue, IRD-URMIS/Paris Diderot, EHESS

1 Introduction

Museums are places to conserve and protect collections of artefacts that represent our heritage, but they are also places to discover, learn, exchange, feel, and entertain without the usual functional constraints associated with objects of everyday life. If the educational role of science or history museums is easy to understand, the interest of visiting an art museum is more subtle.

We seek to demonstrate that exposure to artwork in museums can be used to develop a robot's capability to sharpen and differentiate its way of perceiving the world. In our case, we argue for a synthetic and developmental approach for the study of animal and human cognitive functions [1, 2]. The museum as a place for social learning might be a good framework to evaluate developmental architectures. However, current models are typically applied to isolated or discretized problems and tested in environments that favor a high success rate [3]. Laboratory conditions do not allow researchers to tackle (1) scalability issues nor (2) learning issues related to long-term learning. Working in a museum involves the need for the neural networks (NN) architecture to manage online and real-time social interactions with a multitude of non-expert users. The museum also offers the advantage of short-term interactions (few minutes) which greatly simplifies the constraints in order to obtain acceptable interactions.

Our work with the robot Berenson[1] is a continuing experiment developed in collaboration with the Quai Branly Museum (MQB) in Paris (a museum focusing on different cultures with an anthropological, sociological and artistic approach). Berenson is a mobile and expressive robot designed to be a new kind of visitor to the museum. Berenson develops its own preferences as the result of its interactions with human visitors (see Figure 1). Afterwards, it visits the museum, avoiding obstacles and expressing the preferences developed in the learning phase.

Philippe Gaussier: Neurocybernetic team, ETIS, ENSEA, University of Cergy Pontoise, CNRS UMR 8051; E-mail: gaussier@ensea.fr

1 The name of our robot comes from the nineteenth-century art critic Bernard Berenson.

We began the experiments six years ago with our bio-inspired sensory-motor architecture based on a neural network (NN). The architecture allows Berenson to learn association between the visual objects and the emotional value and to control the robot's direction in order to reach objects according to a given motivation. For example, the robot can be trained to go in the direction of preferred objects and away from objects associated with a negative value. We have shown that Berenson can develop an appreciation of aesthetic preferences by exploiting the principle of social referencing [4]. This skill [2] can be transferred into the NN model, allowing a robot to learn complex tasks without an explicit reward [7, 8].

Berenson's first role is to study the emergence of aesthetic preferences and to test the capability of a neuronal architecture to learn online in a complex social environment. In this paper, we are particularly interested in evaluating the effect of the emergence of aesthetic preferences on a whole population of robots. We argue that the learning variability offered by special environments such as a museum leads to robots' individuation. We also question the interest of teaching artificial systems using a single large database with the goal of improving their performance. Avoiding a uniform response of a population of individuals to an unknown situation increases its chances of success.

Figure 1: The robot "Berenson" in Quai Branly Museum during PERSONA exposition. Berenson generalises its learning on the visitors faces. Here, the visitor and the robot observe each other.

2 Social referencing is a developmental process allowing an infant to seek information from another individual and to use that information to guide his/her behavior [5, 6].

In the following sections, we introduce the case of a single robot by describing the experiment of Berenson in the Quai Branly museum, followed by the presentation of its sensory-motor architecture (used also on robots for the simulated experiment). Then we evaluate this architecture performance in object recognition. Next, we present the simulated experiment that represents the case of populations of robots, and compares the learning experiences of two robot populations. We end with conclusions on the presented work.

2 Related work

This section reviews first the importance and effectiveness of training variability from a psychological point of view. Then it presents various experiments using robots in a museum environment.

Many studies in psychology show that training variability can enhance performance in the long term and improves the capability of transferring that training to related tasks in modified contexts. Schmidt and Bjork [9] demonstrate this in both motor and verbal tasks. Another study [10] finds that the learning rate under conditions of related variation is significantly greater than under conditions of specialization or unrelated variation. The same idea is supported for motor tasks. For example, Wulf et al. [11] studied the effect of the type of practice on motor learning. They found that a variable practice, in general, facilitates recall and recognition for the novel task. Others studies [12] suggest that the durability and the transfer of performance are shown only when the mental procedures developed during training are re-established at testing. More recently, Gonzalez and Madhavan [13] argues for the benefits of categorical diversity in training on the detection of novel items in a visually complex cognitive task. The paper states that, "Humans seem more capable than automated aids at extrapolating from previous knowledge and engaging in adaptive decision making (thinking "outside the box") when faced with novel threat targets". We agree with this statement. Here, we try to go beyond classical object recognition to suggest that a robot could think outside the box in some way using its ability to generalize to the unlearned objects. For the second step, we show the ability of a population of robots to find the object classified as "important" (a kind of accidental discovery) when they have been taught in the laboratory and in the museum (cumulative learning).

In this paper, we are considering a simple case of an accidental discovery of a winner object. In philosophy of science, Karl Popper [14] analyzed the nature of scientific discovery. His work defends an evolutionary view of human innovation. Recently, some studies look forward to integrating computers in scientific reasoning. Some works [15–18] present automated data-driven systems that could discover scientific laws (algebraic equations). Sparkes et al. [19] used the AI techniques to automate some aspects of the scientific discovery process; in this work, the system iteratively creates hypotheses about a problem and later interprets experimental results (closed-loop learning).

Some other works focus on exploring and analysing the nature of the accidental discovery or the chance discovery. Prendinger and Ishizuka [20] presents the chance discovery approach in contrast with the knowledge discovery in databases. They discuss the chance discovery in an open system. They consider the human initiative (thinking "outside the box") as a distinguishing feature of chance discovery as opposed to the knowledge discovery in databases. In [21], a specific neural network for chance discovery was developed. In our case we focus only on the very first step of the chance discovery and we show how chance discovery can emerge as a controlled side effect of cumulative learning (or long life learning) when some environments are unrelated to the task and are associated with a high variability of labels (for instance, visiting a museum with artworks). Learning new categories and correct associations from the success or failure of the trial is not addressed in the present work but it could be performed using our NN architecture.

At another level, many studies focus on the application of robots in museums. For instance, a robot mediator can guide the visitors in the museum. Robots have sensors allowing them to navigate and sometimes interact with visitors. They move along predefined routes and avoid obstacles. One of the first studies in a museum [22] was to develop a model of navigation and localization for a mobile robot. Other studies [23, 24] have taken advantage of certain specific environments such as museums, fairs, and exhibitions to test their algorithms. Their objective was to test existing algorithms in complex environments to highlight the limitations of the algorithms. For example, Ogata's study [25] discusses the communication between autonomous robots and humans through the development of a robot which has an emotion model. On the learning side, Marsland [26] presents a method for performing unsupervised online novelty detection. The method is based on learning to ignore inputs that have been seen before, so that novel inputs are highlighted. Finally, Csikszentmihalyi [27] proposed that the processes essential to creativity are not only to be found in the minds of creators but also in the interactions between individuals and their socio-cultural environment. There is also evidence that variable categories help in generalizing to novel members of each category.

3 Berenson experimental setup

The experiment with Berenson took place at the Quai Branly Museum for 10 days, for 4 hours each day. We limited our experiment to a zone about 800 m^2 corresponding to (40m x 20m) and museum mediators were there to assist in carrying out the experiment. Since the environment is dominated by art, we dressed the robot in a long dark coat and a bowler hat corresponding to an early nineteenth-century style. Berenson is built from a Robulab 10 platform and an expressive head. The Robulab platform from Robosoft embeds a computer [3] and supports the human shape of Berenson. The height of Berenson is 1.80 m, and its weight is almost 20 kg. To avoid obstacles, Berenson is equipped with proximity sensors, 15 infrared sensors and 9 ultrasonic rangers that are arranged all around him. One laser range sensor is also used to avoid human visitors at a smaller distance. One pan camera and one magnetic compass, located on Berenson's head, are used for navigation. Furthermore, the robot has another camera in its right eye to perform artwork recognition task. Berenson's expressive head has 9 degrees of freedom (DoF): 4 for the eyebrows, 3 for the mouth, 1 for the front tilt and 1 that allows the eyes to tilt. The control of the expressions is performed by an SSC-32 (serial servo controller able to control up to 32 servo motors). The embedded computer runs sensors, actuators management (including camera control), obstacle avoidance and artwork recognition.

During the training phase, visitors were asked to select an artwork they found more interesting or impressive than the others (positive appreciation) or on the contrary less interesting than others (negative appreciation). Using a two-button mouse, the mediator assigned each visitor's appreciation value to the object. Berenson associated the object with the given emotional value. Next, if no order was provided to the robot, it went toward artworks associated with a positive value and ignored the objects with a negative or neutral value. Throughout the experiment, the robot displayed either a positive expression (smile), a negative expression (sadness) or a neutral expression according to the value associated to the objects in the field of view of its camera.

3 An Intel i5 processor at 2.50GHz with 4 GB of memory.

4 Berenson sensory-motor architecture for visual processing

Our model is based on a simple neural sensory-motor architecture named PerAc (Perception Action) [28, 29], which allows on-line learning of sensory-motor associations. Figure 2 shows the generic architecture; it involves two data streams associated respectively with perception and action. In the low-level pathway, we suppose we can extract reflex or involuntary information to directly control the robot's actions. The conditioning pathway allows the robot to anticipate the reflex or the involuntary behavior through the learning process. During learning, associations are formed between the recognition of sensory information (high-level) and the reflex or involuntary behavior (low level). By associating the recognized situation with the unconditional input during the learning phase, the system can recognize a situation and react correctly and avoid the reflex or involuntary pathway.

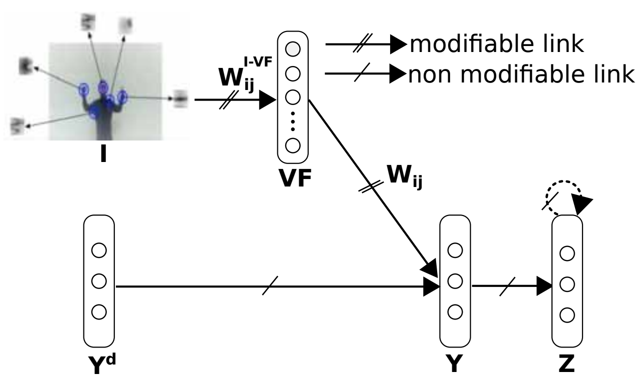

Figure 2: Sensory-motor architecture (PerAc).

Berenson's global sensory-motor architecture includes:

1. the social referencing model (estimation of the *What* channel) performed by using a cascade of PerAc architecture which allows the learning of more and more complex skills;
2. the object pose estimation model (estimation of the *Where* channel);
3. the navigation model (not discussed in this paper).

Next, we briefly present the first and second models constituting the bio-inspired visual system where two types of visual information are processed in parallel: the *What* and *Where* information.

4.0.1 *What* channel estimation

A bio-inspired visual system based on a sequential exploration of focus points allows the learning of different local views, or what we call the *What* information. For each focus point in the image, a local view centered on the focus point is extracted and transformed in log/polar coordinates to allow robustness to scale variation and small perspective changes. The extracted local view is learned and categorized by a group of neurons VF (visual features) using a K-means variant that allows online learning and real-time functions called SAW (Selective Adaptive Winner) [30]:

$$VF_j = net_j \cdot H_\gamma(net_j) \quad (1)$$

The VF_j is the activity of neuron j in the group VF and γ is a vigilance parameter (the threshold of recognition), while $H_\theta(x)$ is the Heaviside function. Incoming local views are compared with learned patterns. The value of net_j is the complement to 1 of the sum of the distances between the input feature and the nearest similar learned feature (see Appendix for the equation). If the maximum activity net_j is below γ (the given threshold), the current local view is learned as a new pattern and associated with a newly recruited neuron (incremental learning). Otherwise, the SAW algorithm adapts the links between the winner neuron and the input pattern. The modification of the weights (W_{ij}) is computed as follows:

$$\Delta W_{ij}^{I-VF} = a_j(t)I_i + \epsilon(I_i - W_{ij})(1 - VF_j) \quad (2)$$

Where I_i is the input visual feature. When a new neuron is recruited $a_j = 1$ and otherwise $a_j = 0$. In the use mode, the robot computes the distance between the current local views and the learned views.

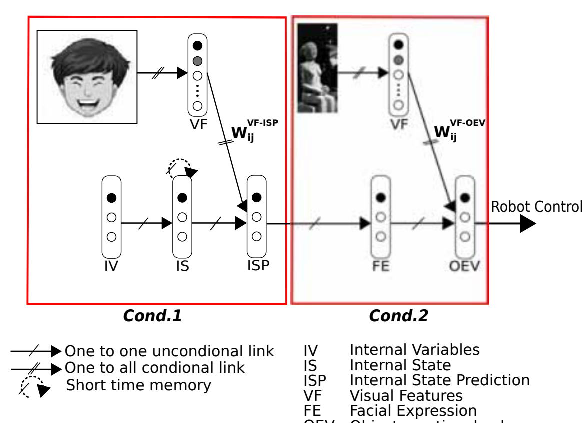

Figure 3: Social referencing model. The recognised facial expressions is used as an unconditional stimulus to associate an emotional state with an object or a scene.

The *What* channel associates the visual features VF extracted from a scene and an emotional value EV (indicated by a visitor). The emotional value association is supervised. It is computed thanks to a Pavlovian conditioning [31] based on the Least Mean Square (LMS) rule [32] as formalized in Equation 3. The LMS minimizes the error between the desired output (one-to-one unconditional link) and the current output (one-to-one conditional link). The groups of neurons ISP and OEV use the same conditioning algorithm (see Figure 3):

$$\Delta W_{ij}^{VF-OEV} = \epsilon_1 \cdot VF_i \cdot (FE_j - OEV_j) \tag{3}$$

Our social referencing model is performed using two PerAc architectures:

1. Facial expression learning: this is based on the hypothesis that the robot makes random facial expressions according to its internal state and the human imitates it during the learning phase (2-3 minutes). The sensory-motor model learns the visual conditioning between the robot's internal state (triggering the facial expression) and the visual feedback provided by the human [33, 34] (see Figure 3: Cond.1). The ISP group associates the human facial expression VF to the robot internal state IS (see equation in Appendix), allowing the robot to recognize and respond to human facial expressions in use mode.
2. Emotional values and objects association: the second conditioning (see Figure 3: Cond.2) is used to associate the visual features of an object with a positive or negative emotional value. The unconditional stimulus is supposed to be the emotional value. It comes from the robot's internal state or from the recognized human facial expression. The OEV associates the recognized facial expression FE to a visual scene VF. At this developmental level, the robot can associate emotional values with each object.

In the museum experiments, we simplified the referencing problem by asking the visitors to use a two-button mouse to directly provide the emotional value (positive or negative) to the artworks when learning them. After several learning iterations, stable connections are reinforced, and non-pertinent associations are decreased. For instance, if a distractor is present in the background of an artwork, it will be associated first with the same value as the artwork. The dissociation occurs if the distractor is present both with positive and negative values. Berenson's facial expression is then the result of the temporal integration of the decisions for each local view. To be acceptable, the frame rate of the global visuomotor loop must be higher than 10 Hz (16 Hz in our case). This allows Berenson to filter the wrong associations.

4.0.2 *Where* channel estimation

In our model, an object is considered as an ensemble of local views. Estimating the object position in the camera's field of vision corresponds to estimating the relative positions of its components according to a given reference point. Figure 4 represents the fundamental schema of the *Where* information processing. The groups θ, θ_L and θ_S use population coding for angle computation. The neurons in the θ group correspond to the focus point's position along the x coordinates in the input image. The θ_L group corresponds to the learned position and the θ_S group corresponds to the shifted position.

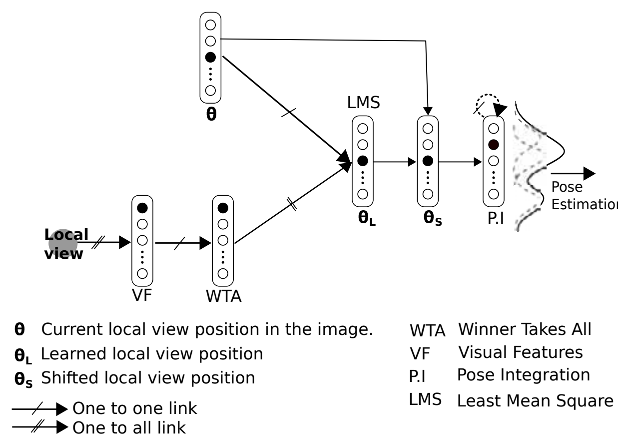

Figure 4: Architecture model for object pose estimation.

During the learning phase, the robot is oriented manually to center the object in the camera field of view. The θ_L group associates each local view explored with the center of the image thanks to the LMS rule (see Equation 4):

$$\theta_L = [W] \cdot VF \tag{4}$$

The learning between neurons associated with the view recognition and the angular position of the object center is done via one-to-all conditional links. They work like a memory, storing the local view positions (see details in Appendix).

The robot chooses a winner direction to control its orientation. It should be associated with the winner value of an object. The system is able to generalize by associating an emotional value to unlearned objects including visitors' faces.

Using only the integrated activity PI cannot lead to the real winner in many cases because it depends on the number of recognized local views. We give two examples of cases which cause problems:

1. many learned objects with a different number of focus points are present in the robot's field of view;
2. the presence of distractors in a heavily textured environment (the number of local views can increase a lot in this case).

To enhance the object recognition and the coherence of winner selection, a modulation with the emotional value was added to the model. Further, a normalization of the object pose estimation fields, and a competition mechanism were also necessary. Those steps are detailed in [35] and Figure 5 represents the final model used on Berenson in the museum experiments. The same model is also used on all of the simulated robots for the following comparison between robot populations.

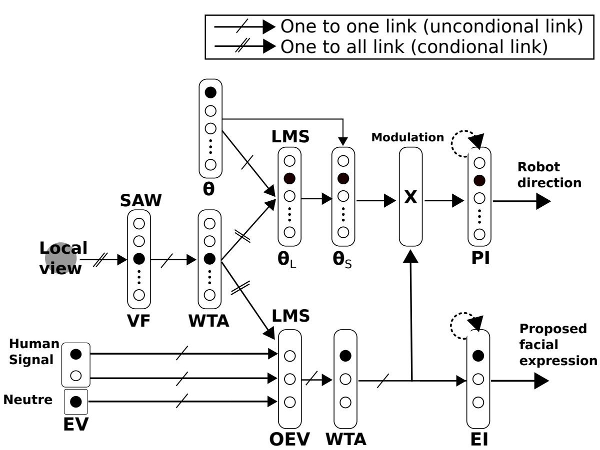

Figure 5: Berenson global sensory-motor architecture.

As mentioned above the robot assigns an emotional value to some artworks, or to some of the visitors' faces that present some similarities with the learned objects. In test mode we put a low vigilance that allows this generalization to unlearned objects. Consequently, the robot can attribute valences to objects it has never seen before. This ability to generalize to unlearned objects is very interesting, as it allows the robot to generalize learned aesthetic preferences.

In the next section, we present a detailed evaluation of the visual system performance in a museum environment followed in Section 6.1 by three summarized tests of the performance needed to approve the populations' comparison conditions.

5 Evaluating Berenson's visual system performance in the museum

Once the experiments at the museum were over, a database was created during an offline analysis of the images recorded by Berenson. An expert used a mouse to provide the object position (the object center) and the correct label (emotional value) for each object in each image of the database. It was found that the objects most frequently indicated by the visitors had the largest number of both positive and also negative values. For instance, the same object was labeled as a negative object by seven visitors and as a positive object by seven other visitors (see Figure 12). Then, the emotional value annotation of objects was changed offline by defining four different emotional categories. These were as follows: negative (objects few selected but mostly negative), positive (objects few selected but mostly positive), surprising or interesting (thus selected both negative and positive objects) and neutral. The surprise category was used to define interesting objects with bimodal emotional value distributions, while the neutral category was used for objects of no particular emotional value such as walls and doors. This was to stop the generalization when the robot was in front of such objects. There is a specific color for each category, blue for negative, red for positive, yellow for neutral, pink for surprising or interesting (see Figure 6).

Figure 6: (Top) represents the learning mode, local views associated with the emotional value. (Bottom) represents the test mode, Berenson associates the recognized object to the learned emotional value. Here the images are taken from the same test but with different number of local views (8 were extracted in learning mode and 20 in test mode), the cercle's color represents the emotional value.

As in the online experience, in the offline test (using the database) five local views per image were extracted and learned in the learning mode, and fifteen local views were extracted in the test mode. The larger number of extracted local views in the test mode is due to the order of focus points' extraction. The order of focus points' extraction is only related to the local contrast and the local curvature extracted thanks to the DOG filter. In an unknown image (test image), there is no reason that the first selected focus points correspond to a learned object (see Figure 6). Hence, in the test image, if there are three objects of similar complexity to one of the learned objects, then the robot should focus on the learned object and off the usual exploration. The issue will be to ensure that distractors will not win or change the winning label because of some wrong generalization over the unknown objects.

We limit the number of neurons or units in the SAW group to 500 neurons which means the learning of 500 local views. This corresponds to the capability to learn 100 images of objects with 5 local views per object; 100 objects if we suppose one view is sufficient or that the robot cannot turn around the object. As there were twenty artworks present in our experiment, these numbers are entirely sufficient. This allows the learning of several 2D views of the same object and to deal with the possible distractors. As discussed above, the vigilance parameter determines how fine the categories will be (see Equation 1). The vigilance parameter is in the interval [0, 1], and was set to 0.98 in the learning mode, and to 0.78 in the test mode. Low vigilance in the test allows a generalization to the unlearned objects. The generalization is one of the bases of our thinking about artwork appreciation and aesthetic experience because we want to observe how Berenson will generalize when confronted with new objects. After the SAW, a Winner Takes All mechanism (WTA) selects the neuron with the maximum of activity and inhibits the others.

Two measurements were done. One was the object recognition rate without taking its position into account, and the other was the object recognition rate taking its position into account. When the object position is taken into account, if the object is found to be too far away from its correct location the winning category will be considered as false. An object is considered as recognized if and only if its label corresponds to the label proposed by the expert and if the distance between the center proposed by the robot and the center proposed by the expert is less than 30 pixels. This means that we accept an error in the position of 10% maximum of the image size (the width of the image being subsampled to 320 pixels to speed up the computations). We also note that when the object is large in the camera's field of view, the expert precision to locate the object center is very low (the variance can be superior to the 30 pixels) and this can create false results. Regardless, this measure provides a good indicator of the robot's capability to navigate in the direction of the correct object.

Figure 7: Learning mode, (Top) and (Middle) represent the objects learning from different distances and angles. (Bottom) represents the object 1 learning from different angles and distances. The circles represent the object's identity.

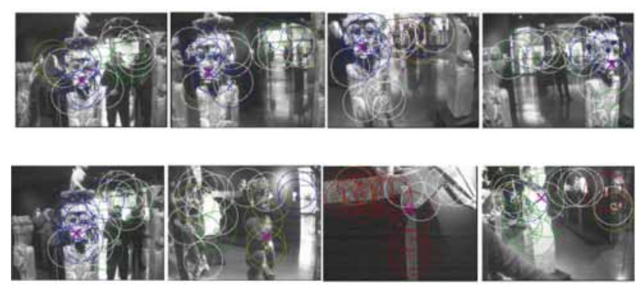

Figure 8: (Top) the object 1 recognition from different angles in test mode. (Bottom) objects recognition in test mode. The circles represent the object's identity.

For the offline learning phase, 13 objects were used with 10 images per object that represent the object under different distances and angles (see Figure 7). The 130 images were extracted from the images recorded by the robot during the online learning episodes in MQB. The conditions are not exactly the same for each object since, according to their location in the museum, some of the objects were visible from long or short distances, while others were located in the center of the area of experience, and were visible from multiple points of view. The different images were associated with four labels: neutral, positive, negative and surprise, as mentioned above. For the test phase, we used 1300 test images randomly selected from Berenson's recording in MQB (in some cases, they corresponded to non-learned objects) (see Figure 8).

The identity of the different objects were correctly found for 892 images (success rate of 68%), but if we take into account the position of the objects within a certain threshold (less than 30 pixels of error), the objects were correctly identified in only 747 images (success rate of

57%), while random decision provides a 25% success rate (see Figure 9 for detailed results).

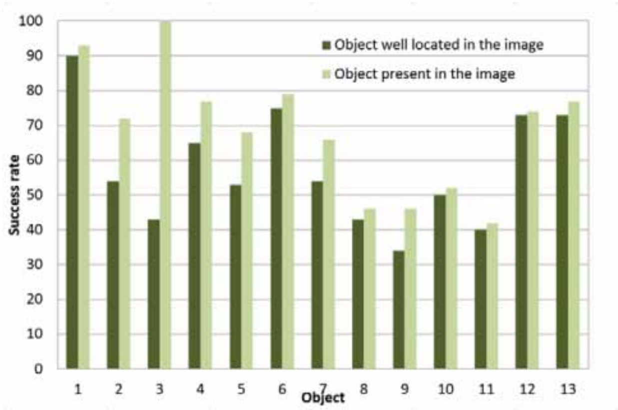

Figure 9: The histogram represents the comparison of two success rates. First, find only the correct object in the image and second, find the correct object and its position in the image.

By analyzing the detailed results shown in (Figure 9), it can be seen that using only the *What* information leads to a high rate of error (false positive) in some cases. For example, the third object for which the robot is over generalizing. The robot can move toward another object and express an emotional value by believing that it is the third object, thereby inducing a behavioral error. This case shows the importance of using both the *What* and the *Where* information together. Other objects such as 1, 12 or 13 have an almost equivalent recognition rate for the two tests (with and without taking the object position into account). Of note is object 8, which seems to show one of the limitations of our model. It seems hard for the system to find Object 8 in the image. However, once found, its position is also well recognized. Looking to the general success rate where the objects were correctly identified and localized (success rate of 57%), the system has only one wrong response every three times, which is sufficient to have a correct behavior. In dynamical conditions during the online experiments, the motor control is much slower than the image processing (more than three images from almost the same position). Then the processing speed of our system (20 images per second online) implies that the robot is not wrong even with 57% of success rate.

6 Effect of cumulative learning in a special environment on population diversity

Now that we have shown that our NN architecture can learn and recognize objects, we can study the effect of a museum visit on the long-term performance of either a robot or a system trained in a laboratory environment. Once the learning is finished, the robot should be able to recognize the emotional value of a learned object and to generalize this to new objects. To illustrate the advantages of the learning in a specific environment like a museum, we propose to compare the variation between two populations' responses in the generalized case to a new environment. We demonstrate that exposure to artwork in museums can be used to develop the robot's capability to sharpen its preferences and to differentiate from others (individuation). The development of individual preferences in a museum using a social referencing approach has an important effect when we take the whole population into consideration. If we imagine a population of individuals that have been taught exactly in the same way, all these individuals will have the same preferences, and they will answer any question in the same way. The online learning based on a social referencing approach leads the robots to develop different preferences. After this kind of learning, we obtain a large diversity of individual answers that could maximize their chance of success when facing new objects and new problems.

In this section, we consider a population of robots that have learned mainly in the laboratory, and we divide this population into two groups. The first group will have a second learning experience in the same laboratory while the second group will have a second learning in the museum. For each learning experience, the robots associate local views of an object with an emotional value (positive or negative) using the same architecture implemented on Berenson, as discussed above.

6.1 The learning task

For each learning experiment, we have a set of annotated images used for the learning databases. Two main databases were created. A laboratory database, and a museum database. In the laboratory database (see Figure 10), each object is annotated with a negative or a positive value that will not change depending on our emotional state. In our laboratory's objects annotation, the value was chosen

Figure 10: Examples of objects learned in our laboratory.

Figure 11: Examples of objects learned in Quai Branly.

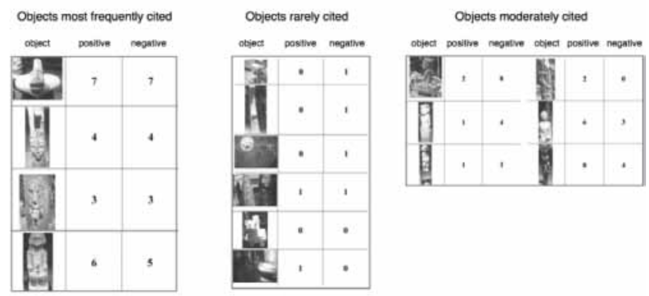

Figure 12: Objects' citation at Quai Branly Museum provided by the visitors during the experiments.

according to the fact that the robot has to reach or avoid these objects. Consequently, their emotional values are the same for all of the robots. For example, a drill press is a negative object because it can be dangerous. We will call this a type of "functional learning".

In the museum database (see Figure 11), all annotated images are taken from a previous experiment where the robot Berenson learned from interactions with the visitors. As the experiment was started from scratch every day, we obtained eight different museum databases. According to the visitors' preferences, some days one object was learned as positive while on another day it was learned as negative. As a result, the same artwork has been labeled with a positive value, or a negative value corresponding to the visitors' opinion (see Figure 12). As previously explained, the objects with higher positive scores were also the ones with higher negative scores. Approximately, the objects that were overall more cited were equally cited positively and negatively.

We simulate an initial population P of 16 robots that had learned in the laboratory during a first period ($T1$). The learning speed in Equation 3 was set to ($\epsilon_1 = 0.3$). For this learning, we used 1024 images from the laboratory database representing 60 different objects (17 images per object). Then, during a second learning period ($T2$) this population was divided into two subpopulations each containing eight robots. We denote by $P1$ the population that has been through a secondary learning in a laboratory, while $P2$ refers to the population that did the secondary learning in a museum. The learning speed in $T2$ was increased to ($\epsilon_1 = 0.6$).

During the second learning period ($T2$) of $P1$ in the laboratory, each robot learns from new images of the same objects used in the first learning period ($T1$). The number of images used in $T2$ is the same as the number of images per day visit in the museum (between 6 and 36 images). This is to avoid the possibility of the first group of robots learning more than the second group of robots.

The robot population $P2$ continue a second learning period ($T2$) at MQB (see Figure 11). The database contains 128 images that represent 20 different artworks. Each of the simulated robots has learned a set of images containing between 6 and 36 images which correspond to one experimentation day of Berenson in the museum. Hence, each robot in $P2$ has a small but specific learning database. Each artwork is then associated with a value corresponding to the subjective visitors' opinions on the selected day of Berenson at MQB. As a result, the database built from eight days of experiments is used as if we brought a new robot from the lab each day to learn from different visitors. The learning speed was increased to 0.6 in $T2$ to compensate for the fewer number of images to be learned compared to the $T1$ period ($T2$ is supposed to be a shorter experiment) but also to give some importance to the novelty of the museum experience. Changing the learning speed can be seen as increasing the robot vigilance because of the novelty of the museum experience. A novelty detector could have been used to automatically change the epsilon value.

Before going further we needed to evaluate the $P2$ cumulative learning. We had to assure that the second learning period ($T2$) in the museum doesn't significantly affect the main learning ($T1$) in the laboratory. We evaluated the performance of the population $P2$ on the museum objects recognition, and we compared its performance on laboratory object recognition before and after the museum experience.

First, we checked the $P2$ performance on laboratory object recognition before the museum experience ($P2$ passed only through the initial learning ($T1$) in the laboratory). We used 1500 test images representing 60 laboratory objects. The performances of $P2$ population's individuals (robots) is very close to each others (see the first column of Table 1). The success rate is about 94% for the whole population of robots. The better performance of our visual

Table 1: The first and the third columns represent the performances of $P2$ robots before and after the learning at the museum. The second column represents the performance of the $P2$ robots tested on learned museum objects.

$P2$	$T1$/laboratory	$T2$/museum	$T2$/laboratory
$R1$	95%	100%	95%
$R2$	95%	81%	95%
$R3$	95%	100%	95%
$R4$	94%	82%	94%
$R5$	94%	93%	94%
$R6$	93%	89%	93%
$R7$	94%	100%	94%
$R8$	95%	88%	95%

system in this test is due to the uniform background of laboratory object images (see Figure 10).

After the second learning period ($T2$) in the museum ($P2$ passed through the initial learning ($T1$) in the laboratory and the second learning ($T2$) in the museum), we tested the recognition performance of each robot of $P2$ on the museum set of images used for its learning (to verify the robots learning). The performances are up to 100% for some robots and not less than 80% for the whole population's individuals. These very high performances are explained because we used exactly the same dataset for the learning and for the test (see the second column of Table 1).

After being sure that all robots learned some museum artworks, we repeated the first test of laboratory object recognition after the museum experiment ($P2$ passed through the initial learning ($T1$) in the laboratory and the second learning ($T2$) in the museum). The $P2$ performances showed very similar results to the first test. The performances of almost all of the population's individuals in laboratory object recognition was still about 94% (see the third column of Table 1). This shows that the museum learning doesn't affect the first laboratory learning, as the difference of recognition level of laboratory objects before and after museum learning is less than 1%. Once the $P2$ cumulative learning is evaluated, we studied its effect on a population of robots. The next section details the chosen test, thereby allowing us to evaluate the impact of the cumulative learning at a special environment such as the museum on population diversity.

Figure 13: Examples of test images that represent new environments.

6.2 Formulating the test

In the simulation, various possible tests can be chosen to analyze the effect of online learning in the museum on the development of the robots' preferences. In our case, the idea is to arbitrarily choose a target object in a new environment, and to measure whether the robot chooses the same object. We then compare the populations' diversity and their success rate (the population success depends on whether at least one of its individuals find the winning object). However, we know that if we use images in this test, all robots will browse the focus points one by one in the same order, and they will associate one focus point at each iteration with an emotional value. Then, we did not simply define a winner focus point but a winner area in each image of test (test images represent new environments). The winner area of each test image will be associated with a winner emotional value. We tested at which iteration the robots would associate a point of the winner area with the winner emotional value, and we named this iteration *the iteration of matching*. The difference between the iterations of matching of each robot reflects the difference between their answers, and reveals their individual preferences.

We arbitrarily chose completely new test images taken from completely different environments to test all the robots (see Figure 13). It would also have been possible to test the differences of preferences with images from the MQB database, but we think that testing the generalized learning on a new database is far better at showing the building of individual preferences. In each test image, we arbitrarily selected a focus point to be part of a "target" (or goal) object, with X_g, Y_g representing the position of the winning focus point. We will annotate it arbitrarily with a winner emotional value (negative or positive), see Figure 14:

$$V_{goal} = V(X_g, Y_g)$$

We defined a rectangle representing the winning area around the winner focus point, allowing a range of error tolerance. This is defined by a threshold of error in x and in y goal position ($\Delta X_g, \Delta Y_g$). In this way, every robot could propose the correct emotional value in the winner area but during a different iteration. In order to compare *the itera-*

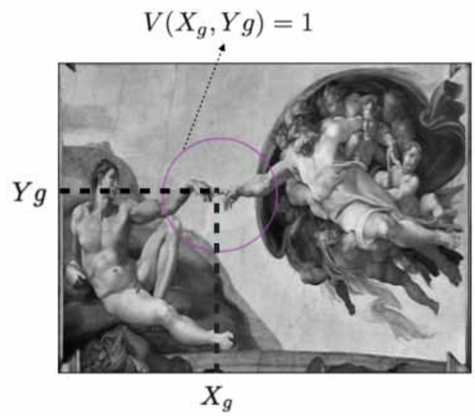

Figure 14: The image represents the position of the goal and the winner emotional value, where positive emotional value = 1, and negative emotional value = 0.

tion of matching of each robot (i.e. when the robot associates any point of the winner area with the correct emotional value) we used the following Equation 5, where i annotates the iteration order and $V(X_i, Y_i)$ annotates the emotional value related to the selected focus point:

$$Nb_{iter} = Argmin|V(X_i, Y_i) - V(X_g, Y_g)| \qquad (5)$$

where $|X_i - X_g| \leq \Delta X_g$, $|Yi - Yg| \leq \Delta Y_g$, Nb_{iter} is the iteration of matching and $i \in Scanpath$.

For testing and comparing the robots' answers, eighteen images taken from completely different environments (new environments) were chosen. A winning zone and a winning emotional value were annotated for each test image. The threshold that represents the range of error tolerance was set to $\Delta X_g = \Delta Y_g = 0.08$.

The test was repeated on a larger test database. We used 106 different images representing completely different environments. As for the first test, a winning zone and a winning emotional value were annotated for each test image and the same constraints were respected.

In the next section, we present the first test results (18 test images) and the second test results (106 test images). We use different calculations (the minimum iteration of matching, the mean value of the robots' iterations of matching, the standard deviation of the robots' iterations of matching and the mean absolute deviation of the robots' iterations of matching) to compare the variations in the answers for the two populations.

6.3 Experimental results and evaluation

6.3.1 First test results and evaluation

We compared the mean of the iterations' averages (each iterations' average represents the iterations of matching of a population's individuals averaged for a single test image), and we see that this number does not change between the main population P and $P1$. It is equal to 82 iterations, but for the $P2$ population learned in the museum, it was slightly higher, 85 iterations.

For each test image (18 in total), we calculated the deviation of the iterations of matching of a population's individuals. First for the initial population P learned at the laboratory, then for $P1$ and $P2$. After that, we averaged those deviations of iterations of matching for all of the test images. This average was equal to 47 for P and $P1$ populations, and increased to 57 for $P2$. This gives us an indication about the highest diversity of choices for the museum learning population $P2$ compared with $P1$ (see Table 5 in Appendix for the $P2$ robots' answers). We can even infer that the preferences of Robot 1 are closer to the winning preferences because it has the minimum average of iterations of matching.

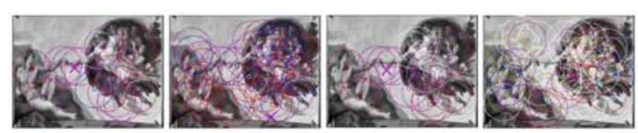

Figure 15: Examples of images that represent the different answers of four robots on a test image. These robots learned in the museum and we notice their different preferences by the colors of the circles representing the emotional category associated with local views.

Table 2: Results of testing on 18 images.

Learning	$P/T1$	$P1/T2$	$P2/T2$
ϵ_1	0.3	0.60	0.60
Mean of iterations	82	82	85
Deviation	47	47	57
Mean of min of iterations	35	35	24

For each population, we considered the minimum of all iterations of matching for each test image. Next, we calculated the average of those minimums for all of the 18 test images. This average was equal to 35 for P and $P1$ popu-

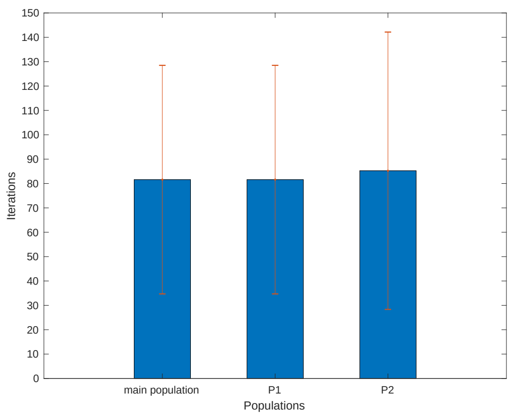

Figure 16: The figure displays the mean and the standard deviation of the average of iteration of matching to find the target object for the robot population over the 18 test images.

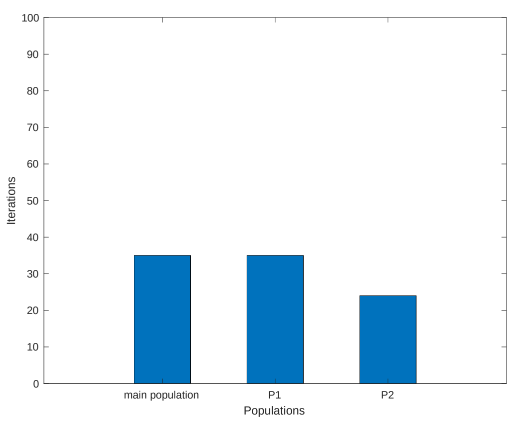

Figure 17: The figure displays the mean of the minimums of iteration of matching to find the target object for the robot population over the 18 test images (the lowest being the best).

lations, and to 24 for $P2$. This result shows that the robot population with a small experience at the museum could find the goal faster than the population taught only in the laboratory. These results point out that learning in the museum increases on average the generalization capability of the robots immersed in this environment. The robots with a second learning phase in MQB show a higher generalization capability than the robots that were only taught in the laboratory (see Figures 16 and 17, as well as Table 2). Moreover, a population with a learning phase in the museum can find the goal faster.

6.3.2 Second test results and evaluation

The same calculations as for the first test were repeated for the 106 test images. As in the first results, there is an indication of a higher diversity of choices for the robots that have gone through a second learning in the museum. This second test made on a larger database showed a clearer difference (significant difference). The average of diversity (the average of the deviations of iterations of matching for all the test images) reflects a clear difference between the two populations $P1$ and $P2$. For the P and $P1$ populations we had 30.5, while for $P2$ we had 45.2. The averaged minimums of all iterations of matching was equal to 46.6 for P and $P1$ populations, and to 20.5 for $P2$ (see Figures 18 and 19, as well as Table 3).

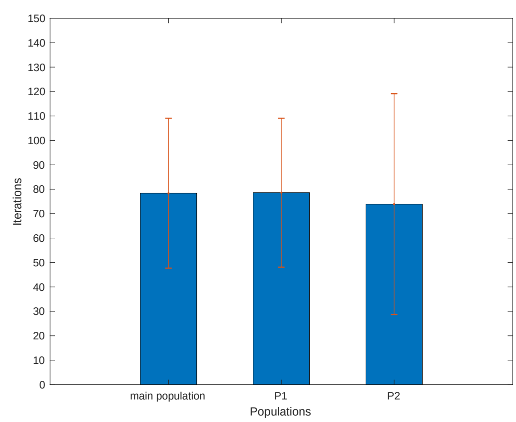

Figure 18: The figure displays the mean and the standard deviation of the average of iteration of matching to find the target object for the robot population over the 106 test images.

Table 3: Results of testing on 106 images.

Learning	$P/T1$	$P1/T2$	$P2/T2$
ϵ_1	0.3	0.60	0.60
Mean of iterations	78.4	78.6	73.9
Deviation	30.7	30.5	45.2
Mean of min of iterations	46.6	46.6	20.5

After taking a closer look at our two populations' robots' answers $P1$ and $P2$, we realized that both populations did not find the goal in 30 test images. For the remaining 76 test images at least one of the two populations

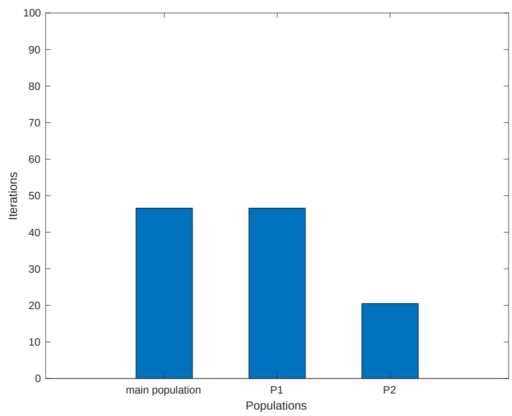

Figure 19: The figure displays the mean of the minimum of iteration of matching to find the target object for the robot population over the 106 test images (the lowest being the best).

Table 4: Detailed results of testing on 106 images.

Learning	$P/T1$	$P1/T2$	$P2/T2$
Success	66	66	76
Failes	40	40	30

had a right answer (see Table 4). Averaging the variances of all images as above doesn't precisely reflect the population's diversity. To obtain a more precise evaluation of our results we used another way to compare the populations' diversity and their capability of an accidental discovery. We calculated the mean absolute deviation (MAD) (equation in Appendix) of robots' answers for each population on each test image. The MAD is a way to describe variation in a data set and helps us to get a sense of how "spread out" the values are (the dispersion of a population individuals' answers in our case). By comparing the MAD of the two populations, we can find out if the museum learning increases the population diversity. We write MAD_{P1} and MAD_{P2} to refer to the mean absolute deviation of the population $P1$ and the population $P2$ respectively. The result shows that the two populations had exactly similar responses for 38 images (the robots belonging to $P1$ answers exactly the same as the robots belonging to $P2$ for a test image). The MAD_{P1} is equal to MAD_{P2} for all 38 images.

For the last 38 test images, the museum population $P2$ has a largely superior diversity ($MAD_{P2} > MAD_{P1}$) in 24 images. We should mention here that in ten images, at least one individual of $P2$ found the goal whereas none of the individuals of $P1$ found the goal. The opposite scenario does not exist in our results. The population $P1$ (laboratory populations) has a very slightly higher diversity for 4 images ($MAD_{P1} > MAD_{P2}$) and for the last 10 test images, the diversity is the same ($MAD_{P1} = MAD_{P2}$) for both populations.

The number of test images where a population has a higher diversity is considered as a frequency data. We chose to use the two sample proportions test to evaluate the significance of our populations' difference.

We used 'significant' to mean statistically significant at the $P = 0.05$ level. The number of trials is $N = 76$ (the number of test images when at least one population has an answer). We used Fisher's exact method to calculate the p-values as we have only four cases when $P1$ has a larger diversity than $P2$ (the number of events is less than 5 in our sample). The test indicates that our two populations have a statistically significant difference ($P < 0.0001$). The results showed a clear difference in diversity between the two populations $P1$ (learned only in the laboratory) and $P2$ (learned in both laboratory and museum). The cumulative learning at the museum played a role on robots' differentiation (individuation) and significantly increased their diversity. Not only does $P2$ show a higher diversity than $P1$, but the fact that one of its individuals found the goal when the whole of the $P1$ population did not find it indicates the higher probability of an accidental discovery for the $P2$ population (in our results the high diversity of $P2$ increases the accidental discovery chances).

6.4 Formal analysis

In a formal framework, we suppose that the robots have first learned to associate the patterns from the set $A = \{A_1, A_2, \dots, A_i, \dots, A_N\}$ to different vectors V_i describing the value of each pattern $V = (v_1, v_2, \dots, v_T)$. Here, N is the number of patterns (supposed to be large) and T is the number of emotional states and $v_p \in \mathbb{R}$. Hence, the learning of the robot k allows the building of a function f_k such as $A_i \overset{f_k}{\mapsto} V_i = f_k(A_i)$. Next, in our experimental framework, some robots are put in another environment (in our case the museum, population $P2$) and learn to associate a small database $B = \{B_1, B_2, \dots, B_M\}$ with $M << N$ to specific V vectors. After learning, the function f_k becomes a new function $f_k^{'}$. To be able to compare those robots with the ones having learned only in the laboratory environment ($P1$ population), the learning is continued for the same number of time steps with images coming from the laboratory environment. This second phase of learning in the laboratory explains the previous results where $P1$ and P have a slightly higher diversity for 4 images. Those 4 images present a higher similarity to the laboratory learning database.

If the dataset B is different enough [4]. from the dataset A, we can suppose that the new learning in the museum is not affecting the first learning in the laboratory. This is because the new patterns are projected onto orthogonal components to the first learned patterns. This condition is satisfied in our experiment where the recognition rate of laboratory objects changed less than 1% before and after museum learning. Of course, this is only true in a first approximation since an image is not recognized as a single input but is considered as a collection of local views voting for a given output label. Hence, some local views can be common to two different patterns belonging respectively to A and B. These induce some changes in the output V if the desired output vectors are different. Yet, since the common local views are limited to a small part of the total number of views (otherwise a new shape should not be recruited), the effect of learning on the f_k function is limited to new patterns and the generalization cases (by definition, a new pattern cannot have a majority of views common to another pattern already present in the database). If the pattern A_i activates only the neuron i of the visual feature group (VF): $\mathrm{VF}_i = 1$ and $\mathrm{VF}_j = 0$ for all $j \neq i$ then only the weights connecting VF_i to OEV_k will be modified since the LMS rule (see Equation 3) modifies only the weights associated to none null inputs. Here, k is the item number of the associated Object Emotional Value.

At the end of learning, the robots having faced two episodes of learning in the laboratory environment will not change their answers from the first phase of learning. Only their weights will be reinforced leaving their outputs identical. There is no difference in answers between the main population P and $P1$. When these robots face the test examples, all of the robots will react in the same way since their weight matrices have converged onto the same solution. This is assuming the learning in the laboratory environment was long enough to ensure all of the robots learned all of the objects the same number of times (this is the reason for a choice of a small learning rate ϵ for the LMS). Hence, if the test examples C_1 to C_L are far away from A_1 to A_N then the robots' reaction will appear as random but identical for all the robots as shown in the experimental results. The number of correct choices over the test examples C_k will decrease as a function of the number L of examples. Yet, for the population P_2 having faced a second learning phase in the museum, the robots can have different reactions to a pattern C_k if, during the museum experiment, the nearest B_j pattern verifies

$$Min||B_j - C_k|| < Min||A_i - C_k||$$

for all $(i, j) \in N \times M$ and has received different output values from the visitors. For the L test examples, if Q examples are nearer to B_j patterns than to A_i patterns, and that all these B_j patterns have received at least one positive and one negative value, then the probability that the population as a whole proposes one correct answer for the Q examples will be 1. The diversity of the population responses will guarantee that at least one robot will find the correct answer for each of the Q examples.

7 Conclusion and discussion

This study began with a presentation of our neural network model which enables Berenson to develop some kind of individual preferences thanks to its interactions with museum visitors. First, we introduced the evaluation of Berenson's visual system in the museum and its ability to recognize the previously learned objects. The results showed the importance of taking into account two types of visual information, *What* and *Where* when processing an object recognition task. The results showed a false positive recognition rate when using only the *What* information. Indeed, the robot does answer correctly to the question "is the object present in the image", but selects a wrong object when it is asked to localize it. This may happen when there are several learned objects in the same image because sometimes the robot overgeneralizes on other objects. We consider the ability to generalize on unlearned objects as essential to generalize learned aesthetic preferences, but it is necessary to have a compromise between generalization and discrimination to maintain a good performance on the object recognition task.

Secondly, we used the capability of Berenson to appreciate artworks to go further and studied the impacts of cumulative learning in special environments such as museums on a whole population of robots. Online learning in the museum environment offers categorical diversity thanks to the real-time interaction with different visitors having distinguished preferences. Such an experiment increases the diversity of choices for the members of a population in a new environment and therefore plays a role in the process of individuation. We mention here that Berenson's way of navigating in a museum was influenced by its previous learning. We can imagine the individuals of a population visiting a place such as a museum. They

4 i.e. for all $(i, j) \in N \times M$, we have $||A_i - B_j|| > \gamma$ with γ a threshold representing the vigilance level used for recruiting new shapes and $||.||$ a norm or distance measure (in our case a Manhattan distance to simplify the computations).

will have different directions toward their preferred objects and they will make different expressions in front of objects. Moreover, the diversity of individual reactions maximizes the chance of success of the whole population when facing new objects and new problems thanks to their different reactions.

The simulation tests indicated the difference in diversity between the two tested populations. According to the learning, the population $P2$ (which attended a second learning phase in the museum) shows a higher diversity than $P1$ (which attended a second learning phase in the laboratory) when facing new objects in new environments. When we repeated the test for a higher number of test images, the difference between the two populations' answers was clearer. Not only was there a difference in the average of iterations of matching between the two populations, but furthermore the population which had learned only in the laboratory didn't find the winner object whereas for the same images of test, the probability that one robot in the museum population found it was high. Comparing the mean absolute deviations of the robots' answers shows a significant difference between the two populations, highlighting the effect of the second learning at the museum on the population diversity. This representative result considers a small population of individuals (robots). Hence, the performances of a robot population visiting the museum (taking or avoiding an unknown object in an unknown visual scene) will be higher if the museum contains a high diversity of artworks associated with a high diversity of appreciations. The fact of getting a high diversity of answers after a museum experience is an important feature to be used in integrating the intelligent systems in scientific reasoning. Such diversity could increase the chances in cases of accidental discovery. However, the nature of scientific discoveries is debated as to whether it is an orderly march, or a kind of random stroll. From a lot of arguments supporting each point of view, we like to highlight that only the persons whose mind is prepared to see things will actually notice them. Then a scientific discovery is unlikely to be purely random. Our point here is to indicate that the museum visit increases the diversity thanks to the learning variability that could increase the chance of accidental discovery without passing by an absolutely random approach.

A Appendix

In Equation 1:

$$net_j = 1 - \frac{1}{N}\sum_{i=1}^{N} |W_{ij} - I_i|$$

$$H_\theta(x) = \begin{cases} 1 & \text{if } \theta < x \\ 0 & \text{otherwise} \end{cases}$$

Here, N is the local view size, I_i is the input visual feature, and the learned features are coded on the neuron's weight W_{ij}. The $H_\theta(x)$ is the Heaviside function.

The facial expression learning is based on the following LMS equation:

$$\Delta W_{ij}^{VF-ISP} = \epsilon_1 \cdot VF_i \cdot (IS_j - ISP_j)$$

The *Where* channel treats the estimation of the object's positions also using the LMS rule. The θ_L group associates the predicted *Where* information with each local view VF in order to compute in θ_S the artwork's shifted position during the use mode. Then, based on θ_S, the angular command is given to the robot to head toward a recognized object. The next equation shows the weight modification for each iteration in learning mode (the time variable t is not represented),

$$\Delta W_{ij}^{VF-\theta_L} = \epsilon_1 VF_i(\theta_j - \theta_{Lj})$$

In use mode, when a local view is recognized, the neuron coding for its learned position is activated. Then θ_S computes the distance between the learned and the current position, as formalized in the next equation. When the local view is at the learning position then $\theta_S = 0$. If the local view is translated by a distance Δd, then $\theta = \theta_L + \Delta d$ and the neuron Δd is activated in θ_S. Here, θ_S represents the vector θ circularly shifted,

$$\theta_S(Circ(x - \arg\max(\theta_L))) = \theta(x)$$

$$Circ(x) = \begin{cases} x, x > 0 \\ x + N, x < 0 \end{cases}$$

Now, if we assume that the learned object was centered in the camera field, the referential becomes the object center (in the x abscissa). The local views belonging to this object predict in θ_S their distance to the object center. The system can estimate the object pose by integrating the activity of the neurons in θ_S,

$$PI(x) = \sum_{t=1}^{\tau} \frac{1}{2\pi\sigma_1^2} e^{\frac{-(x_t - ds_t)^2}{2\sigma_1^2}}$$

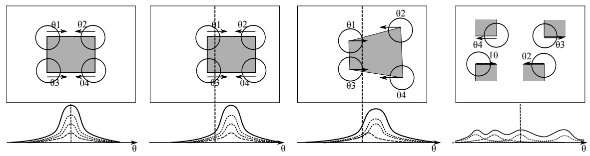

Figure 20: Schematic example of pose estimation applied to a square. (1) is the learned object at the image center. Below is the activity in the *PI* group. It creates a peak at the object location. (2) shows the same square translated and the translation result in the *PI* group. (3) is the same square with deformation, and its estimated pose below. (4) contains same local view as in (1) scattered in the image.

$$ds_t = (Circ(d_t - \arg\max(\theta_L)))$$

In the two previous equations, ds_t is the distance of each local view to the referential (object center), and the *Position Integration* (*PI*) group integrates the local view distances to the referential with a Gaussian kernel summation (more details in [35]). In the learning phase when the object is in the center of the field of view, the sum of the activity creates a peak at the image center. When the object is translated, the peak is also translated in the image referential. The activity in the *PI* group may represent the confidence level the system has in the object's estimated pose and even in the object recognition (see Figure 20).

The following equation describes the calculation of the Mean Absolute Deviation (MAD). The MAD of a data set is the average distance between each data value and the mean:

$$MAD = \frac{\sum |(x_i - \overline{x})|}{n}$$

Table 5: The iterations of matching of all robots in $P2$ for some test images revealing their different preferences.

	I1	I2	I3	I4	I5	I..	Averge
$R1$	10	27	12	41	200	..	66,8
R2	48	27	12	41	200	..	74,4
R3	48	200	12	41	200	..	73,6
R4	48	27	12	41	200	..	74,4
R5	10	55	12	200	10	..	106
R6	42	200	12	41	200	..	121,7
R7	10	27	12	41	200	..	80,2
R8	48	27	12	200	200	..	84,9
SD	19,1	78,5	0	73,6	67,1	..	56,9

Acknowledgements: The authors thank the anonymous reviewers for their valuable suggestions and points for discussion. Their input helped us to significantly extend and improve the paper. The authors thank the Quai Branly museum staff, especially A. C. Taylor and C. Chenu for hosting us. The mediators for their help in animating the experiment. J. Nadel, M. Simon, R. Soussignan and P. Canet for the design and the calibration of the robot head. And the Equipex Robotex. This study was supported by the Labex Patrima (foundation for French patrimony promotion). The authors also thank A.Blanchard, F. De Melo, A. Pitti, A. Melnyk and J. Becker.

References

[1] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, et al., Cognitive developmental robotics: A survey, IEEE Transactions on Autonomous Mental Development, 2009, 1(1), 12–34

[2] M. Lungarella, G. Metta, R. Pfeifer, G. Sandini, Developmental robotics: a survey, Connection Science, 2003, 15(4), 151–190

[3] A. G. Barto, S. Mahadevan, Recent advances in hierarchical reinforcement learning, Discrete Event Dynamic Systems, 2003, 13(1-2), 41–77

[4] A. Karaouzene, P. Gaussier, D. Vidal, A robot to study the development of artwork appreciation through social interactions, In: 2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL), IEEE, 2013, 1–7

[5] M. Klinnert, J. Campos, J. Sorce, R. Emde, M. Svejda, The development of the social referencing in infancy, Emotion in early development, 1983, 2, 57–86

[6] M. D. Klinnert, R. N. Emde, P. Butterfield, J. J. Campos, Social referencing: The infant's use of emotional signals from a friendly adult with mother present, Developmental Psychology, 1986, 22(4), 427

[7] S. Boucenna, P. Gaussier, P. Andry, L. Hafemeister, Imitation as a communication tool for online facial expression learning and recognition, In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2010, 5323–5328

[8] C. Hasson, P. Gaussier, S. Boucenna, Emotions as a dynamical system: the interplay between the meta-control and communication function of emotions, Paladyn, Journal of Behavioral Robotics, 2011, 2(3), 111–125

[9] R. A. Schmidt, R. A. Bjork, New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training, Psychological science, 1992, 3(4), 207–218

[10] M. A. Schilling, P. Vidal, R. E. Ployhart, A. Marangoni, Learning by doing something else: Variation, relatedness, and the learning curve, Management Science, 2003, 49(1), 39–56

[11] G. Wulf, The effect of type of practice on motor learning in children, Applied Cognitive Psychology, 1991, 5(2), 123–134

[12] A. F. Healy, E. L. Wohldmann, E. M. Sutton, L. E. Bourne Jr, Specificity effects in training and transfer of speeded responses, Journal of Experimental Psychology: Learning, Memory, and Cognition, 2006, 32(3), 534–546

[13] C. Gonzalez, P. Madhavan, Diversity during training enhances detection of novel stimuli, Journal of Cognitive Psychology, 2011, 23(3), 342–350
[14] K. Popper, The logic of scientific discovery, Routledge, 2005
[15] P. Langley, H. A. Simon, G. L. Bradshaw, J. M. Zytkow, Scientific discovery: Computational explorations of the creative processes, MIT press, 1987
[16] B. C. Falkenhainer, R. S. Michalski, Integrating quantitative and qualitative discovery: the abacus system, Machine Learning, 1986, 1(4), 367–401
[17] J. M. Zytkow, Automated discovery of empirical laws, Fundamenta Informaticae, 1996, 27(2-3), 299–318
[18] B. Nordhausen, P. Langley, A robust approach to numeric discovery, In: Machine Learning Proceedings 1990, Elsevier, 1990, 411–418
[19] A. Sparkes, W. Aubrey, E. Byrne, A. Clare, M. N. Khan, M. Liakata, et al., Towards robot scientists for autonomous scientific discovery, Automated Experimentation, 2010, 2(1), 1
[20] H. Prendinger, M. Ishizuka, Methodological considerations on chance discovery, In: 26th Annual Conference of the IEEE, Industrial Electronics Society, IECON 2000, IEEE, 2000, 1652–1655
[21] W. Tung, C. Quek, A neurocognitive approach to decision-making in chance discovery, In: Chance discoveries in real world decision making, Springer, 2006, 231–250
[22] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, et al., The interactive museum tour-guide robot, In: AAAI '98/IAAI '98 Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence, 1998, 11–18
[23] I. Macaluso, A. Chella, Machine consciousness in cecerobot, a museum guide robot, In: Proceedings, AAAI Fall 2007 Symposium, Arlington VA, 2007
[24] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, et al., Probabilistic algorithms and the interactive museum tour-guide robot minerva, The International Journal of Robotics Research, 2000, 19(11), 972–999
[25] T. Ogata, S. Sugano, Emotional communication between humans and the autonomous robot which has the emotion model, Sensors, 1999, 10(3)
[26] S. Marsland, U. Nehmzow, J. Shapiro, On-line novelty detection for autonomous mobile robots, Robotics and Autonomous Systems, 2005, 51(2), 191–206
[27] M. Csikszentmihalyi, Society, culture, and person: A systems view of creativity, In: R. J. Sternberg (Ed.), The Nature of Creativity: Contemporary Psychological Perspectives, 1988
[28] P. Gaussier, S. Zrehen, Perac: A neural architecture to control artificial animals, Robotics and Autonomous Systems, 1995, 16(2-4), 291–320
[29] M. A. Goodale, A. D. Milner, Separate visual pathways for perception and action, Trends in neurosciences, 1992, 15(1), 20–25
[30] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, An efficient k-means clustering algorithm: Analysis and implementation, IEEE Transactions of Pattern Analysis and Machine Intelligence, 2002, 24, 881–892
[31] R. A. Rescorla, A. R. Wagner, et al., A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, Classical conditioning II: Current research and theory, 1972, 2, 64–99
[32] B. Widrow, M. E. Hoff, Adaptive switching circuits, In: IRE WESCON, New York: Convention Record, 1960, 96–104
[33] S. Boucenna, P. Gaussier, P. Andry, L. Hafemeister, A robot learns the facial expressions recognition and face/nonface discrimination through an imitation game, International Journal of Social Robotics, 2014, 6(4), 633–652
[34] S. Boucenna, P. Gaussier, L. Hafemeister, Development of first social referencing skills: Emotional interaction as a way to regulate robot behavior, IEEE Transactions on Autonomous Mental Development, 2014, 6(1), 42–55
[35] A. Moualla, A. Karaouzene, S. Boucenna, D. Vidal, P. Gaussier, Readability of the gaze and expressions of a robot museum visitor: impact of the low level sensory-motor control, In: 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2017), IEEE, 2017, 712–719