

Citation for published version:

John Lones, Matthew Lewis, and Lola Cañamero, 'From Sensorimotor Experiences to Cognitive Development: Investigating the Influence of Experiential Diversity on the Development of an Epigenetic Robot', *Frontiers in Robotics and AI*, Vol. 3 (44), August 2016.

DOI:

<https://doi.org/10.3389/frobt.2016.00044>

Document Version:

This is the Published Version.

Copyright and Reuse:

© 2016 The Author(s).

This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Enquiries

If you believe this document infringes copyright, please contact Research & Scholarly Communications at rsc@herts.ac.uk



From Sensorimotor Experiences to Cognitive Development: Investigating the Influence of Experiential Diversity on the Development of an Epigenetic Robot

John Lones*, Matthew Lewis and Lola Cañamero

Embodied Emotion, Cognition and (Inter-)Action Lab, School of Computer Science, University of Hertfordshire, Hatfield, UK

OPEN ACCESS

Edited by:

Guido Schillaci,
Humboldt University,
of Berlin, Germany

Reviewed by:

Matthias Rolf,
Oxford Brookes University, UK
James Law,
The University of Sheffield, UK

*Correspondence:

John Lones
j.lones@herts.ac.uk

Specialty section:

This article was submitted to
Humanoid Robotics,
a section of the journal
Frontiers in Robotics and AI

Received: 10 October 2015

Accepted: 04 July 2016

Published: 26 August 2016

Citation:

Lones J, Lewis M and Cañamero L
(2016) From Sensorimotor
Experiences to Cognitive
Development: Investigating
the Influence of Experiential
Diversity on the Development
of an Epigenetic Robot.
Front. Robot. AI 3:44.
doi: 10.3389/frobt.2016.00044

Using an epigenetic model, in this paper we investigate the importance of sensorimotor experiences and environmental conditions in the emergence of more advanced cognitive abilities in an autonomous robot. We let the robot develop in three environments affording very different (physical and social) sensorimotor experiences: a “normal,” standard environment, with reasonable opportunities for stimulation, a “novel” environment that offers many novel experiences, and a “sensory deprived” environment where the robot has very limited chances to interact. We then (a) assess how these different experiences influence and change the robot’s ongoing development and behavior; (b) compare the said development to the different sensorimotor stages that infants go through; and (c) finally, after each “baby” robot has had time to develop in its environment, we recreate and assess its cognitive abilities using different well-known tests used in developmental psychology such as the violation of expectation (VOE) paradigm. Although our model was not explicitly designed following Piaget’s or any other developmental theory, we observed, and discuss in the paper, that relevant sensorimotor experiences, or the lack of, result in the robot going through unforeseen developmental “stages” bearing some similarities to infant development, and could be interpreted in terms of Piaget’s theory.

Keywords: epigenetic development, developmental robotics, sensorimotor development, cognitive development, social robotics, affective adaptation, human–robot interaction, autonomous robot

1. INTRODUCTION

The first 2 years of life represent a period of rapid cognitive development in human infants. During this 2-year period, behavioral patterns shift from simple reactions to incorporating the use of symbols in mental representations, setting the stage for further cognitive development (Piaget, 1952; DeLoache, 2000). While this cognitive development process is still not fully understood, some evidence does suggest that early stimulation provides the foundation for this process (Piaget, 1952; Fischer, 1980; Fuster, 2002; Bahrick et al., 2004). This theory of development is perhaps best explained in Piaget’s concept of sensorimotor development. According to Piaget’s work (Piaget, 1952), during the sensorimotor period infants go through 6 substages of development, which we will briefly lay out.

The first stage, often simply referred to as the “reflex stage,” lasts from birth until around 1 month of age, with infants limited to simple automatic “innate” behaviors (Piaget and Inhelder, 1969). The second stage, known as “primary circular reactions,” occurs approximately between 1 and 4 months and sees the infant’s behavior begin to incorporate, repeat, and refine reflex behaviors focused on their own bodies. In the third stage, which takes place between the 4th and 8th month, infants start to notice that their actions can have interesting effects on their immediate environment. By around the 8th to 12th month, infants begin to display coordination in their secondary circular reactions facilitating goal-directed behavior. In addition, during this stage, infants begin to show signs of understanding the concept of object permanence (Baird et al., 2002). Between 12 and 18 months, children’s behavior starts to incorporate tertiary circular reactions, where they will now both take greater interest in and experiment with novel objects. Finally, before the end of the 24th month, it is expected that children would have developed some form of mental representation and symbolic thought, whereby they will now engage in both imitation and make-believe behaviors.

In Piaget’s theory, these different substages represent progressive incremental steps in the cognitive development of the infant (Piaget, 1952). However, it should be noted that, in other works, some of the different cognitive developments, such as the understanding of object permanence, have been found to occur in different life stages (Baillargeon et al., 1985; Kagan and Herschkowitz, 2006) and that, even within the Piagetian tradition, the very notion of progressive developmental stages has been questioned in favor of other aspects of development, such as domain specificity (Karmiloff-Smith, 1992). This would suggest that the complexity of human development implies that either the different milestones overlap or are not necessarily entirely dependent on their “predecessor,” with environmental conditions also constituting an important factor influencing the outcome and process of development (Baillargeon, 1993). Alternatively, it cannot be ruled out that flaws or oversights in experimental models may have led to different outcomes (Munakata, 2000). In any case, the milestones put forward by Piaget do seem to represent critical cognitive developments which are likely to set the foundation and facilitate the emergence of more advanced functions (Piaget, 1952; Baird et al., 2002; Fuster, 2002; Bahrack et al., 2004). It is likely then that, if the abilities gained during these stages are indeed important in the development of human cognition, then they would also be significant in the development of adaptive autonomous robots undergoing “similar” sensorimotor experiences as those related to the development of those skills (Asada et al., 2009; Cangelosi et al., 2015). This potential has led to the design of a range of different models using sensorimotor developmental principles.

In this paper, we examine the role that exposure to sensory stimuli may have on the cognitive development of an autonomous robot. Unlike related studies such as (Shaw et al., 2012; Ugur et al., 2015), which explicitly model the developmental process, the model used in our experiments was not designed following a particular sensorimotor developmental theory,

but based on a plausible epigenetic¹ mechanism (Lones and Cañamero, 2013; Lones et al., 2014). However, similar to the work of Cangelosi et al. (2015) and Ugur et al. (2015), our model leads to the emergence of an open-ended learning process achieved by allowing a robot to be able to identify and learn about interesting phenomena, a common goal of developmental models (Marshall et al., 2004). The underlying approach that our model takes to achieve the desired open-ended development also has similarities to other developmental models. Here, we use a novelty-driven approach regulated by intrinsic motivation (see Section 2.2.3). Using novelty as a way to regulate interactions with the environment and drive development has been previously explored for example by Blanchard and Cañamero (2006) and Oudeyer and Smith (2014). While there are significant differences in the way in which curiosity is generated by these models (here through hormone modulation with regard to the robot’s internal and external environment), similar to Oudeyer and Smith (2014), we use the concept of curiosity to model the robot’s novelty-seeking behavior by encouraging it to reduce uncertainty in an appropriate manner given its current internal state. This mechanism drives the robot’s interactions, permitting it to learn and develop in an appropriate manner as it is exposed to different sensorimotor stimuli as a result of its interactions.

Using this model in an autonomous robot, we observed a natural and unforeseen developmental process somewhat similar to the sensorimotor development suggested by Piaget (1952), as we will present in this paper. As we will see, the robot’s progression through, as well as the emergence of, the behaviors associated with the different developmental stages, depend on the robot’s environment, and more specifically on the sensory stimuli that the robot is exposed to over the course of its development. For example, a robot placed in an environment deprived of sensory stimulation did not develop behaviors or abilities associated with the sensorimotor developmental theory. By contrast, a robot given free range in a novel environment, showed the emergence of different stages and abilities associated with the developmental process, i.e., primary and later secondary circular reactions as a consequence of its interactions in the environment. In our model, this would suggest that the emergence of these stages is related to exposure of the robot to environmental stimuli. More importantly, this leads to the research question of whether the emergence of these similar processes and stages would have any consequence for the cognitive development of the robot, or whether the similarities are simply related to a temporal phenomenon. In order to investigate this question, we allowed three robots to develop under different environmental conditions, two of which provided different levels of novelty and sensory stimulation, and the final was equivalent to sensory deprivation. We tested the robots’ cognitive abilities in a range of scenarios ranging from a simple learning task to a more specific violation of expectation paradigm.

¹The term “epigenetic” is used here to refer to mechanisms that lead to changes in gene expression (Schlichting and Pigliucci, 1998; Holliday, 2006; Lones and Cañamero, 2013) rather than the Piagetian notion, widely adopted in the field of developmental robotics, referring to developmental processes not directly stemming from the action of genes (Cangelosi and Riga, 2006).

The model used for these tests is based on an earlier hormone-driven epigenetic model presented in (Lones and Cañamero, 2013). This previous model showed the ability to both rapidly adapt to a range of different dynamic environmental conditions and react appropriately to unexpected stimuli. However, this early epigenetic model was based on reactive behaviors and lacked cognitive development, limiting the ability of the robot to produce and engage in planned behaviors, or to explicitly learn about new aspects of its surroundings; thus, the robot was partly dependent on information about the environment “pre-coded” in its architecture. In an attempt to overcome these limitations, we integrated in the robot’s architecture a new form of neural network that we have since called an “Emergent Neural Network” (ENN), in which nodes and synaptic connections between them are created as the robot is exposed to stimulation. This network should therefore allow the robot to learn about different aspects of the environment with regard to the affordances they provide (Lones et al., 2014).

This extended model uses the same hormonal system as we used previously in (Lones and Cañamero, 2013), this time to modulate the development of the ENN in an “appropriate” manner dependent on the interaction with the external environment. The proposed model not only allowed the robot to learn about different aspects of its environment and engage in planned behavior in an appropriate manner but also gave it the ability to shape its body representation due to its interaction with the external environment. This allowed the robot to adapt successfully to a range of simple, but “real-world” and dynamic environments such as our office environment (Lones et al., 2014). However, the ENN used in (Lones et al., 2014) while successful, had constraints imposed on it due to the focus of that particular study, where the interest lay in the roles that homeostasis and hormones played on modulating curiosity and novelty-seeking behavior. Specifically, the network was explicitly pre-trained in a sterile environment and then frozen, removing the potential for additional learning. In contrast, here, where the interest lies in the roles that stimulation has on the robots’ “cognitive development,” the network did not undergo any pre-training and was fully active throughout. This means that the robot’s learning is dependent on its own sensorimotor experiences within its environment. As we will demonstrate, for this particular model, the quality of these sensorimotor experiences is paramount to the robots’ cognitive development, where a robot which has been exposed to rich sensorimotor experience develops not only greater cognitive abilities but also goes through developmental stages that we had not anticipated and which bear some similarities to infant development.

2. MATERIALS AND METHODS

2.1. Robot and Sensors

For the experiments reported in this paper, we used the medium-sized wheeled Koala II robot by K-Team. This robot is equipped with 14 infra-red (IR) sensors spread around the body, which are used to detect the distance, size, and shape of different objects. In lieu of traditional touch sensors, the IR sensors were also used to detect contact. In addition, to complement these sensors, a Microsoft LifeCam provided a microphone for sound

detection and, along with the OpenCV library, simple color-based vision. The robot’s architecture was written in C++, and control of the model was handled through a serial connection to a computer running Ubuntu.

2.2. Architecture of the Robot

The software architecture giving rise to the behavior of the robot combines three main elements: a number of survival-related homeostatically controlled variables that provide the robot with internal needs to generate behavior; a hormone-driven epigenetic mechanism that controls the development of the robot; and a novel neural network that we have named a “emergent neural network”, which provides the robot with learning capabilities. These three elements of the model interact in cycles or action loops of 62.5 ms in order to allow the robot to develop and adapt to its current environmental conditions as shown in **Figure 1**. This development occurs in the following manner:

- The levels of the survival-related homeostatically controlled variables change as a function of the actions and interactions with its environment (see Section 2.2.1).
- Deficits of the different homeostatic variables trigger the secretion of different artificial hormones (see Section 2.2.2).
- The different hormones, once secreted modulate both the epigenetic mechanism and emergent neural network (see Section 2.2.3).

2.2.1. Homeostatic Variables

Homeostatic imbalances have often been linked to the generation of drives and motivation in biological organisms, providing them with a potential short-term adaptation tool (Berridge, 2004). In a similar manner, in robotics, artificial homeostatic variables have been used as an effective way of generating motivations and behaviors which can be used for adaptive robotic controllers (Cañamero, 1997; Breazeal, 1998; Cañamero et al., 2002; Di Paolo, 2003; Cos et al., 2010). We have endowed our robot with three survival-related variables that it must maintain within appropriate ranges in order to survive: health, energy, and internal temperature. The three variables decrease as a function of the robot’s actions and interaction within its environment. Health is a simulated variable which decreases proportionally in relation to physical contact, as shown in formula 1:

$$\Delta Health = \begin{cases} -C & \text{if } C \geq 5 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where C is the intensity of any contact, and the value 5 represents a threshold/resistance to damage² that must be surpassed in order for health loss to occur. Health deficits can be recovered through the consumption of specific resources found in the environment.

Energy is linked directly to the robot’s battery and decreases at an average of around 15 mAh/min, although the exact amount

²Value of less than 5 are roughly equivalent to gentle strokes or minor contact resulting in no health loss.

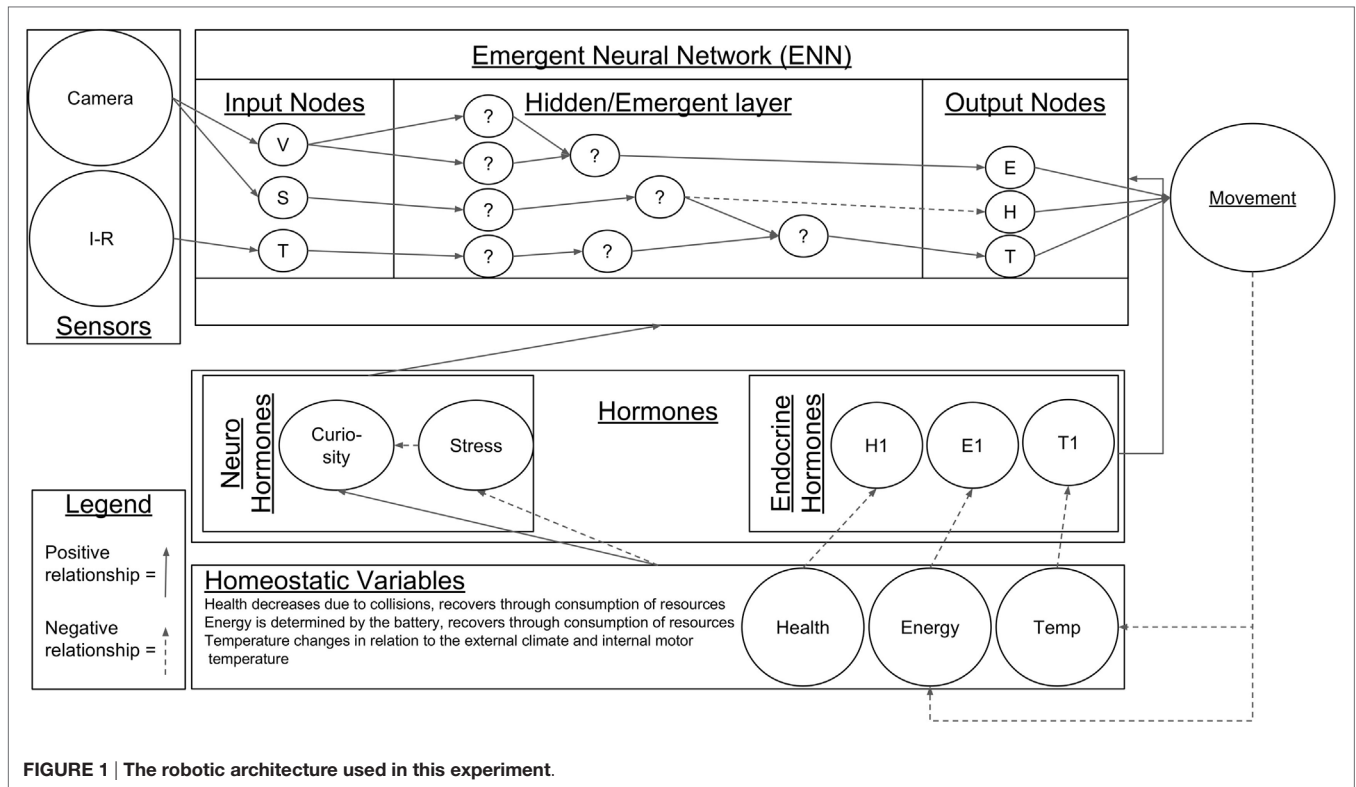


FIGURE 1 | The robotic architecture used in this experiment.

varies as a function of the robot’s motor usage. Although the total battery size is approximately 3500 mAh, the robot has been programmed to only sense a maximum charge of up to 75 mAh (around 5 min of running time) at any given time, creating a virtual battery. This allows us to implement a virtual charging system where the robot needs to find specific energy resources in order to recharge its virtual battery back to the maximum 75 mAh.

For both the energy and health resources, the robot recovers 7.5 and 10 U (10% of maximum capacity), respectively, per action loop when the resource is directly in front of the robot and within roughly 10 cm distance. Finally, the robot’s internal temperature is related to the speed of the motors and the climate, following equation (2):

$$\Delta Temperature = \frac{|speed|}{10} * Climate \quad (2)$$

where |speed| is the current absolute value of speed of the wheels (measured in rotations per action loop) and 10 is a predetermined constant to regulate the temperature gain with regard to movement. Climate refers to the external temperature, usually measured using a heat sensor; however, for these experiments, in order to remove unwanted variations, this was set to be detected as a constant of 24 [i.e., 24°C (75°F)].

The robot’s body temperature is set to dissipate at a constant rate of 5% of the total internal temperature per action loop. Dissipation is the only method available to the robot to reduce its body temperature, meaning that in order to cool down, the robot must either reduce or suppress movement.

Each of the survival-related homeostatic variables has a lethal boundary which, if transgressed, results in the robot’s death. In the case of energy and health, the lethal boundary is set at the bottom end of the range of permissible values (0), in the case of temperature the lethal boundary is at the upper end of the range (100).

2.2.2. Hormone-Driven Epigenetic Mechanism

While survival-related homeostatic imbalances, as previously mentioned, are often used to model motivation and drive behavior in autonomous robots, these imbalances alone may not be enough to ensure adaptive behavior in dynamic or complex environments (Avila-Garcia and Cañamero, 2005). A suggested, and so far successful, addition to the previous architecture, consists of integrating different hormone or endocrine systems into the model (Avila-Garcia and Cañamero, 2005; Timmis et al., 2010; Lones and Cañamero, 2013). These systems borrow from biological examples, where neuromodulatory systems have been shown to regulate behavior and allow rapid and appropriate responses to environmental events (Krichmar, 2008). These models have been used for a range of environments and conditions such as single-robot open field experiments (Krichmar, 2013) to multiple robot setups such as predator-prey scenarios (Avila-Garcia and Cañamero, 2005) to robotic foraging swarms (Timmis et al., 2010).

In our setup, we use a model consisting of five different artificial hormones, both “endocrine” (eH) and “neuro-hormones” (nH) to help the robot maintain the three previously discussed homeostatic variables. While both hormone groups share

common characteristics, they also present some significant differences. The first group of hormones (eH) is secreted by glands as a function of homeostatic deficits; each of the three homeostatic variables has an associated hormone. These hormones – E1 associated with energy, H1 associated with health, and T1 associated with internal temperature – are secreted as shown in equation (3):

$$eHSecretion_h = \psi_h Deficit_h \quad (3)$$

where $eHSecretion_h$ is the amount of hormone h secreted, $\psi_h > 0$ is a constant regulating the amount of hormone h secreted by the gland (it can be thought of as reflecting the gland's "activity level"), and $Deficit_h$ ($0 \leq Deficit_h < 100$) is the value of the relevant homeostatic variable's deficit.

These hormones play a key role, as the robot is unable to directly detect the values of own homeostatic deficits; rather, the concentrations of the different hormones are used to signal homeostatic deficits through modulation of the ENN as discussed in Section 2.2.3, leading to the generation of drives and motivations.

The second group of hormones (nH) consists of two hormones: curiosity (nH_c) and stress (nH_s). These two hormones, which are secreted in relation to internal and external stimuli, are loosely based on the hormones Testosterone and Cortisol. nH_c will encourage outgoing behavior by increasing novelty seeking and suppressing detection of perceived negative stimuli. As an example, common behaviors linked to a high concentration of this hormone would be interest and interactions with novel objects. In contrast, nH_s will reduce novelty-seeking behavior and heighten the detection of any negative stimuli. An example of a behavior linked to a high concentration of this hormone would be the withdrawal to a perceived area of safety, which in our experimental setup often consisted of the edge of the environment – a wall – leading to the emergence of a sort of wall-following behavior. The robot's perception of walls providing safety arises here due to their perceived lack of novelty and they offer protection to one side of the robot. For the full implementation of the nH hormone group, please see Lones et al. (2014), we briefly summarize the behaviors of the hormones in equations (4) and (5).

$$nHSecretion_c = \frac{pS + \sum_v r_v}{nH_s} \quad (4)$$

where pS is the sum of all perceived "positive" stimuli, $r_v \geq 0$ is the (perceived) recovery of a homeostatic variable v during the current action loop, and nH_s is the concentration of the stress hormone which suppresses the secretion of nH_c . By "positive stimuli," we refer to the stimulation associated (by the robot's neural network) with the recovery (i.e., the correction of the deficit) of a homeostatic variable. In other words, positive stimulation pS is the sum of any output associated with the recovery of a homeostatic variable and is calculated by the synaptic function of the output nodes of the neural network, as shown in equation (12).

$$nHSecretion_s = roD \times oS \times nS \quad (5)$$

where roD or the "perceived risk of death" is the sum of all homeostatic deficits, oS or "overall stimulation" is the sum of the total

amount of stimulation (regardless of its type), and nS is the sum of perceived "negative" stimuli. By "negative stimuli," we refer to the stimulation associated (by the robot's neural network) with the worsening (i.e., the increase of the deficit) of a homeostatic variable. In other words, negative stimulation nS is the sum of any output associated with the worsening of a homeostatic variable and is calculated by the synaptic function of the output nodes of the neural network, as shown in equation (12). The overall stimulation oS is also determined by the synaptic function of the output nodes of the network and is the sum of the total synaptic output.

Once secreted, all hormones decay at a constant rate shown in equation (6).

$$hC_{h,t+1} = 0.95 \times hC_{h,t} \quad (6)$$

where $hC_{h,t+1}$ is the hormone concentration in the next action loop.

The second aspect of the hormonal system is the inclusion of 5 hormone receptors, each one associated with a specific hormone. These receptors are part of the ENN and detect the current concentration of their relevant hormone (see Sections 2.3.2 and 2.3.3). The sensitivity of these hormone receptors is not constant; rather, it is modulated by a hormone-driven epigenetic mechanism (Crews, 2008; Zhang and Ho, 2011) that we introduced in Lones and Cañamero (2013). This epigenetic mechanism consists of a feedback loop where the concentrations of the different hormones will lead to either upregulation (increased sensitivity) or downregulation (decreased sensitivity) of their respective receptors, following equation (7):

$$senS_{h,t+1} = senS_{h,t} \times \frac{hC_h}{\sigma} \quad (7)$$

where $senS_{h,t+1}$ is the hormone receptor's sensitivity in the next action loop, hC_h is the relevant hormone concentration, and σ a constant value that regulates the speed of the epigenetic change.

2.2.3. Emergent Neural Network

The emergent neural network consists of a novel design in which nodes are created as a function of the robot's interactions and exposures to different environmental stimuli. This emergent neural network, of which an example can be seen in **Figure 2**, is designed to allow the robot to learn the affordance of different aspects of its environment. Here, the term "affordance" is used in the context of the robot learning the potentialities of an action or interaction with different aspects of its environment in relation to its current internal state. Since the internal state of the robot presented here is dependent on and made up of the three homeostatic variables (see Section 2.2.1), the affordances learned by the robot will be in relation to the ability of actions to affect these said variables. For example, a potential action involving the energy resource will likely have an affordance associated with energy recovery. At this stage, it is important to highlight two aspects of the robotic model:

- First, all behaviors discussed here emerge as a result of the development and modulation of the neural network, simply put there are no pre-designed behaviors or internal states.

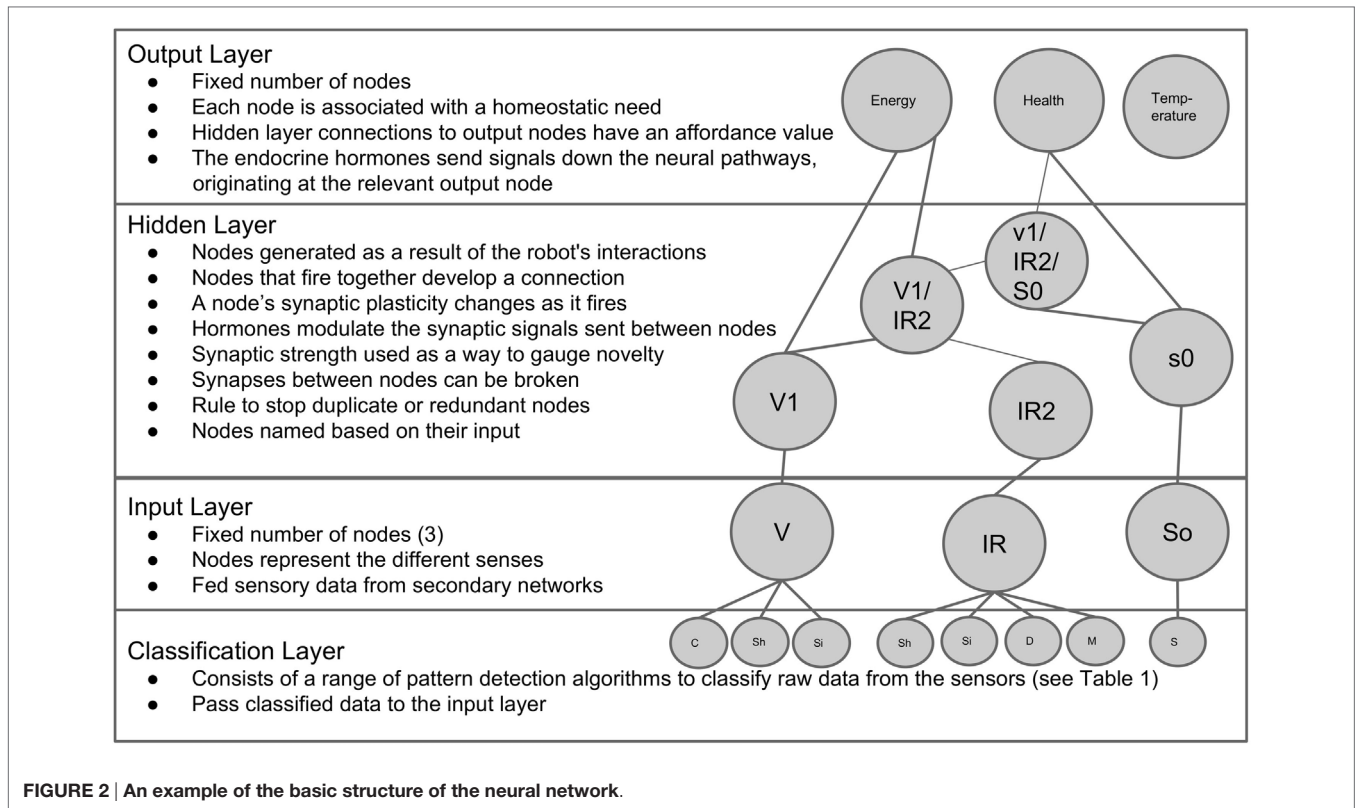


FIGURE 2 | An example of the basic structure of the neural network.

- Second, the development of both the neural network and affordances are based on the robot's perceptions and interactions, therefore robots with different morphological designs or placed in different environmental conditions will likely develop in different ways.

2.3. Detailed Description of the Emergent Neural Network

The emergent neural network used in this paper consists of a three-layer network design shown in **Figure 2**. The first layer of the network consists of an input layer which is fed sensory data from a range of different classification networks. The second layer is the hidden layer in which nodes emerge as a function of the robot's interactions and environmental exposures. This layer is responsible for recognizing different aspects of the environment and assigning an appropriate affordance based on the robot's past experiences. The final layer is the output layer which simply sums the detected affordances.

2.3.1. Classification and Input Nodes

The input layer consists of three fixed nodes, each representing one of the robot's different sensory modalities. These modalities are vision, IR, and sound and receive input data from different pre-processing classification algorithms shown in **Table 1** and **Figure 3**. These three *input nodes* are quite different to conventional neurons found in other networks. In our network, these nodes will fire differently depending on which classification

network is currently feeding input, with each node in the input layer associated with a specific fixed group of classification networks (see **Table 1**). For example, the node representing the vision modality is associated with classification networks that detect color, shape, and size. These input layer nodes work as follows.

For each sensory modality, the output from each of the pre-processing classification networks (shown in **Table 1**) consists of a 4-digit input pattern that feeds into the appropriate node in the input layer. The four digits provide information about the sensory modality used, the type of stimulus, the position of the stimulus with respect to the body coordinates of the robot, and the number of instances of that stimulus detected in that action loop. The number of pre-processing classification networks associated with each node of the input layer depends on the modality of the latter – three for vision, four for IR, and one for sound (see **Table 1**). For each input pattern received, each node in the input layer will either strengthen its connection with a node in the hidden layer corresponding to that input pattern if a node has already been associated with it, or create a new node if the pattern is classified as novel. In any one time frame, a node in the input layer can receive multiple inputs from each pre-processing classification network, and thus it can potentially create multiple new nodes in the hidden layer.

As an example, when perceiving the face depicted in **Figure 4**, the vision node in the input layer would receive an input from the shape pre-processing classifier consisting of the four digits 1 (indicating the vision modality), 0 (representing a circle),

TABLE 1 | The different sensory modalities and their implementation and design.

| | Vision | | | IR | | | | Sound | |
|-----------|---|--|---|---|---|---|---|---|--|
| Sense | Color | Shape | Size | Shape | Size | Distance | Movement | Volume | |
| Sensor | Webcam | Webcam | Webcam | IR | IR | IR | IR | Webcam | |
| Algorithm | OpenCV HSV | OpenCV contour | OpenCV contour | Pattern detection | Pattern detection | IR value | Compare IR value | Sound | |
| Values | White = 0 Black = 1 Green = 2 Red = 3 Yellow = 4 Blue = 5 Unknown = 9 | Circle = 0 Square = 1 Rectangle = 2 Triangle = 3 Crescent = 4 Unknown = 9 | Small = 0 Medium = 1 Large = 2 Unknown = 9 | Flat = 0 Curved = 1 Corner = 2 Hole = 3 Unknown = 9 | Small = 0 Medium = 1 Large = 2 Unknown = 9 | Close = 0 Medium = 1 Far = 2 Unknown = 9 | None = 0 Small = 1 Medium = 2 Large = 4 Unknown = 9 | Silent = 0 Quiet = 1 Medium = 2 Loud = 3 | |

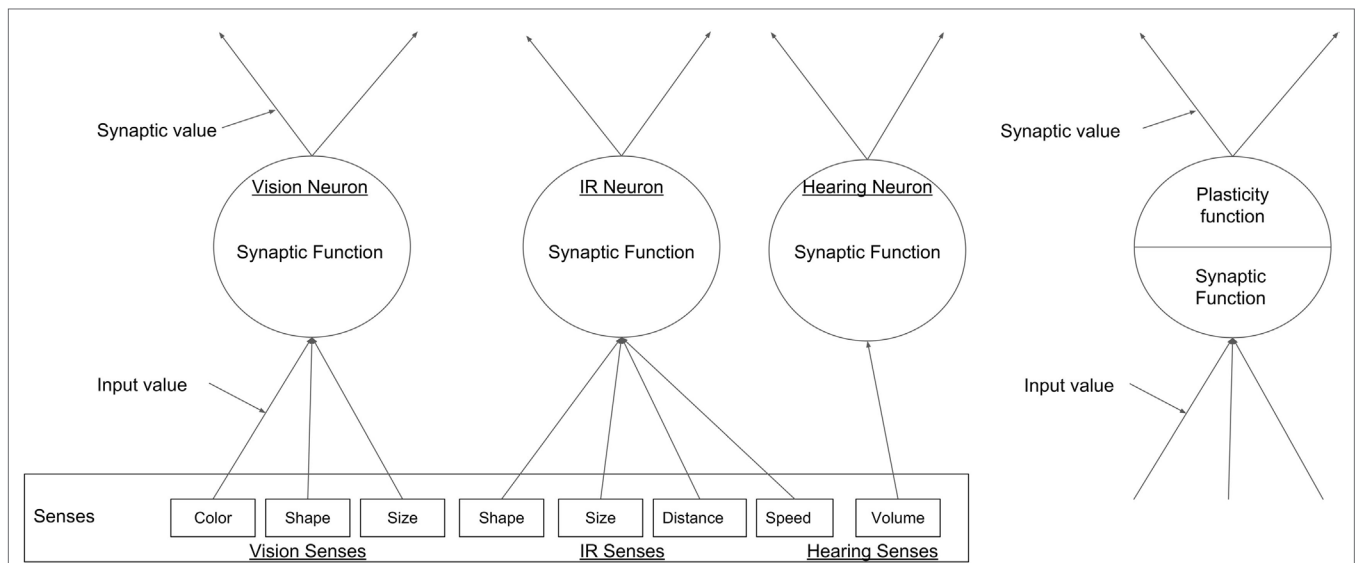


FIGURE 3 | A generic example of the types of nodes in the ENN, the three *input nodes*, each representing one of the robot’s senses can be seen on the left, and a hidden layer node on the right.

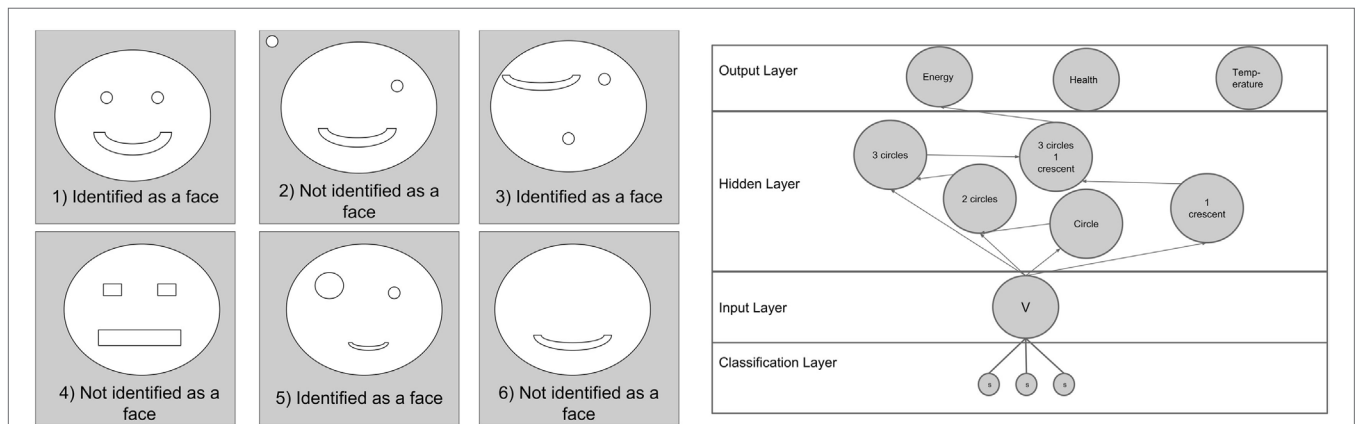


FIGURE 4 | To provide an example of how the robot perceives its environment, we have shown the robot a simple picture of a face, seen in image (1) on the left. A simple example on how the ENN may develop in relation to this picture is then shown on the right. The robot is shown the 6 pictures on the left to see which ones are identified as being the same. In this particular example, the robot has learned to identify the original by the presence of a large circle, 2 smaller circles and a half crescent; hence, samples 1, 3, and 5 on the left are all considered by the network to be the same face.

4 (if the face was directly ahead), and 3 (for the three circles: two eye circles plus the larger enclosing circle). It would additionally receive an input of 1, 4, 4, and 1 (indicating, respectively, vision, crescent, ahead, and one instance).

2.3.2. Hidden Layer

The second layer of the ENN is the hidden layer, which receives data from the input layer and sends data to the output layer. This layer initially starts empty, and nodes are created as a function of the robot's exposure to different stimuli. Creation of nodes takes place under two circumstances:

1. When an input node fires but has no synaptic connection to a relevant node (as described in the section above), or
2. When two or more hidden layer nodes fire at the same time.

When a new node is created in the hidden layer, in addition to being connected to the relevant nodes that led to its creation (which provide the input), it is also fully connected to the nodes of the output layer. However, all these different connections can disappear as the network continues to develop. When a synaptic connection between two nodes i and j is created, it is given a strength of $sP_{ij} = 0.5$. The connection strength is then updated as the robot interacts with its environment, using a sigmoid function, as seen in equation (8).

$$sP_{ij} = \alpha e^{\beta x_{ij}} \quad (8)$$

where α and β are constants, and x_{ij} is the sum of times the nodes i and j have fired together, minus the number of times that they have not fired together, within a range of $-10,000 < x_{ij} < 10,000$. A negative value of x_{ij} thus means that, more often than not, the nodes have not fired together.

Equation (8) results in a synaptic connection with strength in the range $0 < sP_{ij} < 1$. Due to the sigmoid nature of the function, the closer the synaptic strength gets to either end of this range, the lower the rate of change, or plasticity, of the connection. The synaptic strength of a connection between nodes plays a number of roles in this ENN, as will be discussed shortly. One of the most important roles is simply determining if a connection exists between nodes. This is achieved as follows:

- If a synaptic connection between two nodes exists but the synaptic strength drops below 0.05, then the connection is broken.
- If a synaptic connection does not exist but synaptic strength would be above 0.1 if it existed, then a connection is made.

In addition, when a synaptic connection is made from a new node to a node in the output layer, this connection is assigned an affordance – the potential to recover a homeostatic variable. Initially, this affordance is set at the change detected in the related output node's homeostatic variable. For instance, a new synaptic connection to the energy output node during a loop when the robot gained 2 units of energy will result in that connection receiving an initial affordance value of 2. This affordance assigned to the synaptic connection then changes as the robot continues

to interact with that particular aspect of the environment, as follows:

$$\Delta \text{Affordance}_{v,i,j} = \text{Affordance}_{v,i,j} \times sP_{ij} + \text{HomeostaticChange}_v \times (1 - sP_{ij}) \quad (9)$$

where $\text{HomeostaticChange}_v$ is simply the change in a homeostatic variable v in the current action loop compared with the previous one.

In order for a node i in the hidden layer to fire, it must receive a total input that is greater than or equal to its total number of synaptic inputs, thus:

$$\text{output}_{i,d} = \begin{cases} 1 & \text{if } \text{input}_{i,d} \geq sC_i \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

where sC_i is number of input synaptic connections for node i , and d ($0 \leq d < 8$) is the direction of the detected stimulus with respect to 8 equally spaced body coordinates of the robot, the third digit of the input pattern discussed in Section 2.3.1. Using this system, 0 represents the body coordinate directly behind the robot, and then going clockwise each subsequent value represents the next coordinate. For example, 4 represents a coordinate directly in front of the robot.

As shown in equation 10, if the firing threshold is reached, the nodes fire with a value of 1; however, the synaptic function $sF_{i,j,d}$, or output of the hidden layer node i , is then modified depending on the outgoing connection of the node and the directional origin of the stimulus d . If a hidden layer node i is connected to another hidden layer node j , the synaptic function is:

$$sF_{i,d} = \text{output}_{i,d} \times nH\text{modulation}_i \times \sum_v eH\text{Modulation}_{v,i} \quad (11)$$

where $eH\text{Modulation}_{v,i}$ is the modulation of the endocrine hormones eH_v on node i [see equations (13) and (14)] and $nH\text{modulation}_i$ is the combined strength of the modulation from the neuro-hormones stress and curiosity nH [see equation (15)]. The roles of hormones in the ENN are discussed in greater detail in Section 2.3.3.

If the node is connected to an output node for homeostatic variable v then the synaptic function is given by:

$$sF_{i,j,v,d} = nH\text{modulation}_i \times \text{output}_{i,d} \times \text{Affordance}_{v,i,j} \times eH\text{Modulation}_{v,i} \quad (12)$$

A basic example of how the ENN works and allows the robot to identify objects and stimuli can be seen in **Figure 4**, which shows how the robot perceives a simple face. Here, the robot is able to identify the face by the presence of the key characteristics of a large circle, 2 small circles and a crescent. However, the robot cannot detect spatial arrangements, and therefore, as long as the features are close enough, they will be identified as the same object. The characteristics used by the robot to identify objects depend on its past learning. A relatively new robot, for instance, may identify all pictures as being the same, since they all possess a circular shape. In contrast, a robot with greater environmental exposure, such as the one used in this example, will have more specific criteria.

2.3.3. Hormones in the ENN

As shown in equations (11) and (12), different hormone concentrations modulate the synaptic functions of the ENN. As discussed in Section 2.3.2, these hormones modulate the nodes within the ENN as a function of the internal state of the robot. In the case of the *eH* hormones, the strength of the modulation is dependent on the hormone's sensor's sensitivity and the connections of the nodes. For a node i directly connected to the output layer, the modulation by the eH_v is given by:

$$eHModulation_{v,i} = eHConcentration_v \times senS_v \quad (13)$$

where $eHConcentration_v$ is the concentration of the *eH* hormone eH_v , $senS_v$ is the sensitivity to eH_v [see equation (7)].

However, for nodes not directly connected to an output node, the modulation from the *eH* becomes weaker, as shown in equation (14), resulting in a larger modulation of the nodes closer to the output layer and/or with stronger synaptic connections to it. Using hormonal modulation in this manner promotes the activation of nodes that have a higher synaptic strength, and hence promotes behaviors that, in past interactions, have led to better homeostatic balance.

$$eHModulation_{v,i} = \sum_{j \in O(i)} \frac{eHModulation_{v,j} \times sP_{i,j}}{noI_j} \quad (14)$$

where $O(i)$ is the set of output nodes from node i , i.e., the set of nodes that are connected to output of i , $eHModulation_{v,i}$ is the strength of the modulation in the current node, dependent on the sum of the signal passed down from connecting nodes $eHModulation_{v,j}$, and noI_j the number of input connections of node j .

In contrast to the *eH* hormones, the *nH* hormones surround the ENN, affecting all nodes equally. The *nH* behave differently as their role is to either promote or suppress novelty-seeking behavior. This is caused by the combined effect of the curiosity and stress hormones. The curiosity hormone increases the activation of nodes with a low synaptic strength and suppresses nodes with a high synaptic strength. Conversely, the stress hormone increases the activation of nodes with a high synaptic strength and suppresses nodes with a low synaptic strength, as in equations (11) and (12). Therefore, the robot is using the synaptic strength as a way of assessing the novelty value of an object or aspect of the environment, since a high synaptic strength only happens if an object behaves as expected each time the robot interacts with it. This can be seen below in equation (15)

$$nHmodulation_{n,ij} = sP_{ij} \times nHConcentration_s \times senS_s + (1 - sP_{ij}) \times nHConcentration_c \times senS_c \quad (15)$$

where $nHConcentration_s$ is the concentration of the stress hormone (s), $senS_s$ is the receptor's sensitivity to the stress hormone, $nHConcentration_c$ is the concentration of the curiosity hormone (c) and $senS_c$ the receptor's sensitivity to the curiosity hormone.

2.3.4. Output Layer

The final layer of the ENN is the output layer, which consist of a fixed number of nodes equal to the total number of survival-related

homeostatic needs. Each output node simply sums up the total input from the hidden layer in order to calculate the affordance of moving in a certain direction.

$$output_{v,d} = \sum_i sF_{i,v,d} \quad (16)$$

The output of the ENN then feeds directly into (and modulates) the robot's actuators – in this case, the wheels:

$$WheelSpeed_i = \sum_{v,d} output_{v,d} \times set_{i,d} \quad (17)$$

where $WheelSpeed_i$ is the speed of the left (0) or right wheel (1), $set_{i,d}$ are constant vectors equal to $(-10, -10, -5, -3, 1, 3, 5, 10)$ if $i = 0$, or $(-10, 10, 5, 3, 1, -3, -5, -10)$ if $i = 1$. This means that if a single stimulus originating from the left side of the robot is detected, the robot's left wheel moves at a speed of $5 \times output$ and the right wheel moves at a speed of $5 \times output$. Therefore, a positive output will result in the robot turning toward the stimulus and a negative output in turning away from it.

To summarize, the causal chain that leads to internal or external stimuli promoting different behaviors is as follows:

1. As homeostatic deficits occur they lead to the release of the associated endocrine hormone *eH* [see equation (3)].
2. Internal and external stimuli lead to the release of the neurohormones, with curiosity being secreted in relation to perceived positive stimulation and stress in relation to negative stimulation [see equations (4) and (5)].
3. The robot's sensitivity to each hormone is dependent on its historic exposure to it [see equation (7)].
4. The artificial hormones modulate the synaptic function of the hidden layer nodes [see equation (12)] depending on the nodes position in the network [see equations (13) and (14)].
5. The output nodes sum up the synaptic function of connected neural pathway [see equation (16)], with the value dependent on past outcomes associated with pathways activation [see equations (8) and (9)].
6. The output from the network then affects the behavior (wheel speed), by promoting or suppressing the tendency to move in a certain direction at a certain speed. The larger the output, the greater modulating effect it will have on behavior [see equation (17)].

2.4. Experimental Setup

To test the effects of the previously described robot architecture, we allowed the robot to develop in three different environments with a single run of 60 min of duration in each. These three environments were (1) a base/standard environment, (2) a novel environment, and (3) a sensory deprivation environment. For each of these environments, the robot spent the first 10 minutes with a caregiver who looked after it, and introduced the robot to core components of the environment. In the remaining 50 minutes, the robots were then placed in their specific environments. The base and novel environments both consisted of our open lab environment (see **Figure 5**) with some differences that we will discuss in the relevant sections (Sections 1 and 2). In the sensory deprivation environment, the robot was placed and left



FIGURE 5 | Different aspects of the environment used during the experiments. (1) shows a panoramic picture of the standard open environment used during 3.1 and 3.2. (2) shows an example of one of the novel structures used during experiment 3.2. (3) shows the koala robot used during this work and the cardboard box the robot was placed in to create a sensory deprivation environment. (4) shows two AIBO robots used as novel objects in the test described in Section 4.2.

alone inside a cardboard box after the initial 10 minutes with the caregiver. In the three environments, the robot would have access to two sources of each type of resource; in the third experiment, this meant that the resources were placed inside the box along with the robot.

3. EXPERIMENTS AND RESULTS

Before the robots were placed into their different environments, each spent the first 10 min of their “life” with the caregiver. This caregiver provided an identical experience for each of the robots with the primary purpose to teach them the critical aspects needed to survive, such as how to recover from homeostatic deficits. This period essentially consisted of the caregiver satiating the robot’s needs by bringing the relevant resource to them. During this period, the robot’s behaviors were essentially driven by exposure to stimuli. In some ways, these basic behaviors are similar to the so-called reflex acts of a newborn. At this stage, both newborns

and the robots display many “reflex” behaviors; for instance, a newborn will “grasp” objects placed into their hand or suck an object placed against their lips; our robots’ “reflexive” behaviors will generally see them move toward or away from (attraction vs. repulsion) different environmental stimuli.

In this first phase, the interactions between the robots and the caregiver resulted in the emergence of five main reflex behaviors. The first three occur due to the homeostatic variables, which are Attraction/Repulsion, Avoidance, and Recoil. Attraction and Repulsion emerged when the caregiver fed the robots by placing a relevant resource in front of them. This “feeding” by the caregiver, made the robot move toward the caregiver when hungry and then away when satiated. Avoidance emerged when the caregiver moved too close to the robot, making the robot move toward an area with more space. Recoil emerged when physical contact occurred; unlike with the avoidance behavior, here the robot will prefer to move in an opposite direction to the stimulus rather than simply toward more space.

The final two reflex behaviors seen during this period are slightly different. The Exploration behavior emerges due to a combination of the first three behaviors: the Attraction behavior gives the robot the motivation to move forward while the Avoidance and Recoil lead to a motivation to avoid collisions. Finally, Localized Attention, the last innate behavior seen during this period, is based partly on learning and emerges around the 8-minute mark. This behavior sees the robot turn to face a moving object that is roughly within a 30-cm range. The basis behind this behavior can be traced to the fact that the robot at this stage associates movement with the presence of the caregiver³ and therefore the impending “feeding,” which can only occur if the robot is facing the resource (and hence the caregiver holding it). At the end of this initial period, the caregiver would leave the environment and be outside the robot’s view.

3.1. First Experiment: The Standard Environment

In the first of the three experiments, the robot was placed in our open lab environment shown in **Figure 5**. For this experiment, the robot was given free rein of our lab with only limited changes to the environment made. These changes include (a) the use of plywood borders to block access to “problem” areas where the robot’s sensors and actuators would be unsuitable and (b) the placement of resources. Additionally, blackout curtains were used to block natural light, in order to keep lighting conditions comparable through the experiments.

3.1.1. First Experiment: Minutes 10–20

During this period of the experiment, the caregiver was removed from the environment, and with him the feeding interaction between the caregiver and the robot. From now on, in order to maintain its homeostatic balance, the robot would need to seek out the different resources scattered throughout the environment. Resources were placed in manner in which they could be clearly seen by the robot – four resources, one in each corner, alternating in type – with the aim of causing it to move around the environment in order to experience different sensorimotor stimuli. The prior 10-min exposure to resources through the caregiver’s feeding was enough for the robot to have begun to learn some of the key features of the different resources, such as their shape, color, and size, to allow the robot to detect them.

The immediate challenges that the removal of the caregiver presents to the robot are threefold. First, the robot must be able to manage conflicting needs, e.g., if it chooses to replenish energy it must at least temporarily forgo reducing its temperature or replenishing its health. Second, the robot needs to develop tolerance so its consumption pattern – particularly to what level it can let a homeostatic variable drop before replenishing – is appropriate for the current environment. Third, the robot must adapt its sensorimotor behavior – how fast to move and when to turn to avoid collisions – to the current environmental conditions.

³Although the robot did possess color vision at this stage, perhaps due to environmental noise or to slower development of vision, the robot relied on movement to detect the caregiver.

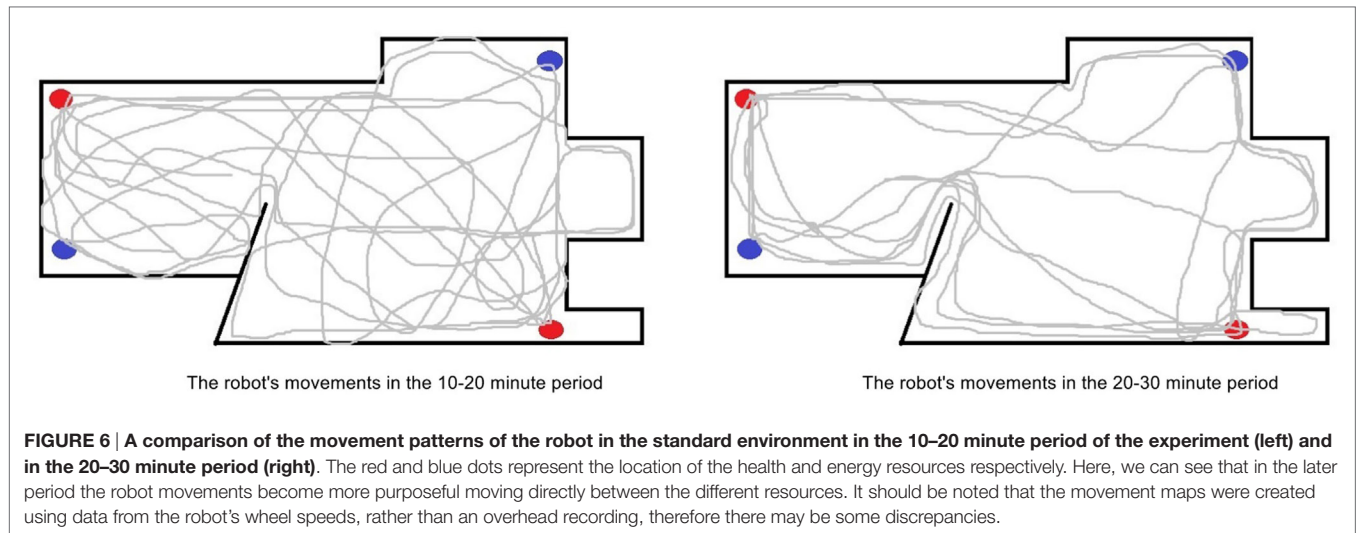
At the beginning of this period, the robot was highly sensitive to its internal needs – attempting to replenish any variable that was roughly below 90%. Due to spacing of the resources, the robot was often able to see at least one of each type at any given time, and therefore at this point, it did not search the environment when a deficit occurred but rather moved to the nearest perceived resource. This movement was often inefficient (see **Figure 6**) as in many cases a closer resource was located outside its immediate field of view, either to the side or behind. However, at this point in time, the robot’s behavior was still largely reflex-driven – seeing the resource made the robot move toward it. When two homeostatic variables were low and the required resources could both be seen, the robot’s choice of which variable to recover first would be determined based on the size of both the internal deficit and the detected stimuli. A problem with satisfying needs in this manner is due to a combination of noise – the perceived size of external stimuli would fluctuate – and homeostatic variables not decreasing linearly or at an equal rate; the robot’s intrinsic motivations would thus fluctuate, and hence its “goals” and executed behaviors often changed before a need was satiated as shown in **Figure 7**.

3.1.2. First Experiment: Minutes 20–30

The inefficiency in the robot’s behavior after the withdrawal of the caregiver initially leads to the robot having issues in maintaining homeostasis. However, after the robot had been sufficiently exposed to its environment and the epigenetic mechanism began to regulate hormone receptors, its behaviors became more appropriate, and the robot was able to recover a homeostatic deficit 54% faster on average. This can be seen in **Figures 6** and **7** which show, respectively, the change in the robot’s movement patterns and motivations.

As shown in **Figures 6** and **7**, the robot’s movements have become much more efficient for its environment, as it now moves more directly between the resources with limited motivation or behavior switching. This occurred first as a result of a change in tolerances to homeostatic deficits. As the robot had consistently lower but stable homeostatic variables due to needing to feed for itself, it soon became tolerant to these lower levels through the epigenetic mechanism. This resulted in reduced urgency in replenishing its internal variables, to the extent that they would now need to reach an average level of around 60% instead of the previous 90% before the robot would become motivated to replenish them. As a consequence of the reduced need to replenish the homeostatic variables of energy and health, the robot was no longer under such internal pressure to move quickly between the resources and could reduce its overall speed, resolving the issues of overheating and increased collisions associated with faster movement in the previous period. Additionally, while the robot maintained a relatively constant speed in previous periods, slowing down only to consume or due to internal overheating, now the robot began to modulate its speed to match the environmental conditions. For example, the robot would move slower near the edges of the environment where it previously had collisions, and faster in the open middle areas.

This period represented an important time in the robot’s development. As described previously, during the early stages of



this experiment, when the robot was first exposed to this environment, its behavior was almost entirely reflex-driven. However, due to motor stimulation, the robot's behavior has started to become adaptive, taking into account the current environmental conditions and its own physical body.

This period therefore potentially bears some similarities to the concept of primary circular reaction in infant development. Much like with infants at this stage, here the robot's focus is on the effects that its behaviors had on its own body – for instance, developing appropriate movement speeds and understanding and adapting to the restraints of the levels of its homeostatic variables. Similarly, for both the robot and the infant, behaviors categorized as primary circular reactions emerge as accidental discoveries (Papalia et al., 1992; Schaffer, 1996).

3.1.3. First Experiment: Minutes 30–40

During the first 30 minutes or so, the robot had begun to adapt its behaviors with regard to maintaining homeostasis by developing behaviors which have similarities to primary circular reactions. However, at this point, the robot began to show the emergence of more complex behaviors that could be considered similar to secondary or even tertiary circular reactions, as we will discuss in more detail below. At around the 33 minutes mark, due to the robot's previously discussed reduced need for, and increased efficiency in, maintaining homeostasis, the robot spent a much smaller proportion of its time attending to homeostatic needs, showing a reduction from 93% of its time actively searching for resources in the first 30 minutes down to 59% in this period shown in **Figure 8**. This reduction in time needed to maintain homeostasis provided the robot with the opportunity to *explore* and interact with other aspects of the environment. During this period of exploration, using the previously discussed novelty mechanism (see Section 2.2.3), the robot's motivations were determined by both the internal and external environment. Such exploration would take different forms, depending on hormonal levels. With high levels of the nH_s , which is associated with positive stimuli and a good level of

homeostatic variables, the robot's attention was focused on the novel aspects of the environment. These novel aspects tended to be objects or areas that the robot had limited knowledge of, and/or objects that had some perceived uncertainty or danger as to the outcome of any interaction. In contrast, with higher levels of the nH_s , which is associated with negative stimuli, over-stimulation and poor homeostasis maintenance, the robot is more attracted to, and will interact with, less novel aspects, such as those it already had some understanding of, or perceives to be safe, e.g., the walls of the environment due to their static nature. In cases where very high levels of the nH_s were present, the robot would simply move to an area of perceived safety and only leave when the nH_s levels had decreased sufficiently.

This period represented an important stage in development of the robot for two critical reasons. First, during this period, the increased exploration is strongly linked to the growth of the ENN (see Section 4.1). Second, this exploration and interaction represent an opportunity for the robot to further understand both its own body and the ways in which it can influence its environment. Due to the relatively static nature of this first environment – most objects were either immovable or too large for the robot to meaningfully interact with them – interaction was relatively limited; it consisted for the most part in pushing an object for a few seconds, before learning that the only outcome of this behavior was a reduction in its health due to the contact, thus reducing future attempts to interact with the said object. However, around the 38th minute, the robot found the resources which consisted of small plastic balls, light, and easy to push, and therefore the robot was able to create an interesting novel experience for itself by pushing the balls.

3.1.4. First Experiment: Minutes 40–60

During the latter stages of this experiment, due to improved efficiency in recovering homeostatic deficits, the robot spent most of the time either idle or interacting with resources. Initially, this interaction consisted of small pushes that took place over a period of around 10 minutes. The motivation for the robot to

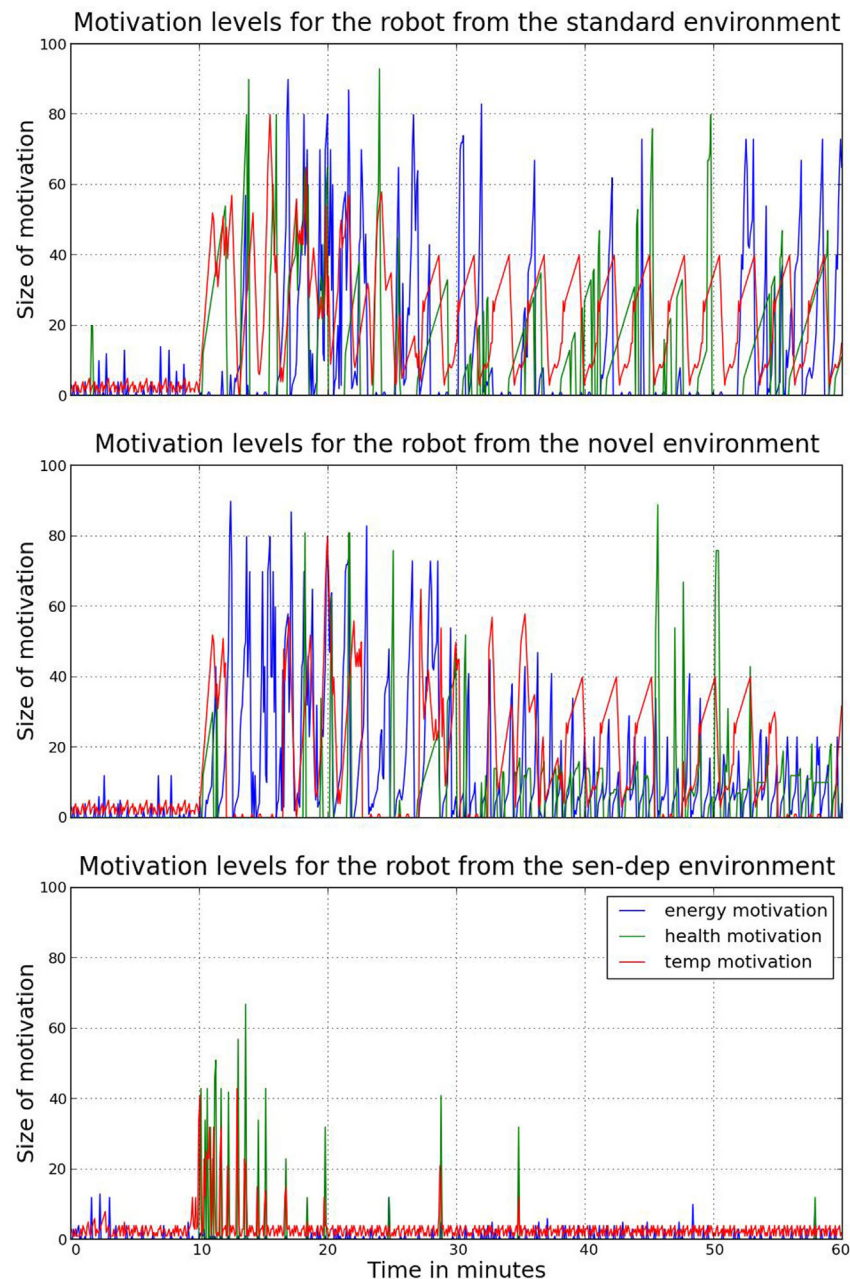
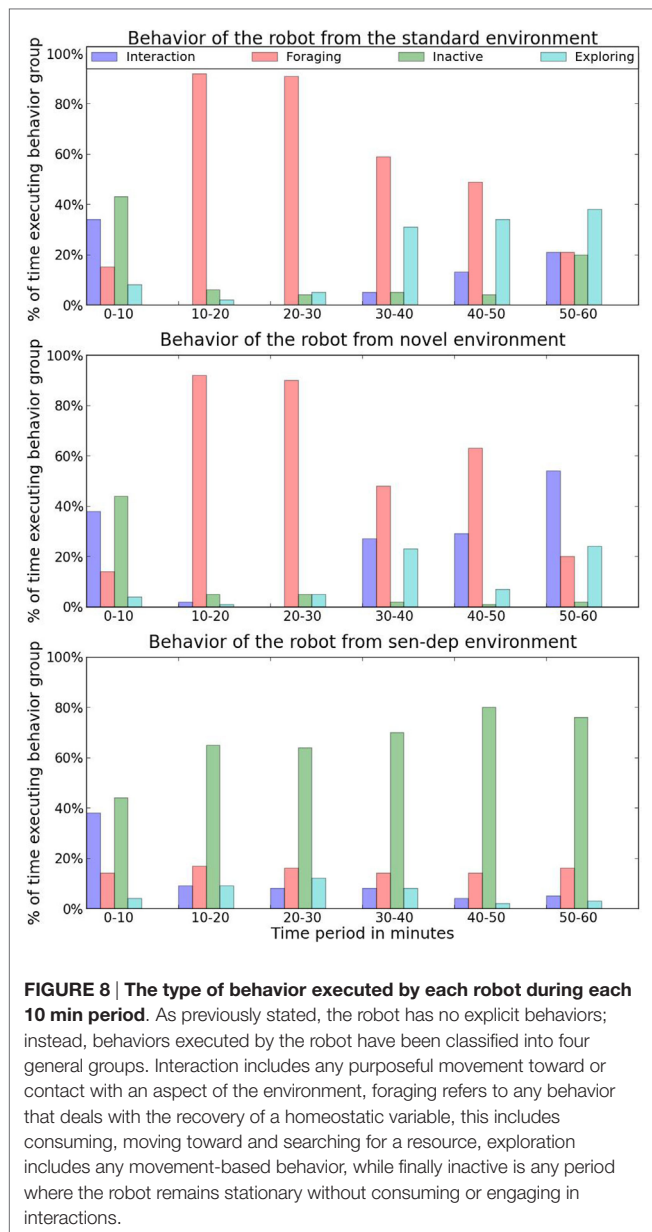


FIGURE 7 | A comparison of the three main motivations (replenishment of a homeostatic variable) for each robot during the experiments. As can be seen, during the period immediately following the initial ten minutes with the caregiver, changes and growth of motivations are much more volatile. This leads to increased occurrences of rapid behavior switching. Due to the volatility of change, this can lead to both inefficiencies and missed opportunities, i.e., constantly moving between two resources without (fully) feeding.

push the balls was twofold. Initially, the pushing was curiosity driven, as the robot tried to learn what the pushing resulted in. After around 5 minutes, however, the pushing became novelty driven, caused by the new element of motion, as mentioned in the previous section. As expected in our model, due to the high novelty that resulted from pushing an object, the robot would only “purposefully” push objects when it had high ratio of nH_c to nH_s concentration.

This emergent behavior presents some similarities with ideas of secondary circular reactions. For example, a child using a rattle and our robot pushing the ball share the fact that the agent is beginning to notice and explore that their actions and behaviors can have interesting effects on their surroundings. Similarly, later the behavior where we see the robot pushing the ball in order to create a novelty source has similarities to progression of secondary circular reaction to coordinated secondary circular reaction,



where the robot is now demonstrating the ability to manipulate an object to achieve a desired effect.

We observed another interesting phenomenon at around the 47 minutes mark, as the robot seemed to develop a search strategy while looking for resources. Previously, when searching for a resource, the robot would randomly explore its environment; however, at this point, the robot began to show some strategy in its search, since instead of the random exploration, it would now move to the walls and follow them to search for the resources, which were placed near the corners of the environment. The emergence of this behavior further reduced the average time spent searching for a resource from the previous 59% down to 47%. As time went on, this behavior continued to develop and the robot began to learn to associate certain easily identifiable

landmarks in the lab, such as a blue screen or a cupboard, with the presence of a particular resource. This ability greatly improved the time needed to find a resource, further reducing the average time spent searching for resources down to 21%. This behavior might suggest that the robot had developed some notion of “object permanence”. However, it may be a simple association between resources and landmarks, which is a significantly simpler concept than object permanence. In order to investigate which of these might be the case, we carried out the experiments reported in Section 4.3.

3.2. Second Experiment: The Novel Environment

In the second experiment, we developed the robot in an environment very similar to the one used in the first experiment, with the difference of the inclusion of a range of different novelty sources. These included light movable objects arranged in various shapes and patterns, as shown in **Figure 5**, as well as two small Khepera robots that moved around randomly. If, at any time, any of these object were knocked over (e.g., due to the Koala robot’s interactions) or stopped functioning as intended, the caregiver would replace or reset them as soon as the robot had moved away.

3.2.1. Second Experiment: Minutes 0–30

As we would expect, in the early stages of this experiment the exposure to additional (with respect to the first environment) sources of novelty had no real effect on the robot due to its effort to maintain homeostasis. Apart from the need to avoid the two additional randomly moving robots and the additional novel objects, the behavior and development of this robot was almost identical to the robot in the first experiment as shown in **Figures 7 and 8**. For this reason, we will not spend time discussing this robot’s early life but will rather move on to the second half of the experiment, when the behavior started to deviate.

3.2.2. Second Experiment: Minutes 30–50

Much like the robot in the first experiment, at around 33 minutes into its development, this robot had adapted to its environment well enough to no longer need to spend the majority of its time looking for resources. The exception to this, shown in **Figures 7 and 8**, occurs between the 40th and 50th minutes. Due to the increased interaction with objects as discussed shortly, the robot suffers additional health damage as it learns how to properly interact; therefore, it spends additional time during this period recovering its health variable.

While the robot in the first environment spent much of its “free time” being idle simply due to a lack of things to do, (i.e., a very limited number of novelty sources to interact with), this robot had a much larger range of possible objects to learn about. As before, the robot’s interest in the novel objects in the environment depended on the concentration of the nH_c and nH_s hormones. Initially, with a high value of the nH_c , the robot’s attention was mostly focused on the randomly moving robots. During this period of high concentration of nH_c , in the initial instances, the robot would simply engage in a following behavior

moving behind the nearest moving robot. After around 2–3 instances of this following behavior, the robot began to intensify its interaction by engaging in both pushing and approaching the small robot from different angles. Since the randomly moving robots had been programmed to stop if contact was detected, the novelty value that the robot would associate with them greatly diminished over a period of around 5 minutes, dropping to almost zero novelty near the start of the 50th minute as shown in **Figure 9**.

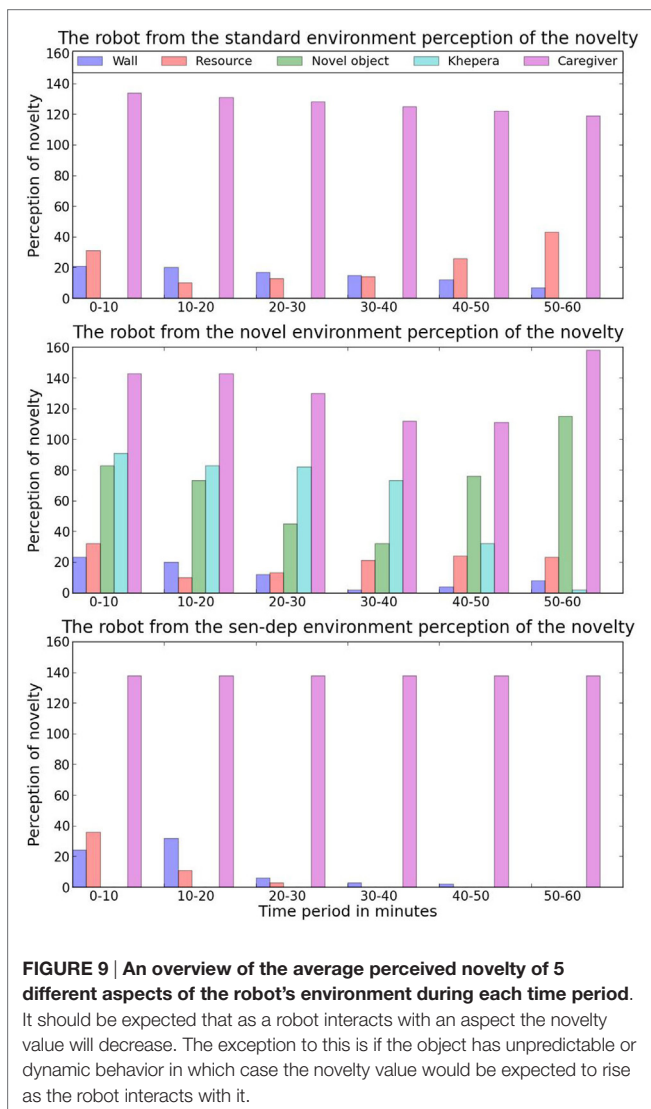
In contrast, with a medium concentration of the nH_c , the robot was attracted to the different arrangements of objects that were constructed with the small tin cans (see **Figure 5**). Initially, the robot would either move close to these structures or slowly circle around them. After a couple of minutes, when the robot was familiar with the structures, it began to make physical contact with them through gentle bumps and pushes. Due to the lightweight nature of the tin cans, any physical contact from the robot would easily knock them over, and this resulted

in the robot detecting not only a large amount of unexpected rapid movement around itself but also collisions, as some of the tin cans hit the robot. Since the robot only had a moderate amount of the nH_c when initially interacting with the structures, their falling resulted in significant over-stimulation, leading to increased secretion of the nH_c , and the robot's withdrawal to a perceived safer location. These implications of the early contact with the structures resulted in the robot associating a higher level of perceived novelty with them due to the uncertainty of the outcome of any interaction. This increase in novelty associated with the structures along with the decrease in novelty associated with the Khepera robots resulted in structures having the highest perceived novelty as shown in **Figure 9**. Due to the increased perceived novelty, the robot would now only interact with the novel structures with high nH_c levels. The higher concentration of nH_c protected the robot from becoming overstimulated due to unpredicted outcomes, which led to more thorough interaction with the structures. In the last 5–10 minutes of this period, the robot engaged with the structures in a number of different ways as it attempted to learn about them – including moving around them at different speeds, stopping near them at different distances, trying to move through them, and pushing them with different intensities.

3.2.3. Second Experiment: Minutes 50–60

At around 54 minutes into its development, the robot started displaying a new behavior: it would gently push over a structure before moving away and stopping. As we previously mentioned, when a structure was knocked down, the caregiver would replace it when the robot had moved away. As soon as the caregiver entered the environment to replace the tin cans, the robot immediately moved toward them and tried to interact with the caregiver. The caregiver, due to a number of factors such as size, shape, and movement, was unsurprisingly perceived as highly novel by the robot (see **Figure 9**). What was, however, interesting is that the robot seemed to engage in this sequence of behaviors “on purpose.” It is likely that, after experimenting with the objects, the robot had learned that by pushing the structures over, it could cause the caregiver to enter the environment and use this to satisfy its own need for novelty. Before the 54th minute, the robot had not displayed this behavior sequence of trying to have the caregiver enter the environment; yet, after the first occurrence, in the remaining 6 minutes of the experiment, this behavior occurred 11 additional times. In all cases, this behavior only occurred with high nH_c and low nH_i levels, supporting the idea that the robot was using this behavioral sequence “on purpose” to satisfy its own need for novelty. Examining the ENN seems to back up this idea, as neurons associated with the caregiver were active when interacting with the tin cans.

This behavior by the robot could be regarded as the emergence of a form of tertiary circular reactions and potentially bear a similarity to a representation of cause and effect. With regard to tertiary circular reactions, the robot was demonstrating the ability to not only manipulate and experiment with different objects in its environment, but also to use these objects in order to change



its environment, thus suggesting some sort of representation of cause and effect, an aspect of tertiary circular reactions (Papalia et al., 1992).

The formation of these representations is less clear, though, since the robot's behavior of knocking over structures in order to bring the unseen caregiver back into the environment could potentially suggest object permanence, which we test later in Section 4.3.

3.3. Third Experiment: Sensory Deprivation

In the final experiment, instead of being allowed to move freely in an open environment like in the previous experiments, after the first 10 minutes of interaction with the caregiver this robot was placed in a small cardboard box with the resources directly in front of it, in an attempt to create a sensory deprivation experience (see **Figure 5**). As would be expected, with both resources directly in front of it and little room to move, the robot remained mostly inactive throughout the sensory deprivation period as shown in **Figure 8**.

4. COMPARING THE COGNITIVE DEVELOPMENT OF OUR ROBOTS

From the overview of the experiments, it appears that the robot that developed in the novel environment (Section 3.2) gained more advanced cognitive abilities than the robots developed in the standard and "sensory deprivation" environments. These advanced cognitive abilities would seem to support the idea that an environment which provides a richer sensorimotor experience over the course of development leads to a greater cognitive development in autonomous robots too. However, we must ask the question whether these more advanced cognitive abilities are a permanent result of the actual developmental process, or a transient phenomenon due to the different environmental conditions. In order to try to understand if these developmental conditions had indeed affected the cognitive development of the robots, in the following section we compare the robots' neural networks and behavior in different developmental tests.

4.1. Comparison of Neural Development and Activity

For a first comparison between the robots, we will look in closer detail at the development of their different neural networks, which can be seen in **Figure 10**.

Figure 10 shows that the robot from the novel environment developed a larger neural network with significant growth occurring in the latter stages of the experiment, coinciding with the robot going through what in Section 3.2.2 we considered related to the coordination of secondary and tertiary reactions during the exploration period. Additionally, we can see again that the robot from the novel environment had a significantly larger number of neurons firing per action loop in the later stages. The increased number of nodes and neural activity from the novel robot can be explained due to this robot developing larger neural pathways. The increased pathways benefited the robot by giving it a better "understanding" of its environment.

4.2. Learning and Association: Introduction of a New Object

We next tested the ability of each of the three versions of the Koala robot – the robots from the experiments carried out in the "standard," "novel," and "sensory deprivation" environments – to learn by introducing two new novel objects that the robots had not seen before – two AIBO robots (one white and one black) shown in **Figure 5**. These novel objects were set to work in a similar manner to the energy resource, recharging the energy of the robot when it was close, although these novel objects provided a much greater rate of energy replenishment – 30 units of energy per second, 4 times faster than the original energy resource. The Koala robots were then given a choice between the novel objects and the original energy resource, with the assumption that if/once the robots learned that the novel resources provided a greater charge, they would prefer them over the original resources.

In order to conduct this experiment, two minor changes were made to the robot's architecture. First, the energy level was set to 20% after every action loop, to ensure the robot had a permanent motivation to recover from energy deficits. Second, the secretion of the nH_c was suppressed to remove the motivation to move to the novel resources based purely on their novelty value. The experiment involved two parts, with results show in **Table 2** and **Figure 11**.

For the first part of the experiment, the first novel object (the white AIBO) was placed directly in front of the Koala robot (close enough to charge) for a period of 10 s, to give the Koala an opportunity to learn about it; after this period, both this novel object and the original energy resource were placed slightly spread in front of the robot at a distance of around 1 m, forcing the robot to choose which one to move to in order to replenish its energy levels. This entire cycle was then repeated 10 times.

The results are reported in **Table 2**, where we can see that the "novel" robot appears to immediately learn the increased energy affordance provided by the first novel object and was significantly more attracted to it. In comparison, the "standard" robot would often pick the novel resource after increased exposure to it, although as seen in **Figure 11** it was only slightly preferred. The "sensory deprived" robot did not show any signs of adaptation, systematically selecting the original energy resource.

For the second part of this experiment, conducted immediately after the first part, we changed the first novel object with the second (the black AIBO). Unlike in the previous part of the experiment, the new novel object was not placed in front of the robot at any time; instead, it was placed 1 m ahead of the robot and slightly spread. Once again each of the versions of the Koala robot underwent another 10 runs with a similar need to replenish its energy level. While the robot had never seen the second novel object before, this object shares similarities with the first, hence here we are testing if the robot can identify that the two novel objects share similarities and therefore may behave in a similar manner, i.e., both would offer rapid replenishment of the energy deficit. The results are shown in **Table 2**, where we can see that even though the novel robot had never seen or interacted with the new novel object, unlike the other robots, due to its more

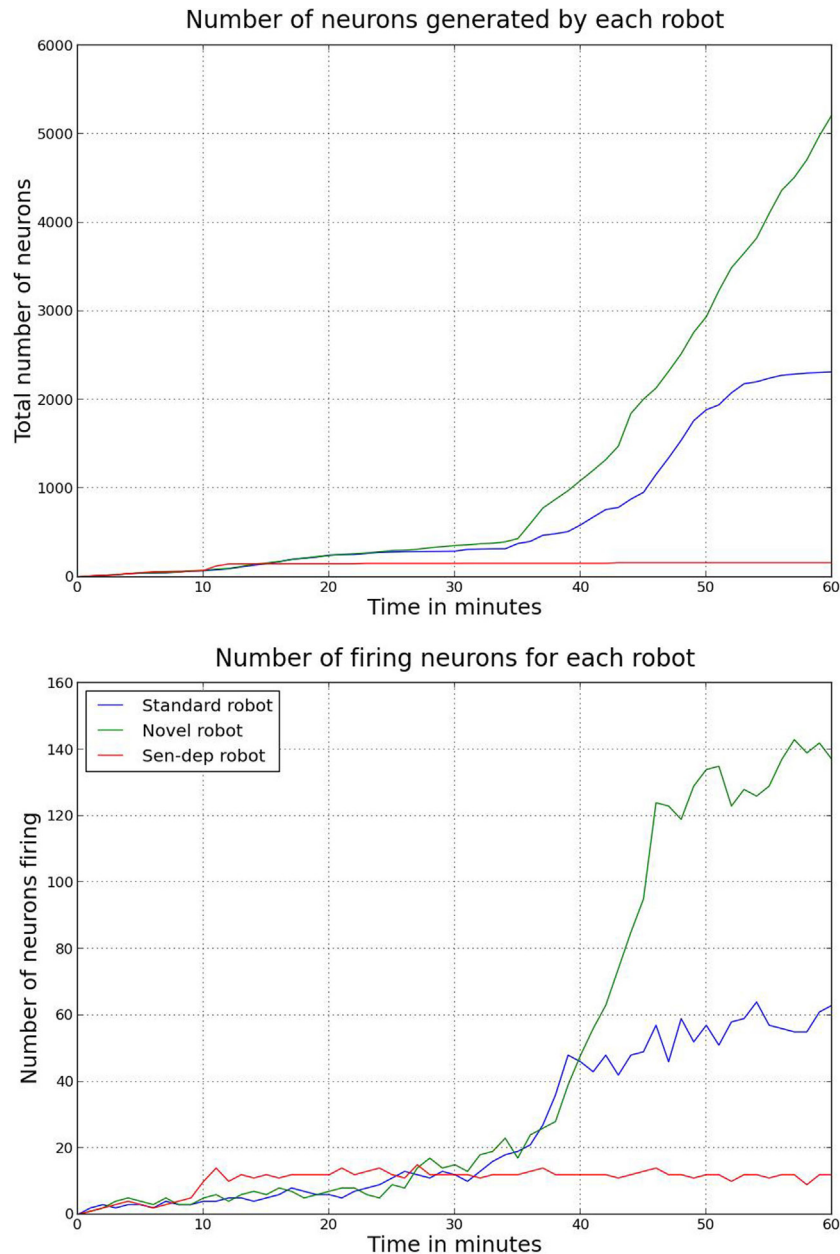


FIGURE 10 | The total number of neurons generated by the ENN for each of the robots over the course of the experiment (top) as well as the total number of firing nodes (bottom). The nodes themselves are generated as a function of the robots interaction with its environment. A higher number of nodes would suggest that the robot has learned about a larger number of, or in more detail about, different objects or aspects of its environment. A higher number of nodes firing appear to be related to the robot either noticing more aspects of the environment or having a greater understanding of the affordance of different aspects of the environment.

developed neural network (Section 4.1), it was able to identify the similarities between the two novel objects and recognize that the second had similar properties (i.e., the ability to provide a rapid charge) to the first. In contrast, while the standard robot did seem to identify some similarities between the two novel objects, leading to a slight perceived affordance of energy recovery with the new objects, the perception was not good enough for it to choose the novel object over the safer original energy resource.

Finally, the sensory deprived robot, which only showed minimal learning in the first stage of this experiment, showed no association between the two objects.

4.3. Object Permanence: Recreation of a Hidden-Toy Test

One of the tests most commonly used in developmental psychology to assess whether infants have acquired the notion

TABLE 2 | The robots' choices between the novel and the original resource in the first (left) and the second part (right) of the learning experiment.

| First experiment | | | | Second experiment | | | |
|------------------|----------------|-------------|-------------------|-------------------|----------------|-------------|-------------------|
| Run | Standard robot | Novel robot | Sensory-dep robot | Run | Standard robot | Novel robot | Sensory-dep robot |
| 1 | Original | Novel | Original | 1 | Original | Novel | Original |
| 2 | Original | Novel | Original | 2 | Original | Novel | Original |
| 3 | Novel | Novel | Original | 3 | Original | Novel | Original |
| 4 | Novel | Novel | Original | 4 | Original | Novel | Original |
| 5 | Novel | Novel | Original | 5 | Original | Novel | Original |
| 6 | Novel | Novel | Original | 6 | Original | Novel | Original |
| 7 | Novel | Novel | Original | 7 | Original | Novel | Original |
| 8 | Original | Novel | Original | 8 | Original | Novel | Original |
| 9 | Novel | Novel | Original | 9 | Original | Novel | Original |
| 10 | Novel | Novel | Original | 10 | Original | Novel | Original |

of object permanence is the hidden toy test⁴ (Piaget, 1952; Baillargeon, 1993; Munakata, 2000). We reproduced this test by placing a needed resource in front of each of the Koala robots at a range of 2 m. As the robot began to move toward the resource, 5 tins cans, used to build the previous novel structure shown in **Figure 5**, were placed directly in front of the resource to block it from the robot's view. If the robot has a representation of object permanence, we would hypothesize that the robot would continue to move toward the object even when it is hidden from sight. If the robot stopped for more than 10 seconds, or 1 minute had passed after the resource had been hidden, the experiment ended to reduce the risk that the robot might find the resource accidentally or as part of its exploratory behavior. This experiment was conducted 10 times for each robot, and the results are shown in **Table 3**.

As shown in **Table 3**, the robots from the novel and standard environments both had some success in finding the resources once hidden; in comparison, the robot from the sensory deprivation environment was unsuccessful every time. If we look at the behavior and neural activity of the robots, shown in **Figure 12**, we can see that the robot from the novel environment was the only one to consistently search for the resource after it was hidden. In addition, this robot was also the only one which consistently (i.e., in every run) had high activity along the neural pathway associated with the detection of the resource even after it had disappeared. This neural activity resulted from the fact that the original signal remained active along the pathway due to the modulation of this pathway by the different hormone concentrations, leading to feedback loops. These loops provide the robot with an ability akin to "active," or "short term" memory.

The 3 occasions when this robot failed to find the resource were due to the fact that the robot moved past the hidden resource. The fact that the neural activity remained high during these failed attempts suggests that, while the robot shows neural

activity associated with the hidden object and the behavior it affords, without the expected feedback from sensory readings regarding the distance and position of the object, the robot cannot consistently locate it. This would appear to back up the previous observation that the first two robots had gained an ability consistent with the "understanding" of object permanence during their developmental runs, rather than having this skill from the start.

4.4. Violation of Expectation Paradigm

For the final experiment, we tested the robots using another common cognitive test, the Violation of Expectation paradigm (VOE). VOE experiments are normally carried out by showing very young infants two different pictures, one of which shows an impossible outcome – often some type optical illusion – while the other is almost identical but without the impossibility (Sirois and Mareschal, 2002). The experiment seeks to assess if the baby can notice the impossibility by measuring which picture it looks at more. The underlying assumption is that if the baby can identify the impossibility in the picture, it must have some expectation about the object represented in that picture, and will look at it for longer than at the image without the impossibility.

We created a version of this experiment suitable for our robots. Here, a white ball was placed in front of the robot. For the possible outcome, we simply measured how long the ball which has not been seen before held the robot's attention. For the "VOE," the white ball was again placed in front of the robot; however, the robot's sensors were manipulated to make it appear as if the ball became smaller as the robot moved toward it. We once again measured how long the ball held the robot's attention. If the robot can identify the "VOE," we would expect it to hold its attention for a longer period of time.

As shown in **Table 4**, the "VOE" held the novel robot's attention for significantly longer than the possible object. In contrast, the robot from the standard environment only showed slightly more interest in the "VOE"; due to the small difference, it is not possible to conclusively suggest that this robot was showing an interest in the "VOE." Finally, the sensory deprived robot showed no real difference in the time spent with both objects, suggesting the VOE paradigm had no real influence on the robot.

Our results suggest that the ability to respond to "VOE" arises as part of the later stages of the sensorimotor development process, which only the robot from the novel environment went through. Incidentally, this finding correlates with Piaget's developmental theory regarding when these skills should emerge. It should be noted that the VOE paradigm has been criticized by other developmental psychologists, who debate whether these skills are indeed learned as suggested by Piaget (1952) in his theory of development, or if they are part of a core knowledge that all infants possess from birth, as suggested by Baillargeon et al. (1985), Baillargeon (1993), and Spelke et al. (1992). These authors have previously used the VOE paradigm to demonstrate that babies are able to identify impossibility much earlier than would be expected if Piaget's theory was correct. In the case of a robot, we can be certain that this was not part of the robot's core knowledge but is developed given the appropriate sensorimotor experiences.

⁴We have also carried out experiments using the A-not-B test with robots that were developed in an environment with human "caregivers". Results of these experiments will be reported in a forthcoming publication.

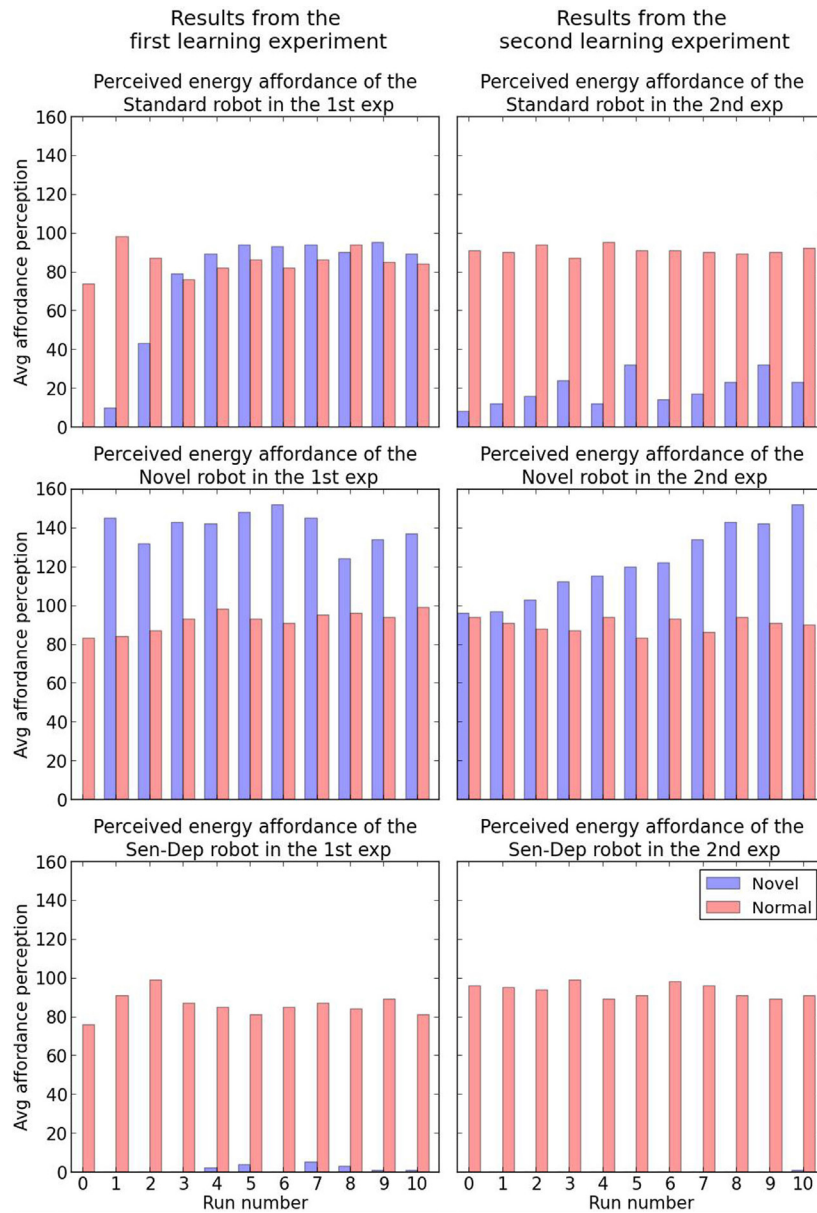


FIGURE 11 | The results of the first (left) and second (right) part of the learning experiment. The units in the y-axis show the strength of the perceived energy affordance of the two objects which is determined by equation (12). Run 0 is the perceived energy affordance before the start of experiments.

TABLE 3 | Results of the hidden-toy test.

| Run | Novel robot's time | Standard robot's time | Sens-dep robot's time |
|-----|--------------------|-----------------------|-----------------------|
| 1 | 32 | 16 | Not found |
| 2 | 14 | Not found | Stopped |
| 3 | 17 | Not found | Stopped |
| 4 | 32 | 32 | Stopped |
| 5 | Not found | Not found | Stopped |
| 6 | Not found | Not found | Stopped |
| 7 | 15 | Not found | Stopped |
| 8 | Not found | 43 | Not found |
| 9 | 23 | Not found | Stopped |
| 10 | 19 | 12 | Stopped |

Time is measured in seconds.

There have also been debates (Baillargeon et al., 1985; Baillargeon, 1993) regarding whether the violation of expectation paradigm used with infants may have been biased, suggesting that the impossible variation offered other additional stimuli (e.g., more activity or increased number of elements in the impossible picture), which attracts the infants' attention rather than their ability to identify or be attracted to the perceived impossibility. We could also imagine that the higher novelty of the impossible object might to some extent be responsible for the response of the infants. In the case of our robot experiments, we would be happy to accept that the different response that the novel robot displays in the face of impossible experiences might be due to novelty. However, the novelty offered by the impossible

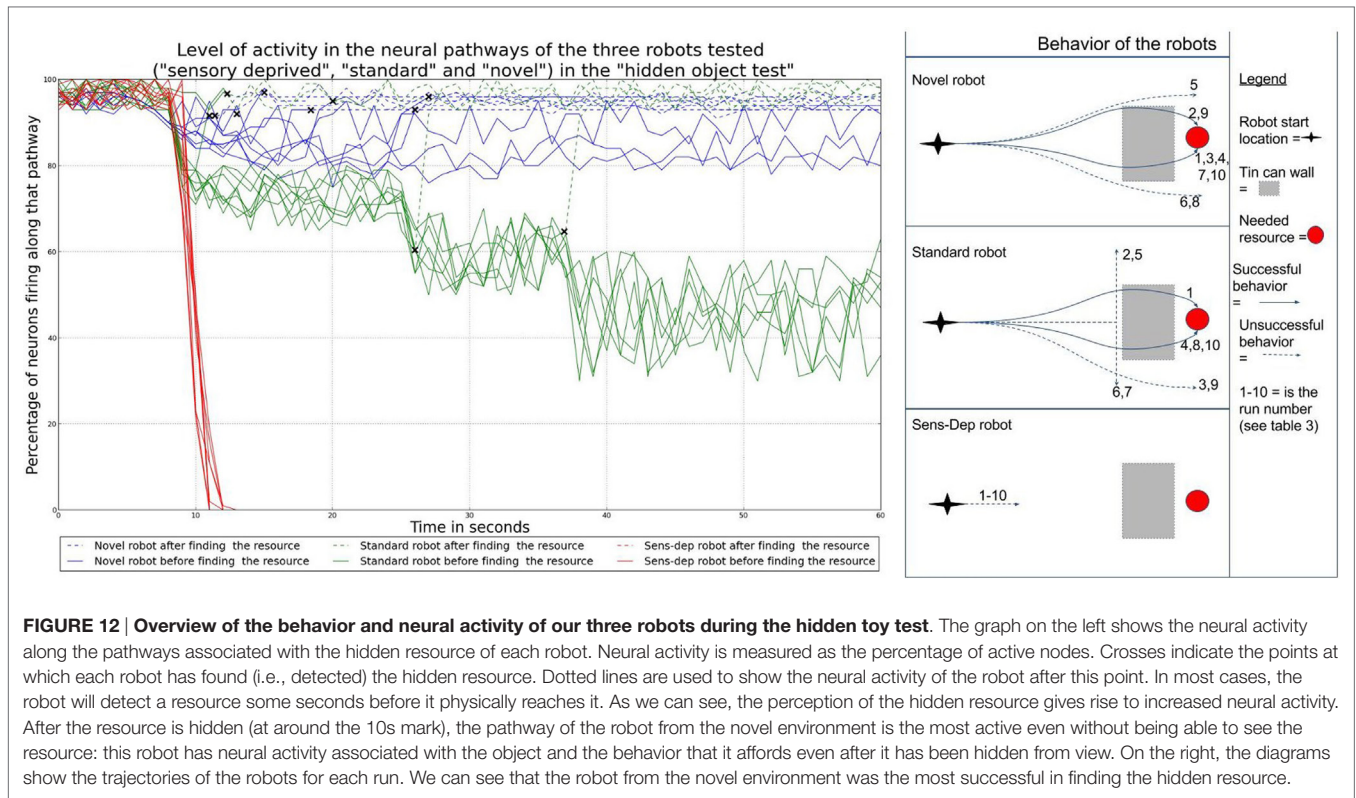


FIGURE 12 | Overview of the behavior and neural activity of our three robots during the hidden toy test. The graph on the left shows the neural activity along the pathways associated with the hidden resource of each robot. Neural activity is measured as the percentage of active nodes. Crosses indicate the points at which each robot has found (i.e., detected) the hidden resource. Dotted lines are used to show the neural activity of the robot after this point. In most cases, the robot will detect a resource some seconds before it physically reaches it. As we can see, the perception of the hidden resource gives rise to increased neural activity. After the resource is hidden (at around the 10s mark), the pathway of the robot from the novel environment is the most active even without being able to see the resource: this robot has neural activity associated with the object and the behavior that it affords even after it has been hidden from view. On the right, the diagrams show the trajectories of the robots for each run. We can see that the robot from the novel environment was the most successful in finding the hidden resource.

TABLE 4 | Results from the VOE experiment showing the time (in seconds) spent focusing on or interacting with the impossible and possible object.

| Run | Novel robot's time | | Base robot's time | | Sens-dep robot's time | |
|-----|--------------------|----------|-------------------|----------|-----------------------|----------|
| | Impossible | Possible | Impossible | Possible | Impossible | Possible |
| 1 | 46 | 17 | 21 | 18 | 19 | 20 |
| 2 | 48 | 15 | 22 | 19 | 23 | 21 |
| 3 | 43 | 17 | 19 | 22 | 21 | 23 |
| 4 | 49 | 18 | 17 | 15 | 20 | 20 |
| 5 | 54 | 14 | 19 | 12 | 15 | 15 |
| 6 | 52 | 18 | 16 | 15 | 19 | 20 |
| 7 | 49 | 19 | 15 | 14 | 18 | 19 |
| 8 | 56 | 16 | 14 | 13 | 19 | 20 |
| 9 | 49 | 14 | 16 | 16 | 14 | 13 |
| 10 | 54 | 17 | 21 | 15 | 18 | 16 |

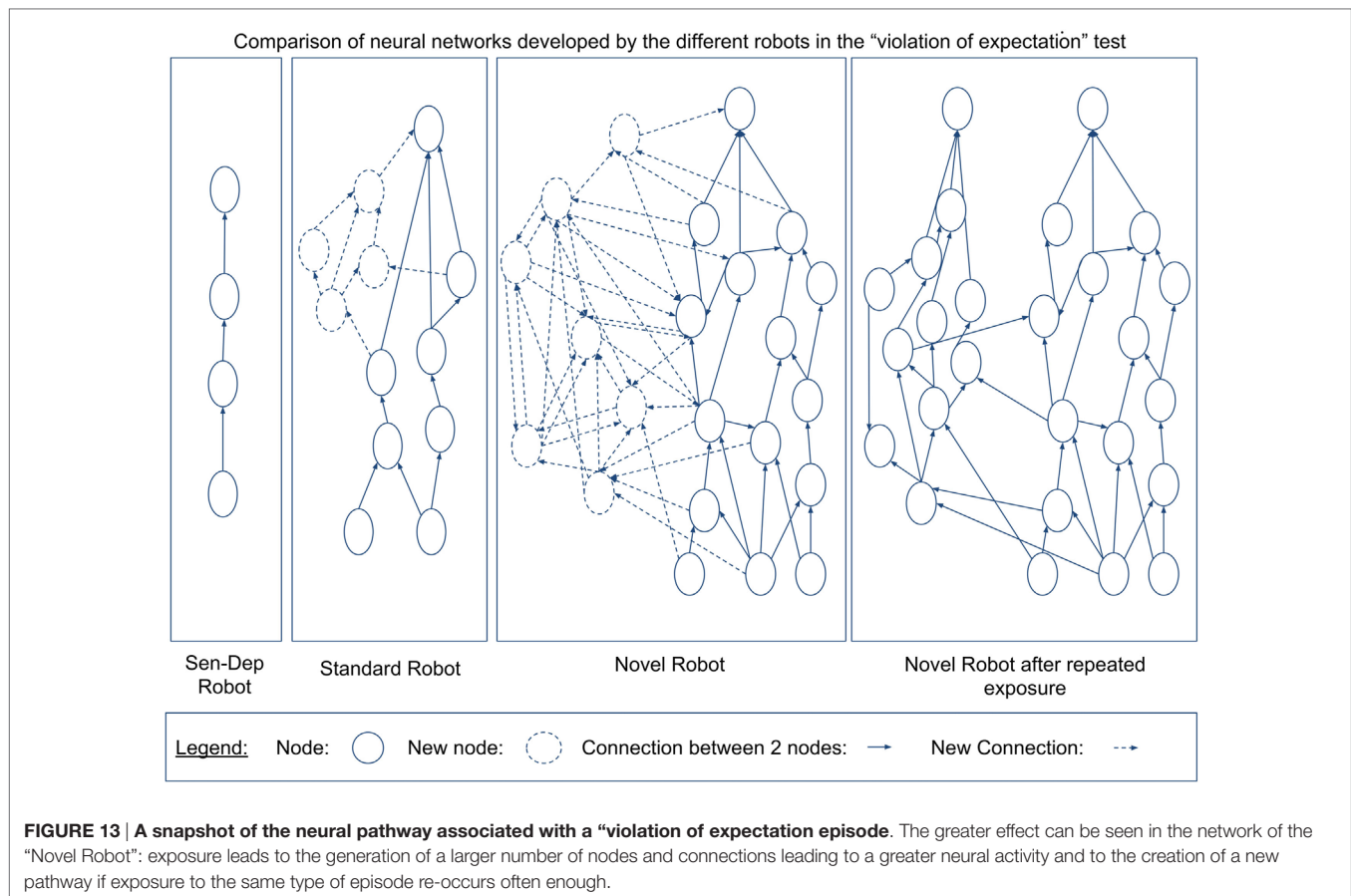
object (a white ball which behaves in a way that violates all the robot's previous sensorimotor experiences) is very different from the type of novelty offered by the perception of the novel object (white ball).

Although both exposure to a new object and a "VOE" episode produce "novelty," which we could define as the lack of behaviors and representations associated with an object, there is a qualitative difference in the effects that both experiences have on the robot's neural network. Novelty that arises from exposure to a new object is related to the level of plasticity between two connecting neurons [see equation (15)]: the higher the novelty of the object, the lower the level of plasticity, and a totally new object will give rise to new nodes and connections. Novelty related to

a violation of expectation episode involves, in addition to the above, a change in the neural pathway associated with closely related previous experience. Specifically, an existing pathway is activated but a number of new nodes rapidly emerge along this pathway, linked to the elements of the new experience that violate the expectations from previous experiences (see **Figure 13**, label "Novel Robot"). Due to the activity of the pathway and the rapid growth of new nodes along it, a large number of "messy" and overlapping connections between nodes are quickly generated, increasing perceived novelty due to their high plasticity. These overlapping connections also effectively lead to the emergence of a sort of positive feedback loop within the pathway. This results in a greater level of activity along that pathway and therefore increased novelty. Over time, if the robot is repeatedly exposed to the "violation of expectation," the original nodes and the new nodes will separate along two distinct pathways, the overlapping connections get "pruned," and the feedback loops disappear (see Section 2.3.2), as shown in **Figure 13**, label "Novel Robot after repeated exposure". This results in the consolidation of the new pathway and thus the previous "violation of expectation" becomes a normal "experience" for the robot. However, if exposure to the "VOE" stimulus is infrequent, the pathways will not split and the novelty associated with it will persist.

5. CONCLUSION AND DISCUSSION

In this paper, we have demonstrated the importance of sensorimotor experiences and environmental conditions in the emergence



of more advanced cognitive abilities in an autonomous robot. The robot exposed to a wide range of novel sensorimotor experiences and different stimuli showed greater cognitive abilities than the robots from the other environments. Particularly, the robot raised in a context of sensory deprivation showed no additional behaviors or abilities outside simple reflex-like behavior that each robot started with. Furthermore, this sensory deprived robot performed badly at adapting and in learning tasks compared with the other robots. Our model thus shows that a richer sensorimotor experience during early development correlates with greater cognitive ability.

We have also shown how an autonomous robot implementing an epigenetic architecture has the potential to go through developmental stages in a similar manner as outlined in Piaget's sensorimotor theory. Our robot starts with a simple reflex-like behavior, yet through interactions with the external environment and stimulation from its internal environment, it develops more complex behaviors and cognitive abilities. Our robot was not explicitly designed around a developmental theory, but these developmental substages emerged purely due to the interactions among the different aspects of the architecture – the hormonal, epigenetic, and ENN – and the external environment. In our past studies such as Lones and Cañamero (2013), when not all the previous components were present, the developmental phenomena described here did not emerge. In particular, the addition of the ENN, with its learning and

representational capabilities, in interaction with the other elements of the architecture, was a key factor in the emergence of these developmental phenomena. Our model thus offers potentially useful insights to bridge gaps between studies of epigenetic mechanisms such as Crews' (2010) and developmental epigenetic theories such as Piaget's (1952) by showing how the former can lead to the emergence of the latter.

AUTHOR CONTRIBUTIONS

JL: lead author, designed, conducted, and analyzed the experiments, wrote the first draft of the paper, and addressed the revisions. ML: second Ph.D. supervisor of the lead author, provided substantial advice regarding formalization and analysis of the experiments and substantial contributions to improving, editing, and rewriting various drafts and proofs of the paper. LC: main Ph.D. supervisor of the lead author and senior author of the paper, provided general advice on this work and substantial contributions to improving, editing, and rewriting various drafts and proofs of the paper.

FUNDING

This research was partly funded by a PhD studentship awarded to JL by the University of Hertfordshire and by the European Commission under FP7 grant ALIZ-E to LC and ML.

REFERENCES

- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., et al. (2009). Cognitive developmental robotics: a survey. *IEEE Trans. Auton. Ment. Dev.* 1, 12–34. doi:10.1109/TAMD.2009.2021702
- Avila-Garcia, O., and Cañamero, L. (2005). “Hormonal modulation of perception in motivation-based action selection architectures,” in *Procs of the Symposium on Agents that Want and Like* (Hatfield: SSAISB), 9–17.
- Bahrack, L. E., Lickliter, R., and Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Curr. Dir. Psychol. Sci.* 13, 99–102. doi:10.1111/j.0963-7214.2004.00283.x
- Baillargeon, R. (1993). The object concept revisited: new directions in the investigation of infants physical knowledge. *Vis. Percept. Cogn. Infancy* 23, 265–315.
- Baillargeon, R., Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition* 20, 191–208. doi:10.1016/0010-0277(85)90008-3
- Baird, A. A., Kagan, J., Gaudette, T., Walz, K. A., Herschlag, N., and Boas, D. A. (2002). Frontal lobe activation during object permanence: data from near-infrared spectroscopy. *Neuroimage* 16, 1120–1126. doi:10.1006/nimg.2002.1170
- Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiol. Behav.* 81, 179–209. doi:10.1016/j.physbeh.2004.02.004
- Blanchard, A., and Cañamero, L. (2006). “Developing affect-modulated behaviors: stability, exploration, exploitation or imitation,” in *Proc. 6th Intl. Workshop on Epigenetic Robotics*. Lund University Cognitive Studies (LUCS) Vol. 128, 17–24.
- Breazeal, C. (1998). “A motivational system for regulating human-robot interaction,” in *The Fifteenth National Conference On Artificial Intelligence* (Madison: Aaai/iaai), 54–61.
- Cañamero, D. (1997). “Modeling motivations and emotions as a basis for intelligent behavior,” in *Proceedings of the First International Conference on Autonomous Agents* (Marina del Rey, CA: ACM), 148–155.
- Cañamero, L., Avila-Garcia, O., and Hafner, E. (2002). “First experiments relating behavior selection architectures to environmental complexity,” in *Procs IEEE/RSJ Int Conf on Intelligent Robots and Systems* (Lausanne: IEEE).
- Cangelosi, A., and Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: experiments with epigenetic robots. *Cogn. Sci.* 30, 673–689. doi:10.1207/s15516709cog0000_72
- Cangelosi, A., Schlesinger, M., and Smith, L. B. (2015). *Developmental Robotics: From Babies to Robots*. Cambridge, MA: MIT Press.
- Cos, I., Cañamero, L., and Hayes, G. M. (2010). Learning affordances of consummatory behaviors: motivation-driven adaptive perception. *Adapt. Behav.* 18, 285–314. doi:10.1177/1059712310375471
- Crews, D. (2008). Epigenetics and its implications for behavioral neuroendocrinology. *Front. Neuroendocrinol.* 29:344–357. doi:10.1016/j.yfrne.2008.01.003
- Crews, D. (2010). Epigenetics, brain, behavior, and the environment. *Hormones (Athens)* 9, 41–50. doi:10.14310/horm.2002.1251
- DeLoache, J. S. (2000). Dual representation and young children’s use of scale models. *Child Dev.* 71, 329–338. doi:10.1111/1467-8624.00148
- Di Paolo, E. A. (2003). “Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop,” in *Dynamical Systems Approach to Embodiment and Sociality, Advanced Knowledge International* (Adelaide: Citeseer), 19–42.
- Fischer, K. W. (1980). A theory of cognitive development: the control and construction of hierarchies of skills. *Psychol. Rev.* 87, 477. doi:10.1037/0033-295X.87.6.477
- Fuster, J. M. (2002). Frontal lobe and cognitive development. *J. Neurocytol.* 31, 373–385. doi:10.1023/A:1024190429920
- Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics* 1, 76–80. doi:10.4161/epi.1.2.2762
- Kagan, J., and Herschkowitz, N. (2006). *A Young Mind in a Growing Brain*. New York: Psychology Press.
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT press.
- Krichmar, J. L. (2008). The neuromodulatory system: a framework for survival and adaptive behavior in a challenging world. *Adapt. Behav.* 16, 385–399. doi:10.1177/1059712308095775
- Krichmar, J. L. (2013). A neurorobotic platform to test the influence of neuromodulatory signaling on anxious and curious behavior. *Front. Neurobot.* 7:1. doi:10.3389/fnbot.2013.00001
- Lones, J., and Cañamero, L. (2013). “Epigenetic Adaptation through Hormone Modulation in Autonomous Robots,” in *2013 IEEE 3rd Joint Int Conf on Development and Learning and Epigenetic Robotics*. Osaka: IEEE.
- Lones, J., Lewis, M., and Cañamero, L. (2014). “Hormonal modulation of development and behaviour permits a robot to adapt to novel interactions,” in *ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems*, Vol. 14. New York, 184–191.
- Marshall, J., Blank, D., and Meeden, L. (2004). “An emergent framework for self-motivation in developmental robotics,” in *Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)* (San Diego: Salk Institute), 10.
- Munakata, Y. (2000). Challenges to the violation-of-expectation paradigm: throwing the conceptual baby out with the perceptual processing bathwater? *Infancy* 1, 471–477. doi:10.1207/S15327078IN0104_7
- Oudeyer, P.-Y., and Smith, L. (2014). How evolution may work through curiosity-driven developmental process. *Topics Cogn. Sci.* 8, 492–502.
- Papalia, D. E., Olds, S. W., and Feldman, R. D. (1992). *Human Development*. New York: McGraw-Hill.
- Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: International University Press.
- Piaget, J., and Inhelder, B. (1969). *The Psychology of the Child*. New York: Basic Books.
- Schaffer, H. R. (1996). *Social Development*. Oxford: Blackwell Publishing.
- Schlichting, C. D., and Pigliucci, M. (1998). *Phenotypic Evolution: A Reaction Norm Perspective*. Sunderland, MA: Sinauer Associates Incorporated.
- Shaw, P., Law, J., and Lee, M. (2012). An evaluation of environmental constraints for biologically constrained development of gaze control on an icub robot. *Paladyn* 3, 147–155. doi:10.2478/s13230-013-0103-y
- Sirois, S., and Mareschal, D. (2002). Models of habituation in infancy. *Trends Cogn. Sci.* 6, 293–298. doi:10.1016/S1364-6613(02)01926-5
- Spelke, E. S., Breinlinger, K., Macomber, J., and Jacobson, K. (1992). Origins of knowledge. *Psychol. Rev.* 99, 605. doi:10.1037/0033-295X.99.4.605
- Timmis, J., Murray, L., and Neal, M. (2010). “A neural-endocrine architecture for foraging in swarm robotic systems,” in *NICSO 2010: Nature Inspired Cooperative Strategies for Optimization*, Vol. 284. eds J. R. Gonzalez, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor (Springer-Verlag Berlin), 319–330.
- Ugur, E., Nagai, Y., Sahin, E., and Oztop, E. (2015). Staged development of robot skills: behavior formation, affordance learning and imitation with motionese. *IEEE Trans Auton Ment Dev.* 7, 119–139.
- Zhang, X., and Ho, S.-M. (2011). Epigenetics meets endocrinology. *J. Mol. Endocrinol.* 46, R11–R32. doi:10.1677/JME-10-0053

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Lones, Lewis and Cañamero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.