Article

Interaction Histories and Short-Term Memory: Enactive Development of Turn-Taking Behaviours in a Childlike Humanoid Robot

Frank Broz [1,2,]*[iD]**, Chrystopher L. Nehaniv [1,3], Hatice Kose [1,4] and Kerstin Dautenhahn [1,3]**

[1] Centre for Computer Science & Informatics Research, University of Hertfordshire Adaptive Systems Research Group, Hatfield AL10 9AB, UK; chrystopher.nehaniv@uwaterloo.ca (C.L.N.); hatice.kose@itu.edu.tr (H.K.); kerstin.dautenhahn@uwaterloo.ca (K.D.)
[2] Department of Computer Science, Heriot-Watt University, Edinburgh EH14 4AS, UK
[3] Departments of Electrical and Computer Engineering and Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[4] Faculty of Computer and Informatics, Technical University of Istanbul, Istanbul 34469, Turkey
[*] Correspondence: f.broz@hw.ac.uk

Abstract: In this article, an enactive architecture is described that allows a humanoid robot to learn to compose simple actions into turn-taking behaviours while playing interaction games with a human partner. The robot's action choices are reinforced by social feedback from the human in the form of visual attention and measures of behavioural synchronisation. We demonstrate that the system can acquire and switch between behaviours learned through interaction based on social feedback from the human partner. The role of reinforcement based on a short-term memory of the interaction was experimentally investigated. Results indicate that feedback based only on the immediate experience was insufficient to learn longer, more complex turn-taking behaviours. Therefore, some history of the interaction must be considered in the acquisition of turn-taking, which can be efficiently handled through the use of short-term memory.

1. Introduction

Social interaction plays a vital role in a child's development. At an early age, children acquire basic social skills, such as the ability to read social gaze and play simple games that involve turn-taking; these serve as a scaffold for more sophisticated forms of socially-mediated learning. Children acquire these skills through interaction with their caretakers, motivated by intrinsic social drives that cause them to seek out social engagement as a fundamentally rewarding experience. Research on the innately social nature of child development has highlighted the primary role of intersubjective experience of an infant with a carer (e.g., Trevarthen [1,2]) and of contingency and communicative imitation (e.g., Nadel et al. [3] and Kaye [4]).

It is desirable to have robots learn to interact in a similar manner, both to gain insight into how social skills may develop and to achieve the goal of more natural human–robot interaction. This paper describes a system for the learning of behaviour sequences based on rewards arising from social cues, allowing a childlike humanoid robot to engage a human interaction partner in a social interaction game. This system builds on earlier behaviour learning work using interaction histories by Mirza and colleagues ([5,6]) as part of the RobotCub project [7] on the iCub humanoid developed by our consortium. This work also expands on prior work by the authors studying the timing of emergent

interactive behaviours, enabling a range of turn-taking behaviours to be acquired and coordinated by a single learning system [8].

Our work follows an enactive paradigm that has been identified as a promising framework for scaling behaviour-based approaches to fully learning and developing cognitive robotic systems (these concepts are expanded on in Section 2.4) [9]. Autonomy, embodiment, emergence, and experience play key roles via situated interaction (including social interaction) in the course of the ontogeny of cognition [10–12]. Temporal grounding via the remembering and reapplication of experience, as well as the social dimension of interaction, are the basis of the learning mechanisms in the interaction history architecture (IHA), which has been extended here. This work fits within the general approach of bottom–up ontogenetic robotics [13].

In this article, the results of a human–robot interaction experiment are presented that demonstrate how this enactive learning architecture supports the acquisition of turn-taking behaviours, how the inclusion of a short-term memory in the architecture supports the acquisition of more complex turn-taking behaviours, and how learned behaviours can be switched between based on social cues from the human partner.

The desire for social contact is a basic drive that motivates behaviour in humans from a very young age. Studies indicate that infants are able to recognise and are attracted to human faces [14]. They are especially attracted to the sound of the human voice as compared to other noises [15]. They are capable of recognising when they are the object of another's gaze and look longer at the faces of those looking at them [16]. They are also able to recognise contingent behaviour and prefer it to noncontingent behaviour [3]. Many developmental psychologists argue that from the very earliest stages of life, children begin to engage in simple interactions with aspects of turn-taking (e.g., imitating the facial gestures of others [17]), in intersubjective development [18], and in contingent interaction and communicative imitation [3,4,19]).

Infants thus seem to find many aspects of social interaction innately rewarding. This suggests the possibility of using these features to provide reward to a developmental learning system. Such a system would enable robot behaviour to be shaped through interaction in a way that is inspired by the infant–caretaker relationship and to engage humans in a manner that could assist in scaffolding the development of behaviour. Some possible sources of reward for such a system (based on the examples given above) could be:

- Presence of a person;
- Vocalisation/prosody;
- Visual attention;
- Contingency of another's actions;
- Synchronisation of both partners' actions.

This list is organised roughly in order from more immediate perceptual feedback to qualities of an interaction that appear to require some history of the relationship between of one's own and another's behaviour over time. In this study, feedback based on immediate perception and on a history of interaction was used to reinforce behaviours in the learning architecture. A simple computational model of short-term memory is introduced in order to transiently capture some of this recent history and facilitate the learning of social interaction behaviours.

2. Background

Research in both infant development and in robotics informed our approach to the developmental learning architecture and learning scenario presented in this paper. The scaffolding behaviours that were investigated in the context of this work on social development are gaze and turn-taking. The importance of short-term memory in social interaction was explored, and relevant research on the development of short-term memory in infants is presented. The use of childlike robots in the study of human–robot interaction and developmental learning is discussed. Finally, the enactive perspective on development that motivates our approach is discussed.

2.1. Social Behaviours

2.1.1. Gaze

Gaze is an especially powerful social cue. It is also one that becomes significant at an early developmental stage; even young infants are responsive to others' gaze direction [20]. Infants are able to follow eye movements from around 18 months of age [21]. The simplest response to a gaze cue is the recognition of having another's visual attention, and this recognition may be the basis of and developmental precursor to more complex gaze behaviours, such as joint attention [22]. This ability is crucial in a social context, as it provides valuable feedback about whether a nearby person is ready to interact or whether they are attending to something else. There is also some evidence that social gaze is inherently rewarding to people, e.g., one fMRI study found activation of the ventral striatum, an area of the brain involved in processing reward, during a joint attention task [23].

2.1.2. Turn-Taking Games

Turn-taking plays a fundamental role in regulating human–human social interaction and communication from an early age. The ability to engage in turn-taking games is a skill that children begin to develop early in life [24]. Turn-taking dynamics do not depend on the behaviour of a single individual but emerge from the interaction between partners. Turn-taking "proto-conversations" between infants and caretakers set the stage for more complex social interaction and later language learning [18]. Mothers control and shape early vocal play with infants into turn-taking interactions [25]. Carers appear to train infants how to engage in conversational turn-taking through interaction [26]. The regularity due to restriction on possible actions and repetition of simple social turn-taking games allow infants to learn to predict the order of events and increase their agency in the interaction by reversing roles during turn-taking [27].

Peek-a-boo is a simple game played with small children where an adult blocks the child's view to the adult's face with their hands and then lowers them to reveal their face while saying "peek-a-boo". This game has often been used as a scenario for the study of social learning in developmental psychology, for example, to investigate how infants acquire the "rules" and structure of interaction behaviour in interactions with parents [28] and to study infants' ability to recognise regular or irregular patterns of action in familiar interactions [29]. As children's mode of play transitions from passive to active during development, they come to understand the nature of the game, and it can take on a turn-taking form with the adult and child alternating who is the one hiding their face [30].

Drumming has also been used to study social interaction in older children. Children participating in social drumming tasks (where they drum with a present human partner rather than a disembodied beat) were more likely to adapt their drumming to that of their interaction partner [31]. The ability to synchronise with a partner while drumming is age-dependent, improving as children approach adulthood [32]. Drums and other percussion instruments are commonly used in music therapy for children with autism that seeks to improve social skills [33]; and, moreover, music therapy that includes drum play can lead to improvements in joint attention and turn-taking behaviours in children with autism [34].

2.2. Short-Term Memory

There is increasing experimental evidence that people rely on short-term memory in order to evaluate others' social behaviour. Decoding nonverbal social cues requires the use of short-term memory [35]. The prefrontal cortex is widely believed to be involved in storing and processing short-term memory [36]. The prefrontal cortex's response to human faces (as opposed to that of other parts of the brain) is focused on the eyes and may be involved in gaze processing [37]. Functional magnetic resonance imaging (fMRI) studies suggest that the prefrontal cortex is involved in making judgements about social gaze related to gaze duration [38]. The prefrontal cortex has also been observed to be involved in the processing of temporal information in short-term memory tasks in

primates, such as the recognition of the frequency pattern of a stimulus [39] or remembering the temporal ordering of events for an action [40].

In child development, researchers are in agreement that infants are born with limited working memory capacity that significantly improves by the latter half of their first year. Most experimental research on short-term memory in infants focuses on the identification of objects, which makes it somewhat difficult to find direct implications for open-ended, movement-based play interactions (this can be seen in the prevalent focus of research studies into infant memory, e.g., [41]). However, some researchers approach the measurement of short-term memory in terms of time rather than symbolic chunks [42,43]. Studies of working memory in infants indicate that they can perform short-term memory based tasks requiring a memory of up to 10 s by around 9 to 12 months of age [44,45].

As an example of how short-term memory is modelled by cognitive scientists in humans, consider Baddley's model of working memory, which includes an episodic buffer of auditory and visiospatial information (to reflect the fact that working memory appears to form associations across different sensory perceptions and make use of this associated information) [46]. Baddeley's episodic buffer operates on multimodal input and preserves temporal information, allowing the analysis of the sequential history of recent experience. Two further submodules in his proposed model of working memory deal with the specialised processing of auditory and visual and spatial information. A central executive manages switching between tasks and selective attention, as well as coordinating the transfer of information between the other parts of the system. The form that our short-term memory (STM) module takes (described in Section 4.3), operating over a temporal sequence of past sensorimotor data, is similar to Baddeley's model in that it treats short-term memory as multimodal rather than keeping data streams separate. However, in keeping with our enactive approach, in our model of memory there is no central executive controlling information flow or representation of tasks to be switched between.

2.3. Social and Developmental Robotics

Developmental robotics is a growing area of robotics in which approaches to artificial cognition take inspiration from human development (see Cangelosi and Schlesinger [47] for a recent overview, and also Tani [13]). The use of a childlike robot appearance to encourage people to interact in a way that mimics the scaffolding of behaviour that supports human learning is an approach that has been adopted by a number of researchers in this field. The robot Affetto was designed to be infant-like in order to study the role of the attachment relationship that forms between infants and their caretakers in child development [48]. The CB2 robot was designed to learn social skills related to joint attention, communication, and empathy from humans [49]. The iCub robot, the robot used in this experiment, is another childlike robot designed as a platform for research in cognitive systems and developmental robotics, including the acquisition of social competencies learned from humans [50]. Other social robots designed for human–robot interaction, such as Simon [51], Nexi [52], Robovie [53], and the Nao [54] were designed to be childlike in either size, facial appearance, or both. However, for these robots, the design decision was made primarily to make the robot appealing to interact with rather than specifically to support developmental approaches to learning through interaction. Prior to availability of the iCub, the forerunner of the present architecture [6] was implemented on KASPAR, a minimally expressive, childlike humanoid developed in our lab at the University of Hertfordshire for therapeutic interaction with children with autism [55].

The iCub is a 53 degree-of-freedom (dof), toddler-sized humanoid robot. It has a color camera embedded in the centre of each eye. Its facial expression can be changed by turning on and off arrays of LEDs underneath the translucent cover of its face to create different shapes for the mouth and eyebrows. The iCub's control software is open-source, and the source code is available for download as part of the iCub software repository (https://github.com/robotology). The modules that make up its control system communicate using the YARP middleware for interprocess communication [56].

Roboticists such as Breazeal and Scassellati [57,58] and child development researchers such as Meltzoff [59] have proposed using robots to study the development of social cognition with a

focus on the development of theory of mind and its role in imitation learning. In our work, we are exploring other mechanisms that can enable social development. An infant does not need to understand the motivations or beliefs of another person to derive enjoyment from their presence and contingent interaction. Rather than focusing on intentionality, this system focuses on the coordination of behaviour and social feedback as mechanisms for scaffolding turn-taking interactions.

2.4. Enactive Development: Ontogeny without Representation

This work can be viewed as compatible with an enactive approach to cognitive systems that has been suggested and described by philosophers such as Varela, Hutto, and Noë and used as a guiding principle for the implementation of working systems by artificial intelligence researchers such as Vernon, Dautenhahn, and Nehaniv [5,6,10–12,60–64]. The enactive approach's focus on behaviour and interaction makes it a promising framework for the development of cognitive systems. For a system to acquire and select behaviours in interaction, there is no need to assume the existence of static entities with properties characterisable by truth-functional propositional logic and semantics. This approach to interactive behaviour is inspired by Brooks in that there is no model of the world, nor of the self, nor entities in it ("The world is its own best model") [65]. However, unlike classical behaviour-based systems, our system learns new behaviours to select from. Moreover, as in child development, learning occurs continuously and is not artificially separated into discrete 'training' and 'testing' phases.

In an enactive cognitive system, the ontogeny of cognitive capacities depends on the processes of sensorimotor and social interaction and on the capacity of the cognitive architecture to allow behaviour to be shaped by experience (operationally defined as the temporally extended flow of values over sensorimotor and internal variables) and feedback (operationally defined as measures of social engagement), rather than on modelling, symbols or logics. Different sequences of actions are acquired as repeatable behaviours through interaction. These behavioural trajectories can be viewed as "paths laid down in walking" [10] in a space of possibilities for social interaction that shape habits and scaffold future development. This type of interactive history architecture exhibits the capacity for anticipatory prospection [66] and constructing its own behaviours in interactive development through recurrent structural coupling with its environment [6].

3. Related Work

The study of the emergence of turn-taking has vital implications for many areas of human–robot interaction (HRI), such as robot-assisted therapy. This is especially the case for autism therapy, where turn-taking games have been used to engage children in social interaction [67]. Because of the developmental approach of this work (inspired by interaction with preverbal children), we will discuss only HRI research on movement-based (rather than conversational) turn-taking interactions. In many turn-taking studies, the focus has been on evaluating turn-taking quality rather than learning how to take turns from the human. Delays and gaps in a robot's response timing can lead to negative evaluations by human participants [68]. An imitation game with the humanoid robot QRIO was used to study joint attention and turn-taking [69]. In this game, the human participants tried to find the action patterns (previously learned by QRIO) by moving synchronously with the robot. The robot KISMET used social cues for regulating turn-taking in nonverbal interactions with people [70]. Turn-taking between KISMET and humans emerged from the robot's internal drives and its perceptions of cues from its interaction partner, but the robot did not learn behaviours. In a wizard of Oz study of a touchscreen-based task, turn-taking was found to emerge as children adapted to a robot's behaviour, even though the robot's actions were not contingent on the child's [71].

In a data collection study by Chao and colleagues, humans played a turn-taking imitation game with a humanoid robot in order to investigate the qualities of successful turn-taking in multimodal interaction [51]. Based on these findings, Chao developed a system for turn-taking control (including interruption of behaviours) using timed petri-nets [72]. This approach is not developmentally inspired

and involves a complex human–robot collaboration task (assumed to be known) for which robot behaviour is coordinated with the human's.

Peek-a-boo has been used previously as the interaction scenario to demonstrate developmentally motivated learning algorithms for robots. Work by Mirza et al. first demonstrated a robot learning to play peek-a-boo through interaction with a human using an earlier version of the interaction history architecture (to be described in the next section) [6]. Ogino et al. proposed and implemented a developmental learning system that used reward to select when to transfer information from an agent's short-term memory to long-term memory. The test scenario used to demonstrate this learning system was a task of recognising correct or incorrect performance of peek-a-boo based on an experiment with human infants [73]. In both of these studies, peek-a-boo was the main target behaviour to be acquired, so the systems did not demonstrate the learning of more than one interaction task within a single scenario.

Drumming has also been used as an interactive behaviour to study turn-taking. Kose-Bagci and colleagues studied emergent role-switching during turn-taking using a drumming interaction with a childlike humanoid robot that compared various rules for initiating and ending turns [74]. The aim of this research was not to produce psychologically plausible models of human turn-taking behaviour but to employ simple, minimal generative mechanisms to create different robotic turn-taking strategies based on social cues. The robot dynamically adjusted its turn length based on the human's last turn but did not learn to sequence the drumming actions themselves. A system designed for human–robot musical collaboration implemented a leader–follower turn-taking model for drumming with a human [75]. This robot was designed for interaction with adult musicians and used concepts from musical performance to inform its turn-taking model.

Kuriyama and colleagues have conducted research on robot learning of social games through interaction [76]. They focus on different aspects of interaction learning than the work described in this article, using causality analysis to learn interaction "rules" by discovering correlations in the low-level sensorimotor data stream. Our work differs in that it makes use of social cues, such as human gaze directed at the robot, that are fundamental to human interaction and representative of the perceptual capabilities of a human infant (though they may appear more high-level from a data-processing standpoint). Further, the social games we seek to learn in this study are known to the human interaction partner. The robot has only very minimal social and interactive motivations for what constitutes success in its behaviour in the form of reward for social engagement. The focus of the learning is on sequencing simple actions to produce behaviour in response to human social feedback and on coordinating the robot's behaviour with the human's in the context of the interaction. This could be used for acquiring various behaviours, including in principle ones not tested in our experiments.

There are a number of cognitive architectures for robots that are similar to the extended interaction history architecture (EIHA) in that they are also designed to allow bottom–up learning through enactive experience. For example, the HAMMER architecture learns associations between motor commands and sensor data and can use the models obtained from self-exploration to imitate human gestures [77]. Wieser and Cheng proposed a developmental architecture for the acquisition of sensory-motor behaviours through multiple stages of learning and verifying predictors through both mental simulation and physical trials [78]. These architectures have not been applied to learning interactive turn-taking behaviours, and their development and evaluation focuses more on self-exploration to learn to predictably interact with the environment than on socially scaffolded learning from humans.

The Interaction History Architecture

This research extends past work on the iCub using the interaction history architecture (IHA). To avoid confusion, the original system will be referred to as IHA and the modified system from this article will be referred to as the extended interaction history architecture (EIHA). EIHA differs from

the previous version in that the types of sensor input available to the robot have been expanded and a short-term memory module has been added to the architecture.

IHA is a system for learning behaviour sequences for interaction based on grounded sensorimotor histories. While the robot acts, it builds up a memory of past "experiences" (distributions of sensors, encoders, and internal variables based on a short-term temporal window). Each experience is associated with the action the robot was executing when it occurred, as well as a reward value based on properties of the experience. These experiences are organised for the purpose of recall using Crutchfield's information metric as a distance measure [79]. The information metric (which the authors will also refer to as the information distance) is the sum of the conditional entropies of two random variables, each conditioned on the other. Unlike other popular ways of comparing distributions (such as mutual information or Kullback–Leibler divergence), it meets the conditions necessary to define a metric space. Given two random variables, the information distance between them is defined as the sum of the conditional entropy of each variable conditioned on the other.

$$d(X,Y) = H(X|Y) + H(Y|X) \tag{1}$$

In IHA, the random variables are sensor inputs or actuator readings whose values are measured over a fixed window of time (defined as a parameter of the system). These values define a distribution for each sensor or actuator that is assumed to be stationary over the duration of time defined by the window. This vector of variables provides an operationalisation of temporally extended experience for the robot. As the robot acts in the world, the informational profile (probability distribution) of this sensorimotor experience (including internal variables) and reward is collected as an 'experience' in a growing body of such experiences that is stored within a dynamically changing metric space. The pairwise sum of the information distances between the variables making up the current experience vector and a past experience vector are calculated and used as a distance metric. As the robot acts, the most similar past sensorimotor experience to its current state is found, and new actions are probabilistically selected based on their reward value (which is dynamically changing based on the interaction). For a full description of the architecture, see prior publications by Mirza et al. [5,6].

This architecture is inspired by earlier work by Dautenhahn and Christaller proposing cognitive architectures for embodied social agents in which remembering and behaviour are tightly intertwined via internal structural changes and agent–environment interaction dynamics [11]. We view both the embodiment of the system and its embedding in a social context to be critical aspects of the system's cognition. Our focus on learning through social interactions rather than self-exploration is based on the hypothesis that social interaction was the catalyst to the development of human intelligence and is therefore a promising enabler of the development of robot intelligence [80].

While the original version of IHA implements a model of remembering in which sensorimotor experiential data may be recalled in a way that allows experiences to be compared to one another, it lacks certain characteristics that are useful, possibly even necessary for learning about social interaction. IHA makes no use of how recently experiences have occurred. Further, while experiences are themselves temporally extended, it is only over a short duration temporal horizon, and only single experiences rather than sequences of experiences are compared.

Additionally, in IHA, all rewards are calculated based solely on the current sensor data. For many behaviours, especially interactive behaviours, this is not sufficient to determine whether the phenomena that should be reinforced are occurring unless the temporal horizon of experiences has been chosen to match that behavioir [81]. In many cases, it is necessary to examine a sequential history of the sensor data for relationships between an agent's actions and those of the agent with which it interacts. We hypothesise that fluid turn-taking requires attention to the recent history of both one's own and the other's actions in order to anticipate and prepare for the shift in roles. In light of this, EIHA incorporates a short-term memory over the recent history of sensor data relevant to the regulation of turn-taking (to be described in Section 4.3). A diagram of how this module fits into the architecture can be seen in Figure 1.
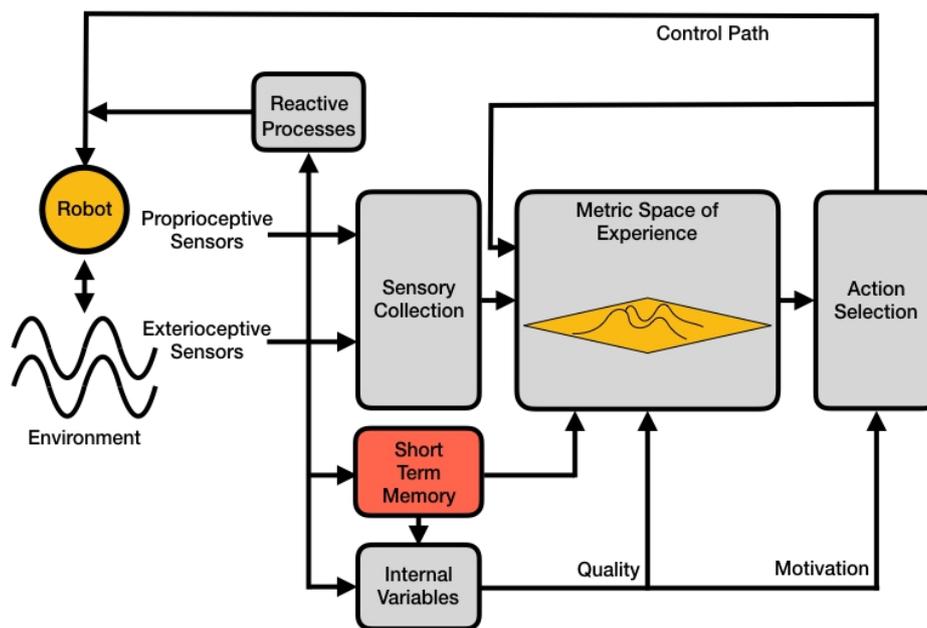
Figure 1. Diagram of the extended interaction history architecture (EIHA), showing the added short-term memory module.

4. Architecture Design and Implementation

EIHA is intended to support the robot developing different socially communicative, scaffolded behaviours in the course of temporally extended social interactions with humans by making use of social drives and its own firsthand experience of sensorimotor flow during social interaction dynamics. The robot comes to associate sequences of simple actions, such as waving an arm or hitting the drum, executed under certain conditions with successful interaction based on its past experience.

In order to demonstrate these concepts, rewards based on social drives are designed to influence the development of behaviour in an open-ended face-to-face interaction game between the iCub and a human. The human interaction partner interacts with the robot and may provide it with positive social feedback by directing their face and gaze toward the robot, as well as by engaging in drum play. The rewards reflecting these social drives for human attention and synchronised turn-taking may be based on the current state of the robot's sensorimotor experience or on the recent history of experience transiently kept in its short-term memory. The robot uses this feedback to acquire behaviour that leads to sustained interaction with the human.

The system described in this article focuses on the use of visual attention as a form of social feedback, and on the learning of turn-taking behaviour to explore issues of contingency and synchronisation in interaction. Two forms of nonverbal turn-taking are specified as targets to be learned by the iCub through interaction with a human, peek-a-boo and drumming. A diagram of the implementation of EIHA used for learning the social interaction game is shown in Figure 2.

In the rest of this section, the data used as input to the learning system and the actions available to the robot for this scenario are described. The role of the short-term memory module in assigning reward and the design of the rewards based on immediate sensor input and short-term memory are explained. Finally, a description of how this learning architecture supports switching between learned behaviours during interaction is given.
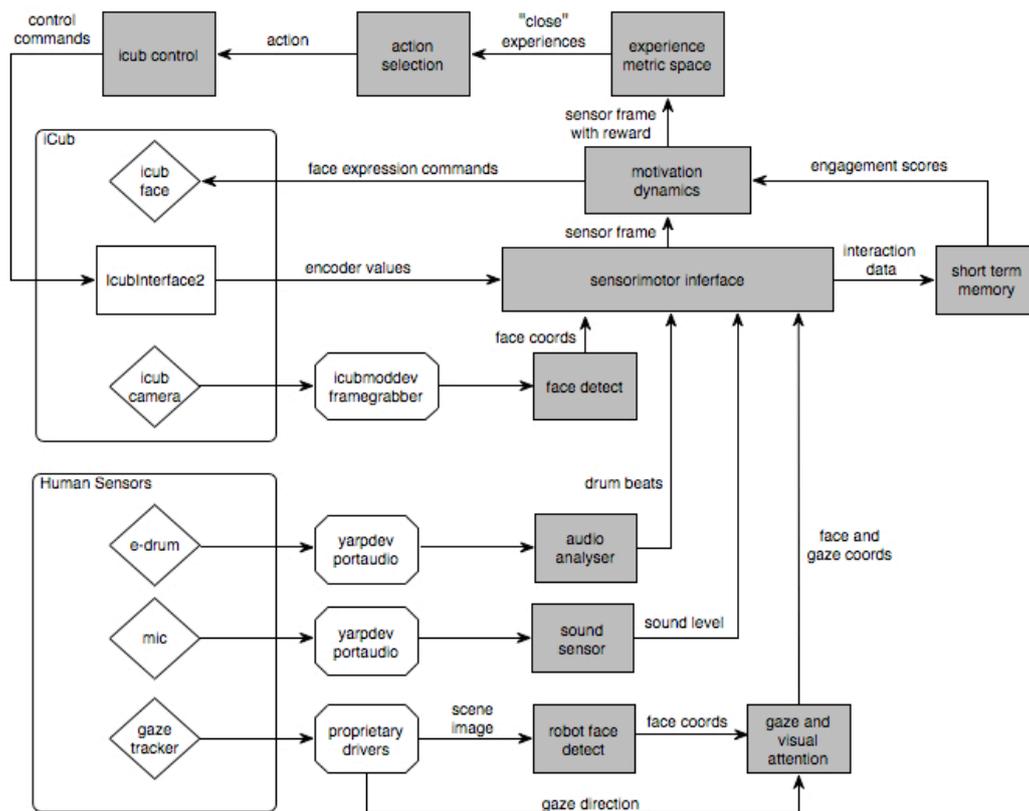
Figure 2. Diagram of the extended interaction history architecture (EIHA), sensor input, and robot controller.

4.1. Sensorimotor Data

The sensorimotor data whose joint distribution over temporally extended intervals make up an experience are a collection of variables describing the robot's own joint positions and both raw and processed sensor input. Continuous-valued data were discretised using 8 equally-sized bins over the range of each variable. The raw camera image was captured as a very low resolution (64 pixel) intensity image. High-level information from processed sensor data was included as binary variables reporting the presence of a face, the detection of a drum hit by the human drummer, and whether the human is currently looking at the robot's face. A full list of data variables is given in Table 1.

The system takes sensor input from the robot's eye camera, an external mic (this input source was not used for the experiment described in this article), an electronic drum, and an Applied Science Laboratories MobileEye gaze-tracking system (see Figure 3) [82]. The wearable gaze tracking system included a head-mounted camera that records the scene in front of the person. The system outputs the person's gaze target within this scene image in pixel coordinates.

Certain sensor input was processed to provide high-level information about the person's current activities. Face detection was used to locate the human interaction partner's face in the camera image from the robot's eye, and a binary variable reported whether or not a person was currently visible to the robot (face detection in Table 1). Face detection was also applied to the gaze tracker's scene camera image to locate the robot's face, compare the location to the person's gaze direction, and report whether the person was looking at the robot as a binary variable (visual attention score). The Haar wavelet based face detection algorithm implementation in OpenCV was used. The audio stream from the electronic drum was filtered to extract drumbeats using the method implemented for the drum-mate system described in the appendix of Kose-Bagci et al. [74]. Note also that variables referring to the "engagement score" for the drumming and hide-and-seek interactions are also a part of the robot's experience. Their inclusion links the immediate experience of the robot to its recent history of observed

behaviour as interpreted by the short-term memory module. This relationship will be further explained in Section 4.3.

Figure 3. The human interaction partner's visual attention toward the robot is measured using a gaze tracking system.

Table 1. Data variables that make up an experience in the experience metric space.

Data	# of Variables
Proprioceptive Data	
head position	3
eyes position	3
left arm position	7
left hand position	9
right arm position	7
right hand position	9
Exterioceptive Data	
left eye intensity image	64
face detection	1
beat detection	1
Internal Variables	
visual attention score	1
Internal Variables (STM-based)	
drum engagement score	1
hide engagement score	1
Total	107

4.2. Actions and Behaviours

The robot's actions are preprogrammed sequences of joint positions and velocities that make up a recognisable unit of motion. Ten actions plus a short-duration no-op action were available to the robot, listed in Table 2 (note that actions which have the same end point or are reflections of one another are represented by one image). In the context of interaction, the robot discovers the reward values of actions for a particular experience during execution. EIHA probabilistically chooses actions associated

with high reward. Over time, the robot learns behaviour sequences by experiencing sequences of high reward actions and associating them with a sequence of experiences, and selecting those actions again when those remembered experiences are similar to its current experience. It is important to note that while low-level actions may have some recognisable "purpose" to the human interactor, the behaviours that are to be learned are made up of sequences of these actions. For example, the peek-a-boo behaviour is made up of some sequence containing "hide-face" followed by "home-position", and a single turn of drumming is made up of "start-drum", an arbitrary number of "drum-hit" actions, and then "right-arm-down" or "home-position". These sequences may also have an arbitrary number of "no-op" actions interspersed throughout them, affecting their timing.

Table 2. The 10 actions that the robot may select among during execution, with corresponding image of the ending pose shown. There is also an 11th no-op action (no image shown).

right-arm-down, left-arm-down, home	hide-face	right-arm-up, left-arm-up	right-arm-wave, left-arm-wave	start-drum	drum-hit

4.3. Short-Term Memory Module

In addition to dynamic remembering of sensorimotor experience and associated rewards, it is also useful to have a more detailed, fully sequential memory of very recent experience. This is especially true for skills such as turn-taking, where the recent history of relationships between one's own and another's actions must be attended and responded to. In IHA, while the experience metric space preserves the ordering of experiences (so that rewards over future horizons may be computed), there is not a mechanism to recall the most recent experience, only the most similar. Additionally, experiences aggregate data over a window of time, eliminating potentially useful fine-grained information about changes in sensor values. The short-term memory preserves temporal information within selected sensorimotor data over a fixed timespan of the recent history. Because the short-term memory captures relevant changes in sensor data in the recent history of experience and an experience accumulates sensor data, it is reasonable for the length of the short-term memory to be of a longer duration than the experience length, probably several times its length. The effect of the relationship between experience length and the duration of the short-term memory will be experimentally investigated in Section 6.

This additional temporal extension of data directly influencing action selection is especially important for guiding social interactions as it allows rewards to be assigned based on these histories of interaction, rather than just the instantaneous state of the interaction that the robot is currently experiencing. History-based engagement scores are included as variables in the vector that defines the experience space. Multiple engagement scores may be defined to process specific relationships between data within the memory module. These scores are computed over the duration of a sliding window of past data values (length defined as a parameter of the system). This places a performance measure of the recent history of the robot's behaviour within an experience, making experiences involving similar immediate contexts that are part of either successful or unsuccessful interaction distinguishable to the system.

4.4. Reward

Of the possible sources of reward arising from internal motivations discussed in Section 2.1, three were selected for the interaction experiment described in this article. The reward used by the system was the sum of engagement scores related to the various forms of social feedback used. One, visual

attention from the interaction partner, was based on immediate perceptual information. The other two sources of motivation were based on engagement and behaviour coordination, and the short-term memory module was used to compute their scores.

The two turn-taking behaviours assessed were peek-a-boo and drumming. Though engagement was calculated differently for each turn-taking behaviour, the criteria for determining the quality of interaction were the same. In both cases, the criteria used to compute the engagement score were

- Reciprocity—*Are both the robot and the human playing this game?*
- Synchronisation—*Is each of them waiting for the other to take their turn?*

The details of how the engagement scores were computed for each source are described below.

4.4.1. Visual Attention Score

Visual attention was detected using the gaze direction coordinates reported by the gaze tracker and the result of running the face detection algorithm on the gaze tracker's scene camera. The bounding box for the robot's face was found in the image. If the human's gaze direction (also reported in pixel coordinates of the gaze tracker's scene camera) fell within this bounding box, visual attention was set to a value of one for that timestep. Otherwise, it was set to zero.

4.4.2. Engagement Scores

Both of the turn-taking games use knowledge of the *characteristic action* for the game to compare the behaviour of the human and the robot. The characteristic action is the action in a sequence making up a behaviour that is necessary in order for the behaviour to be recognised as being related to the game. For peek-a-boo, the characteristic action is "hide-face" and for drumming, it is "hit-drum". Both games also computed their engagement scores over the sensor history of the short-term memory module, whose duration was a run-time parameter of the EIHA system. The details of the engagement score function for each game are described below. Pseudocode for the computation of the scores is shown in Figure 4.

For peek-a-boo, at each timestep, it was determined whether the human or the robot were hiding. These variables were kept in the short-term memory. Positive engagement was given when the human had been hiding for a certain length of time (between 0.5 and 2.5 s) while the robot is not executing the hide action. This duration of time represents the expected length of time that a person would hide their face during peek-a-boo. Shorter losses of the face are likely caused by transient failures in face detection. Longer losses are likely to be caused by a person being absent or having their face (and therefore attention) directed away from the robot. When the human's hiding action fell within this range of time, the reward was computed as the normalised sum of the duration of robot and the human's hide actions over the memory length. The case where the human is hiding and the robot is not was used to assign an engagement value because in the reverse scenario (robot hiding and human not hiding), the robot partially obscures its view of the interaction partner with its hands during its hide action, causing face detection failures that make it difficult to detect whether or not the human is actually hiding.

For the drumming behaviour, the variables of interest were whether the robot and human were drumming and whether they were both drumming at the same timestep. The robot was determined to be drumming only when it was currently selecting the characteristic action (so not while it was starting or ending a turn of drumming). The human was determined to be drumming when the audio analyser had detected a beat in the e-drum input within the last timestep. If the robot chose the hide action while the human was drumming, it received a fixed size negative engagement score reflecting that in this context, that action appeared inattentive to the human. Further, the robot could only receive positive engagement for drumming when the human had been drumming within the history of the short-term memory's length. This constraint limits the engagement score to situations where the human was also engaging in the behaviour, rather than allowing the robot to increase its score

simply by drumming at any time. The engagement score is a normalised sum of the history of the human and robot's drumming, penalised by the sum of the timesteps when both were drumming at the same time rather than synchronising their turn-taking in alternation. Note that, unlike peek-a-boo, engagement may be negative. In this game, poor synchronisation can be detected and penalised (unlike in peek-a-boo, where occlusion of the vision system during hiding may limit perception of the human's actions).
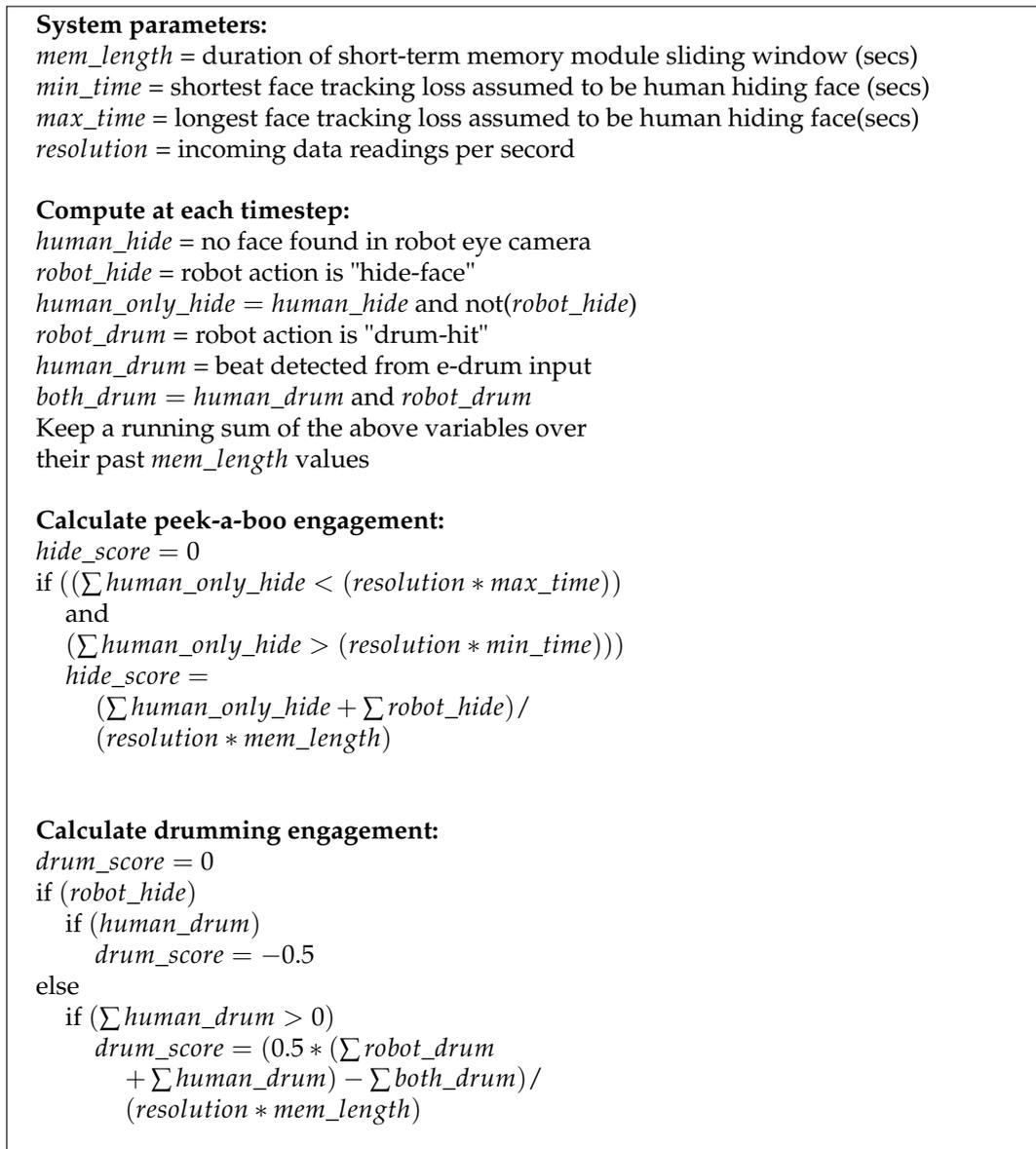

System parameters:
mem_length = duration of short-term memory module sliding window (secs)
min_time = shortest face tracking loss assumed to be human hiding face (secs)
max_time = longest face tracking loss assumed to be human hiding face(secs)
resolution = incoming data readings per secord

Compute at each timestep:
human_hide = no face found in robot eye camera
robot_hide = robot action is "hide-face"
$human\_only\_hide = human\_hide$ and $not(robot\_hide)$
robot_drum = robot action is "drum-hit"
human_drum = beat detected from e-drum input
$both\_drum = human\_drum$ and $robot\_drum$
Keep a running sum of the above variables over
their past *mem_length* values

Calculate peek-a-boo engagement:
$hide\_score = 0$
if $((\sum human\_only\_hide < (resolution * max\_time))$
   and
   $(\sum human\_only\_hide > (resolution * min\_time)))$
   $hide\_score =$
     $(\sum human\_only\_hide + \sum robot\_hide)/$
     $(resolution * mem\_length)$

Calculate drumming engagement:
$drum\_score = 0$
if $(robot\_hide)$
   if $(human\_drum)$
     $drum\_score = -0.5$
else
   if $(\sum human\_drum > 0)$
     $drum\_score = (0.5 * (\sum robot\_drum$
       $+ \sum human\_drum) - \sum both\_drum)/$
     $(resolution * mem\_length)$

Figure 4. Definitions of the memory-based engagement scores for peek-a-boo and drumming.

4.5. Switching between Behaviours

One important feature of this system is the ability to switch between the learned behaviours.[1] Both interactions (drumming and peek-a-boo) can be learned during a single episode of extended

---

1   For an example of the robot switching between learned behaviours, see the video documentation of the system at http://wiki.icub.org/misc/icubvideos/ihaNew_short2_web.mov. In this video, there are also some examples of the robot initiating attempts at switching between interaction games. In particular, the robot need not necessarily always play the role of follower but can start an interaction game of a particular kind itself.

interaction with the human teacher. Once learned, the robot can change between the two learned behaviours to maximise their reward based on social feedback from the human. For example, if the robot were playing peek-a-boo with the human and the human quit playing and attending to the robot, the robot would eventually stop repeating the peek-a-boo behaviour and start to explore new action sequences, including possibly initiating the drumming behaviour. If the human responded by drumming back, the robot would continue the drumming interaction.

Behaviour switching is achieved in the architecture as follows. At first, when the human quits playing, the robot's most similar experiences from the experience metric space are experiences from the history of reciprocal, high-reward turn-taking interaction with the human. As the human fails to respond, either new experiences of lower reward for the hide-face action (and thus, the peek-a-boo sequence) are created from the recent sensorimotor data or the reward associated with that action is revised downward for the recalled experience. Note that whether experiences are added to the experience space or existing experiences are dynamically revised in acts of remembering depends on current sensor values and on system parameters of EIHA. Due to the lower reward for the characteristic action of peek-a-boo, the probabilistic action selection rule is more likely to explore other actions. Note that any action selected (such as waving) can be reinforced using visual attention, so it is possible for the robot to be taught behaviours during interaction other than drumming and peek-a-boo. When the "start drum" action is selected, the robot will recall experiences from its recent drumming interaction with the human, making it more probable that high-reward actions associated with this history of interaction will be chosen. If the human responds to the drumming by drumming in turn, the learned drumming behaviour is further reinforced.

5. Experiment

For artificially intelligent humanoid enactive robotics to demonstrate for the first time an ontogenetic capacity to both acquire behaviours and adaptively to switch between them in a socially situated manner, we designed and conducted a human–robot interaction experiment to illustrate the impact of short-term memory on the learning of turn-taking behaviours by IHA. Our hypothesis is that the inclusion of short-term memory would aid in the learning of turn-taking behaviour and that it would be necessary in order to learn more complex turn-taking behaviours (complexity here defined as both the number of actions that make up the behaviour and the amount of variation that can occur within successful turn-taking interactions).

In order to evaluate the system, we investigated whether and under what conditions EIHA can support the learning of (and subsequent engagement in) turn-taking, demonstrated using the selected activities of drumming and peek-a-boo. The experimental setup was that of a face-to-face interaction with a human teacher/playmate. The human sat approximately 1 m away from the iCub with a small table between them, upon which was placed the e-drum that the human used for drumming and the toy drum that the iCub used. An example of the setup is shown in Figure 5. During the experiment, the experimenter wore a wearable gaze tracker like the one shown in Figure 3.

The impact of the short-term memory module's addition in EIHA was experimentally investigated by allowing the robot to play the game with a human for multiple trials and examining the outcome in terms of which behaviours were successfully learned. Three conditions were compared: No STM, short-term memory (STM), and truncated STM. In the no STM condition, the short-term memory module is not used by the system. The only source of reinforcement in this condition was visual attention through gaze because the other sources of reward rely on engagement scores calculated using information from the short-term memory module. Note that in this condition, EIHA operates in the same manner as IHA, with the only difference between the systems being the extended sensor input in this implementation. In the STM condition, the robot received reinforcement from both visual attention and the engagement scores from the short-term memory module. In this condition, the duration of the interaction history used by the short-term memory module was 4 s, a length of time chosen empirically based on the duration of the target turn-taking behaviours. This length of time should typically capture

at least one transition between the robot's and human's turn during well-coordinated turn-taking for both of the behaviours to be learned. In the truncated STM condition, the short-term memory module computed engagement scores based only on a 1 s history of interaction data, a length of time expected not to be long enough to be able to capture the turn-taking dynamics of both behaviours. The experience length used by this system was 2 s in duration at a 10 Hz sensor rate. The parameter settings for EIHA were taken from the earlier implementation of peek-a-boo learning on the iCub with the original version of IHA. The robot performed 5 game trials per condition. Each trial lasted until either both turn-taking behaviours had been learned successfully (the criteria by which this was determined are described in the next section) or 10 min had elapsed.
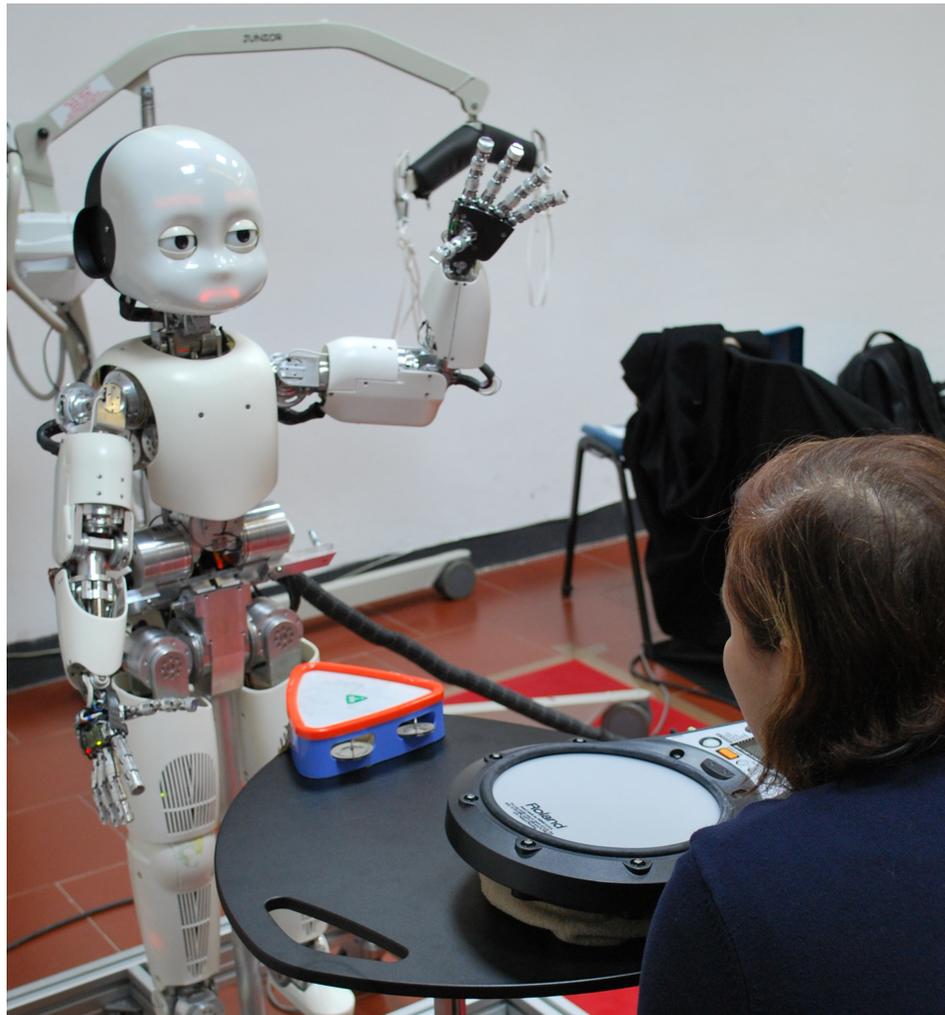
Figure 5. The iCub interacts with the human interaction partner using gestures and drumming.

All of the trials were performed with a single trained experimenter acting as the teacher/interaction partner. The experiment was conducted in this manner (rather than having the robot interact with several different users) in order to focus on differences due to the complexity of the interaction games rather than individual differences in interaction styles. Further, the use of an experienced interaction partner eliminated the possibility of training effects over the course of the trials as the teacher adapts to the system. The focus of this experiment is on understanding how the properties of this cognitive system influence its learning, not on understanding how a range of users might interact with the system. Using a skilled experimenter is appropriate to assess and demonstrate the capabilities of the system. In particular, its exhibiting particular behaviours with this interaction partner shows that it is possible for the system to achieve them in principle, and its failure

to learn certain behaviours under particular settings even with this partner suggests that is unlikely to achieve them with more naïve interaction partners. This allows us to argue that certain components of the architecture are either sufficient or necessary to achieve certain behaviours using this enactive cognitive architecture.

6. Results

6.1. Criterion for Learning Success

There were two criteria used to evaluate the learning performance of the different models. First, whether a model could support the learning of a behaviour within the 10 min trial length was considered. In cases where different models were capable of learning a behaviour, differences in performance were evaluated in terms of the amount of time it took for the system to learn the behaviour, measured from the first time the robot was observed to perform the *characteristic action* for a behaviour.

In order to measure success at learning turn-taking behaviours, we needed to decide on a criterion for success that could be observed directly from the interaction. This criterion was arrived at empirically based on the behaviour observed during free-form training of the system. The intuition for the criterion was that it should indicate that a turn-taking behaviour could be continued by the human partner continuing reinforcement and that, should the human cease reinforcement and change to the other turn-taking behaviour, the originally learned behaviour could be returned to later (i.e., the learned behaviour was not lost by learning the other turn-taking behaviour).The criterion used to determine whether a turn-taking behaviour was learned is the occurrence of three consecutive robot-human turns of that behaviour. The "consecutive" requirement means that if the robot introduces additional random actions that are not part of the desired behaviour sequence into a turn-taking interaction, turn-taking is considered to not have been learned successfully, and previous turns are not counted. Turn-taking was evaluated by manually coding video recorded of the experimental interactions with the robot. This method was used rather than relying on data logged from the robot in order to ensure that the human's turn-taking actions were correctly identified.

6.2. Incidence of Learning Success

The difference in the capability of the models to learn the behaviours within the given trial time was compared. The percentages of the trials that resulted in the successful learning of each behaviour are shown in Figure 6. It can be seen that all of the models successfully learned the peek-a-boo behaviour during almost all of the trials. The differences in learning success for the drumming behaviour were greater. None of the models successfully learned drumming in every trial, and the no STM model did not learn it during any trial. Fischer's exact test was used to evaluate the significance of the differences between the proportions of successes and failures of the STM model and the other models (see Table 3). Based on the number of successes versus failures alone, the no STM model's results showed a statistically significant difference for the drumming interaction (p-value = 0.02). This is a strong indication that drumming was the more difficult to learn of the two forms of turn-taking, and that the short-term memory module played a role in the system's ability to learn it.
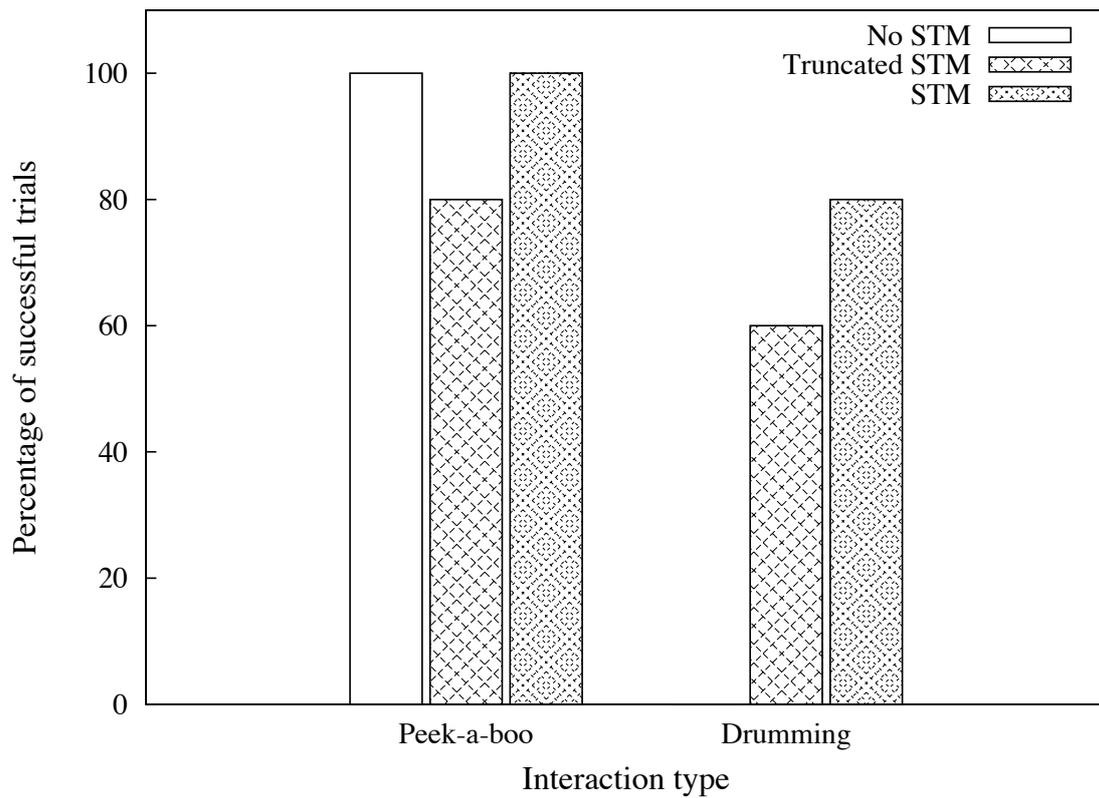
Figure 6. The percentage of trials which resulted in the successful learning of either turn-taking behaviour. Note that the no short-term memory (STM) model always failed to learn drumming.

Table 3. Results of Fischer's exact test comparing the proportion of successes and failures for the memory model versus the other models. * indicates a statistically significant result at the 0.05 alpha level.

Peek-a-boo	
Model	*p*-value
Truncated STM	0.5
No STM	1.0
Drumming	
Model	*p*-value
Truncated STM	0.42
No STM	0.02 *

6.3. Time until Learning Success

In addition to differences in the number of successful trials, it was also expected that there might be differences in the amount of time it took the system to have the desired behaviours sufficiently reinforced to learn turn-taking. The metric used to measure the time it took to acquire a turn-taking behaviour was as follows: The start time was determined to be the first time the robot performed the *characteristic action* for a mode of turn-taking interaction, and the end time was determined to be the completion of the criteria for a behaviour to be considered learned as given in Section 6.1. As previously described, for the peek-a-boo interaction, the characteristic action was "hide-face". For the drumming interaction, the characteristic action was "drum-hit". The times for the trials are shown in Figures 7 and 8. Times for unsuccessful trials are shown as missing bars (marked with an "X") in the graphs. Note that of the 5 trials, the robot failed to learn drumming turn-taking once under the STM model condition. The robot failed to learn peek-a-boo once and drumming twice under the

truncated STM condition. Under the no STM condition, the robot did not learn drumming during any of the trials.
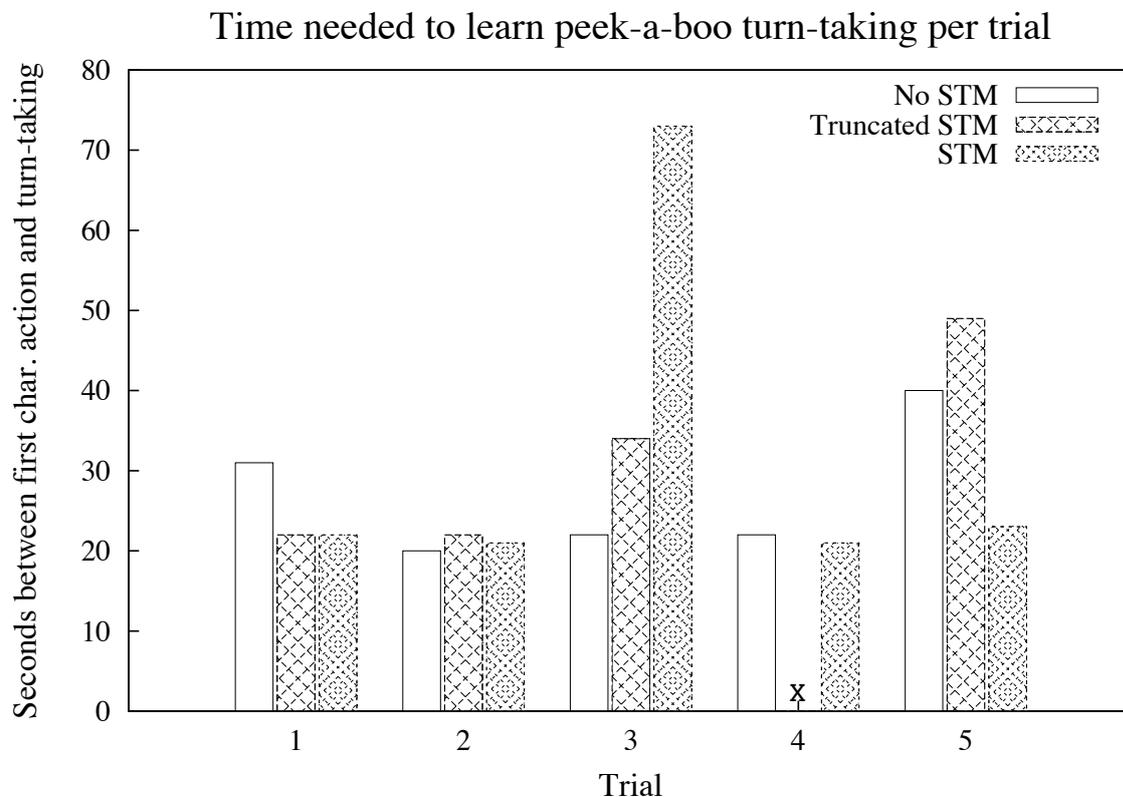

Figure 7. The time in seconds between the first instance of the peek-a-boo behaviour and the completion of three rounds of successful peek-a-boo turn-taking for each trial. Xs indicate that the behaviour was not learned before the time limit for that trial.

In order to determine whether the difference between the average times for the models to learn a turn-taking behaviour were statistically significant, randomised permutation tests using 1000 samples were performed. This test was chosen because of its ability to handle the small sample sizes and uneven number of samples (due to the unsuccessful trials) present in the data. The results are presented in Table 4. No significant differences were found between the average learning times for the successful peek-a-boo trials. The difference between the average times to learn drumming for the successful trials of the STM and truncated STM models is moderately significant.[2] This statistical support, combined with the difference in the number of successful trials (4 for the STM model and 3 for the truncated STM model), suggests that there is a real difference in the system's ability to learn the drumming interaction depending on the length of the short-term memory used. The learning time for the no STM model could not be compared for drumming, since there were no successful trials, demonstrating that the STM module was a necessary component of EIHA in order to learn this interaction.

---

[2]   The randomised permutation test is not an exact test, so the 0.05 significance level is approximate.

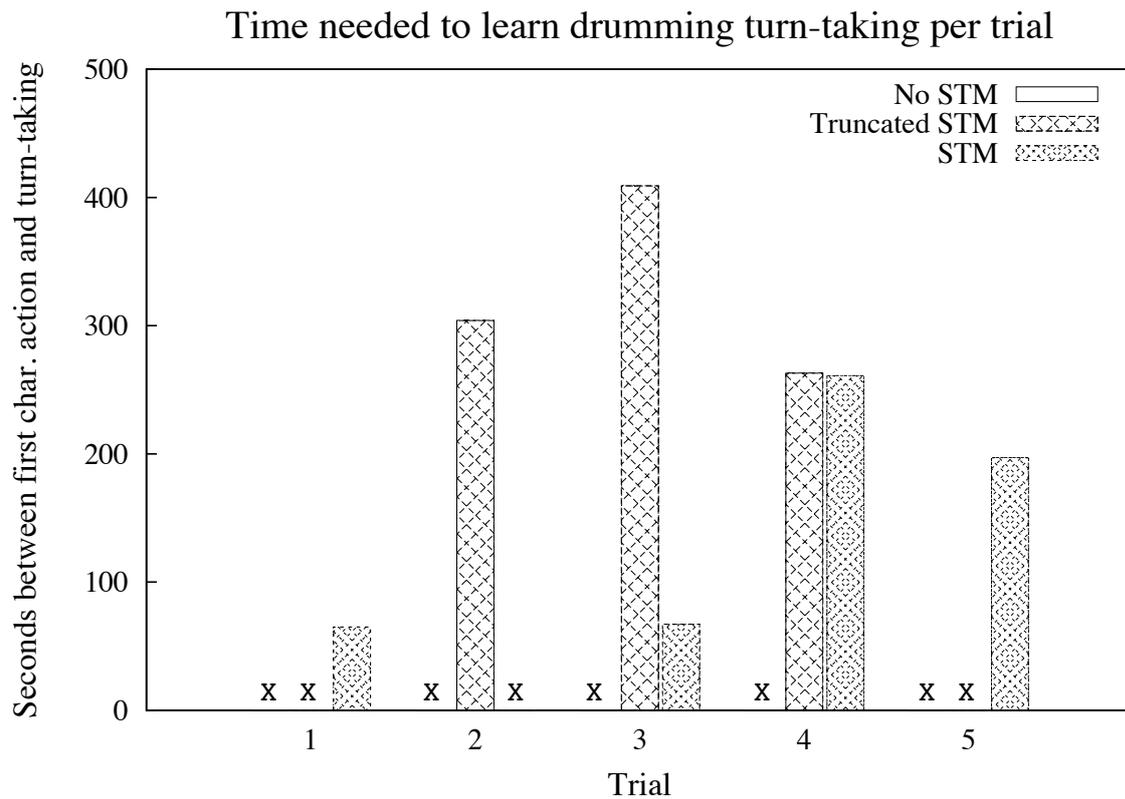## Time needed to learn drumming turn-taking per trial

Figure 8. The time in seconds between the first instance of the drumming behaviour and the completion of three rounds of successful drumming turn-taking for each trial. Xs indicate that the behaviour was not learned before the time limit for that trial.

Table 4. Difference between average times for memory model and other models for behaviour acquisition with significance results from randomised permutation tests. * indicates a statistically significant result at the 0.05 alpha level.

Peek-a-boo		
Model	Avg diff (s)	*p*-value
Truncated STM	−0.25	0.97
No STM	−5.0	0.85
Drumming		
Model	Avg diff (s)	*p*-value
Truncated STM	177.8	0.05 *
No STM	–	–

6.4. Discussion

These results suggest that a short-term memory of interaction is beneficial for the learning of some types of turn-taking in EIHA but may not have an impact on others. In order to understand why peek-a-boo could be learned quickly by all of the models while drumming could not, it is useful to look at the characteristics of the two behaviours. Peek-a-boo was a simpler interaction overall, requiring the learning of shorter action sequences with less variability in the durations of turns.

The no STM model could not learn the drumming interaction at all, showing that visual attention feedback alone was not sufficient for the human teacher to shape this behaviour. The results also demonstrate that the length of the short-term memory had an impact on the amount of time it took to learn the drumming turn-taking, with the truncated STM model requiring a longer time to learn the behaviour than the memory model. This is likely because the shorter duration of the

truncated memory contained less information about the transitions in turn-taking between robot and human (and vice-versa), making it harder to distinguish examples of successful turn-taking by their associated reward, which is partially determined by engagement scores based on information in the short-term memory.

In Mirza and colleagues' work on peek-a-boo learning with the original version of IHA, the learning success was found to be dependent upon the experience length [5]. It was hypothesised that the duration of an experience should be related to the duration of an action performed by the robot in order to support effective learning. The short-term memory module provides a way of relating the recent history of the robot's actions to those of its interaction partner. It allows reward values to be based on changes and relationships in the sensor data over time in a fine-grained way. We see in this experiment that the duration of the short-term memory also has an impact on learning success. The short-term memory should be long enough to capture the dynamics of the actions of both participants in an interaction and therefore is most effective when it is longer than the experience length and also based on the duration of the human partner's actions.

7. Future Work

The interactions learned by the system are "meaningful" in the sense of Wittgenstein [83] in that behaviours are used in structuring interaction games via enaction of the acquired behavioural sequences appropriately in context leading to rewarding social engagement (related to Peircean semiosis [61]). The role of learning in EIHA is currently restricted to the learning of the experience space and the associations between experiences and rewards in order to find effective behaviour sequences. However, there are opportunities to extend the capabilities of this system in order to expand the number and types of turn-taking interactions that could be learned by making the discovery of engagement patterns indicative of successful turn-taking in the datastream automatic. While the relationships between sensors monitored for feedback about turn-taking were predefined in this case, one could instead use statistical methods to discover which sensor channels are associated and predictive of one another. One possibility would be to explore the use of the cross-modal experience metric proposed by Nehaniv, Mirza, and Olsson [84]. Interpersonal maps, as described by Hafner and Kaplan, are another way of identifying these relationships using information metrics [85]. This would allow for interaction-specific turn-taking cues to be discovered, as well as general cues that are invariant over a range of interactions.

Switching between acquired behaviours has been demonstrated here. However, scaffolding of more complex, new behaviours that build on simpler acquired ones remains a near-term goal for this type of approach (which the architecture should with little or no extension be able to support). By following exploration in a zone of proximal development (ZPD) (Vygostsky [86]), variants and new actions could be selected in the context of interactions based on already acquired behaviours. Arguably, switching behaviours appropriately in response to social engagement cues is already a primitive case of this.

Imitation and timing play important communicative roles in structuring the contingent interactions achieved here, as they do in the case of human cognitive development [1,19]. The crucial role of timing and of relating different scales of remembering at different scales of temporally extended experience for behaviour is hinted at here and in previous work by Mirza et al. [5]. Short-term memory would seem to have an impact on the developmental capacity of cognitive systems. Further study of its role in development as well as of types of longer-term narrative memory would be fruitful directions for future research.

8. Conclusions

An enactive cognitive architecture based on experiential histories of interaction for social behaviour acquisition that uses reward based on a short-term memory of interaction was presented. An implementation that allowed the acquisition of and selection between behaviours for interaction

games using rewards from social engagement cues (visual attention and behaviour-specific measures of engagement and synchronisation) was described. The architecture provides the developing agent with prospection and behaviour ontogeny based on an operationalisation of the temporally extended experience and compares current experience to prior experiences to guide behaviour and action selection. The significance of the short-term memory and the importance of using a short-term memory length that is capable of capturing characteristics of the desired interaction behaviours in such an architecture was demonstrated experimentally, showing that the short-term memory had a beneficial effect on the amount of time required to acquire turn-taking behaviours and necessary to acquire longer, more complex behaviours. The capacity not only to learn new behaviours but to actively switch between them depending on the social cues in the context of interaction with a human has also been demonstrated.

Abbreviations

The following abbreviations are used in this manuscript:

IHA	Interaction history architecture
EIHA	Extended interaction history architecture
fMRI	Functional magnetic resonance imaging
HRI	Human–robot interaction
STM	Short-term memory
ZPD	Zone of proximal development

References

1. Trevarthen, C. Communication and cooperation in early infancy: A description of primary intersubjectivity. In *Before Speech: The Beginning of Interpersonal Communication*; Bullowa, M., Ed.; Cambridge Universit Press: Cambridge, UK, 1979; pp. 321–347.
2. Trevarthen, C. Intrinsic motives for companionship in understanding: Their origin, development, and significance for infant mental health. *Infant Ment. Health J.* **2001**, *22*, 95–131. [CrossRef]
3. Nadel, J.; Carchon, I.; Kervella, C.; Marcelli, D.; Reserbat-Plantey, D. Expectancies for social contingency in 2-month-olds. *Dev. Sci.* **1999**, *2*, 164–173. [CrossRef]
4. Kaye, K. *The Mental and Social Life of Babies: How Parents Create Persons*; University of Chicago Press: Chicago, IL, USA, 1982.
5. Mirza, N.A.; Nehaniv, C.L.; Dautenhahn, K.; te Boekhorst, R. Grounded Sensorimotor Interaction Histories in an Information Theoretic Metric Space for Robot Ontogeny. *Adapt. Behav.* **2007**, *15*, 167–187. [CrossRef]
6. Mirza, N.; Nehaniv, C.; Dautenhahn, K.; te Boekhorst, R. Developing social action capabilities in a humanoid robot using an interaction history architecture. In Proceedings of the 8th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Daejeon, Korea, 1–3 December 2008; pp. 609–616. [CrossRef]
7. RobotCub.org. 2010. Available online: http://www.robotcub.org/ (accessed on 10 February 2019).
8. Kose-Bagci, H.; Broz, F.; Shen, Q.; Dautenhahn, K.; Nehaniv, C.L. As Time Goes By: Representing and Reasoning About Timing in Human-Robot Interaction Studies. It's All in the Timing. In Proceedings of the 2010 AAAI Spring Symposium, Stanford, CA, USA, 22–24 March 2010; Technical Report SS-10-06.

9.   Nehaniv, C.L.; Förster, F.; Saunders, J.; Broz, F.; Antonova, E.; Köse, H.; Lyon, C.; Lehmann, H.; Sato, Y.; Dautenhahn, K. Interaction and experience in enactive intelligence and humanoid robotics. In Proceedings of the 2013 IEEE Symposium on Artificial Life (ALife), Singapore, 16–19 April 2013; pp. 148–155. [CrossRef]

10.   Varela, F.; Thompson, E.; Rosch, E. *The Embodied Mind*; MIT Press: Cambridge, MA, USA, 1991.

11.   Dautenhahn, K.; Christaller, T. Remembering, Rehearsal and Empathy—Towards a Social and Embodied Cognitive Psychology for Artifacts. In *Two Sciences of the Mind: Readings in Cognitive Science and Consciousness*; O'Nuallian, S., McKevitt, P., Eds.; John Benjamins North America, Inc.: Philadelphia, PA, USA, 1996; pp. 257–282.

12.   Vernon, D. Enaction as a Conceptual Framework for Developmental Cognitive Robotics. *Paladyn J. Behav. Robot.* **2010**, *1*, 89–98. [CrossRef]

13.   Tani, J. *Exploring Robotic Minds: Actions, Symbols, and Consciousness as Self-Organizing Dynamic Phenomena*; Oxford University Press: Oxford, UK, 2016.

14.   Johnson, M.H.; Morton, J. *Biology and Cognitive Development: The Case of Face Recognition*; Blackwell: Oxford, UK, 1991.

15.   Juscyck, P.W. Developing phonological categories from the speech signal. In *Phonological Development: Models, Research, Implications*; Ferguson, C.A., Menn, L., Stoel-Gammon, C., Eds.; York Press: York, UK, 1992; pp. 17–64.

16.   Farroni, T.; Csibra, G.; Simion, F.; Johnson, M.H. Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9602–9605. [CrossRef]

17.   Meltzoff, A.N.; Moore, M.K. Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Dev. Psychol.* **1989**, *25*, 954–962. [CrossRef]

18.   Trevarthen, C.; Aitken, K.J. Infant Intersubjectivity: Research, Theory, and Clinical Applications. *J. Child Psychol. Psychiatry Allied Discip.* **2001**, *42*, 3–48. [CrossRef]

19.   Nadel, J.; Guerini, C.; Peze, A.; Rivet, C. The evolving nature of imitation as a format for communication. In *Imitation in Infancy*; Cambridge University Press: New York, NY, USA, 1999; pp. 209–234.

20.   Hains, S.M.; Muir, D.W. Infant sensitivity to adult eye direction. *Child Dev.* **1996**, *67*, 1940–1951. [CrossRef]

21.   Corkum, V.; Moore, C. Development of joint visual attention in infants. In *Joint Attention: Its Origins and Role in Development*; Moore, C., Dunham, P., Eds.; Erlbaum: Hillsdale, NJ, USA, 1995.

22.   Frischen, A.; Bayliss, A.P.; Tipper, S.P. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* **2007**, *133*, 694–724. [CrossRef]

23.   Schilbach, L.; Eickhoff, S.B.; Cieslik, E.; Shah, N.J.; Fink, G.R.; Vogeley, K. Eyes on me: An fMRI study of the effects of social gaze on action control. *Soc. Cogn. Affect. Neurosci.* **2010**, *6*, 393–403. [CrossRef]

24.   Ross, H.S.; Lollis, S.P. Communication Within Infant Social Games. *Dev. Psychol.* **1987**, *23*, 241–248. [CrossRef]

25.   Elias, G.; Hayes, A.; Broerse, J. Maternal control of co-vocalization and inter-speaker silences in mother-ineant vocal engagements. *J. Child Psychol. Psychiatry* **1986**, *27*, 409–415. [CrossRef]

26.   Rutter, D.R.; Durkin, K. Turn-Taking in Mother-Infant Interaction: An Examination of Vocalizations and Gaze. *Dev. Psychol.* **1987**, *23*, 54–61. [CrossRef]

27.   Ratner, N.; Bruner, J. Games, social exchange and the acquisition of language. *J. Child Lang.* **1978**, *5*, 391–401. [CrossRef] [PubMed]

28.   Bruner, J.S.; Sherwood, V., Peek-a-boo and the learning of rule structures. In *Play: Its Role in Development and Evolution*; Bruner, J., Jolly, A., Sylva, K., Eds.; Penguin: Middlesex, UK, 1976; pp. 277–287.

29.   Rochat, P.; Querido, J.G.; Striano, T. Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Dev. Psychol.* **1999**, *35*, 950–957. [CrossRef]

30.   Gustafson, G.E.; Green, J.A.; West, M.J. The infant's changing role in mother-infant games: The growth of social skills. *Infant Behav. Dev.* **1979**, *2*, 301–308. [CrossRef]

31.   Kirschner, S.; Tomasello, M. Joint drumming: Social context facilitates synchronization in preschool children. *J. Exp. Child Psychol.* **2009**, *102*, 299–314. [CrossRef] [PubMed]

32.   Kleinspehn-Ammerlahn, A.; Riediger, M.; Schmiedek, F.; von Oertzen, T.; Li, S.C.; Lindenberger, U. Dyadic drumming across the lifespan reveals a zone of proximal development in children. *Dev. Psychol.* **2011**, *47*, 632–644. [CrossRef] [PubMed]

33.   Accordino, R.; Comer, R.; Heller, W.B. Searching for music's potential: A critical examination of research on music therapy with individuals with autism. *Res. Autism Spectr. Disord.* **2007**, *1*, 101–115. [CrossRef]

34. Kim, J.; Wigram, T.; Gold, C. The Effects of Improvisational Music Therapy on Joint Attention Behaviors in Autistic Children: A Randomized Controlled Study. *J. Autism Dev. Disord.* **2008**, *38*, 1758–1766. [CrossRef]

35. Phillips, L.; Tunstall, M.; Channon, S. Exploring the Role of Working Memory in Dynamic Social Cue Decoding Using Dual Task Methodology. *J. Nonverbal Behav.* **2007**, *31*, 137–152. [CrossRef]

36. Fuster, J.M.; Alexander, G.E. Neuron Activity Related to Short-Term Memory. *Science* **1971**, *173*, 652–654. [CrossRef] [PubMed]

37. Chan, A.W.Y.; Downing, P.E. Faces and Eyes in Human Lateral Prefrontal Cortex. *Front. Hum. Neurosci.* **2011**, *5*, 51. [CrossRef] [PubMed]

38. Kuzmanovic, B.; Georgescu, A.L.; Eickhoff, S.B.; Shah, N.J.; Bente, G.; Fink, G.R.; Vogeley, K. Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage* **2009**, *46*, 1154–1163. [CrossRef]

39. Romo, R.; Brody, C.D.; Lemus, L. Neuronal Correlates of Parametric Working Memory in the Prefrontal Cortex. *Nature* **1999**, *399*, 470. [CrossRef] [PubMed]

40. Ninokura, Y.; Mushiake, H.; Tanji, J. Representation of the Temporal Order of Visual Objects in the Primate Lateral Prefrontal Cortex. *J. Neurophysiol.* **2003**, *89*, 2868–2873. [CrossRef]

41. Oakes, L.M.; Bauer, P.J. (Eds.) *Short- and Long-Term Memory in Infancy and Early Childhood: Taking the First Steps towards Remembering*; Oxford University Press: Oxford, UK, 2007.

42. Chen, Z.; Cowan, N. Chunk limits and length limits in immediate recall: A reconciliation. *J. Exp. Psychol. Learn. Mem. Cogn.* **2005**, *31*, 1235–1249; Research Support, N.I.H., Extramural. [CrossRef] [PubMed]

43. Cowan, N. *Working Memory Capacity (Essays in Cognitive Psychology)*; Psychology Press: New York, NY, USA, 2005.

44. Diamond, A.; Doar, B. The performance of human infants on a measure of frontal cortex function, the delayed response task. *Dev. Psychobiol.* **1989**, *22*, 271–294. [CrossRef]

45. Schwartz, B.B.; Reznick, J.S. Measuring Infant Spatial Working Memory Using a Modified Delayed-response Procedure. *Memory* **1999**, *7*, 1–17. [CrossRef]

46. Baddeley, A. The episodic buffer: A new component of working memory? *Trends Cogn. Sci.* **2000**, *4*, 417–423. [CrossRef]

47. Cangelosi, A.; Schlesinger, M. *Developmental Robotics: From Babies to Robots*; MIT Press: Cambridge, MA, USA, 2015.

48. Ishihara, H.; Yoshikawa, Y.; Asada, M. Realistic child robot "affetto" for understanding the caregiver-child attachment relationship that guides the child development. In Proceedings of the 1st Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, Frankfurt am Main, Germany, 24–27 August 2011.

49. Asada, M.; Hosoda, K.; Kuniyoshi, Y.; Ishiguro, H.; Inui, T.; Yoshikawa, Y.; Ogino, M.; Yoshida, C. Cognitive Developmental Robotics: A Survey. *IEEE Trans. Auton. Ment. Dev.* **2009**, *1*, 12–34. [CrossRef]

50. Tsakarakis, N.; Metta, G.; Sandini, G.; Vernon, D.; Beira, R.; Becchi, F.; Righetti, L.; Santos-Victor, J.; Ijspeert, A.; Carrozza, M.; et al. iCub—The Design and Realization of an Open Humanoid Platform for Cognitive and Neuroscience Research. *J. Adv. Robot. Spec. Issue Robot. Platf. Res. Neurosci.* **2007**, *21*, 1151–1175. [CrossRef]

51. Chao, C.; Lee, J.; Begum, M.; Thomaz, A.L. Simon plays Simon says: The Timing of Turn-taking in an Imitation Game. In Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Atlanta, GA, USA, 31 July–3 August 2011.

52. MIT Media Lab Personal Robotics Group. Nexi. Available online: https://robotic.media.mit.edu/portfolio/nexi/ (accessed on 10 February 2019).

53. Imai, M.; Ono, T.; Ishiguro, H. Robovie: Communication technologies for a social robot. *Artif. Life Robot.* **2002**, *6*, 73–77. [CrossRef]

54. Aldebaran Robotics. NAO Robot. Available online: https://www.softbankrobotics.com/emea/en/nao (accessed on 10 February 2019).

55. Dautenhahn, K.; Nehaniv, C.L.; Walters, M.L.; Robins, B.; Kose-Bagci, H.; Mirza, N.A.; Blow, M. KASPAR—A Minimally Expressive Humanoid Robot for Human-Robot Interaction Research. *Appl. Bion. Biomech.* **2009**, *6*, 369–397; Special Issue on "Humanoid Robots". [CrossRef]

56. Metta, G.; Fitzpatrick, P.; Natale, L. YARP: Yet Another Robot Platform. *Int. J. Adv. Robot. Syst.* **2006**, *3*, 8. [CrossRef]

57. Scassellati, B. Theory of Mind for a Humanoid Robot. *Auton. Robots* **2002**, *12*, 13–24. [CrossRef]

58. Breazeal, C.; Buchsbaum, D.; Gray, J.; Gatenby, D.; Blumberg, B. Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots. *Artif. Life* **2005**, *11*, 31–62. [CrossRef] [PubMed]

59. Meltzoff, A.N. Like me: A foundation for social cognition. *Dev. Sci.* **2007**, *10*, 126–134. [CrossRef]

60. Quick, T.; Dautenhahn, K.; Nehaniv, C.; Roberts, G. Understanding Embodiment, System-environment coupling and the emergence of adaptive behaviour. In *Intelligence Artificielle Situe*; Drogoul, A., Meyer, J.A., Eds.; Hermes Science Publications: Paris, France, 1999; pp. 13–31.

61. Nehaniv, C.L. The Making of Meaning in Societies: Semiotic & Information-Theoretic Background to the Evolution of Communication. In Proceedings of the AISB Symposium: Starting from Society—The Application of Social Analogies to Computational Systems, Hillsdale, MI, USA, 19–20 April 2000; pp. 73–84.

62. Vernon, D.; Metta, G.; Sandini, G. A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents. *IEEE Trans. Evolut. Comput.* **2007**, *11*, 151–180. [CrossRef]

63. Hutto, D.D.; Myin, E. *Radicalizing Enactivism: Basic Minds without Content*; MIT Press: Cambridge, MA, USA, 2019; in press.

64. Noë, A. *Out of Our Heads: Why You Are Not Your Brain and Other Lessons from the Biology of Consciousness*; Hill and Wang: New York, NY, USA, 2009.

65. Brooks, R. Intelligence Without Representation. *Artif. Intell.* **1991**, *47*, 139–159. [CrossRef]

66. Mirza, N.; Nehaniv, C.; Dautenhahn, K.; te Boekhorst, R. Anticipating Future Experience using Grounded Sensorimotor Informational Relationships. In *Artificial Life XI*; MIT Press: Cambridge, MA, USA, 2008.

67. Robins, B.; Dautenhahn, K.; te Boekhorst, R.; Billard, A.; Keates, S.; Clarkson, J.; Langdon, P.; Robinson, P. Effects of repeated exposure of a humanoid robot on children with autism. In *Designing a More Inclusive World*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 225–236.

68. Koizumi, S.; Kanda, T.; Shiomi, M.; Ishiguro, H.; Hagita, N. Preliminary Field Trial for Teleoperated Communication Robots. In Proceedings of the IEEE International Symposium on Robots and Human Interactive Communications (RO-MAN2006), Hatfield, UK, 6–8 September 2006; pp. 145–150.

69. Ito, M.; Tani, J. Joint attention between a humanoid robot and users in imitation game. In Proceedings of the Third International Conference on Development and Learning ICDL 2004, La Jolla, CA, USA, 20–22 October 2004.

70. Breazeal, C. Towards Sociable Robots. *Robot. Auton. Syst.* **2002**, *42*, 167–175. [CrossRef]

71. Baxter, P.; Wood, R.; Baroni, I.; Kennedy, J.; Nalin, M.; Belpaeme, T. Emergence of Turn-taking in Unstructured Child-robot Social Interactions. In Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction, Tokyo, Japan, 3–6 March 2013; IEEE Press: Piscataway, NJ, USA, 2013; pp. 77–78.

72. Chao, C.; Thomaz, A.L. Timing in Multimodal Turn-taking Interactions: Control and Analysis Using Timed Petri Nets. *J. Hum.-Robot Interact.* **2012**, *1*, 4–25. [CrossRef]

73. Ogino, M.; Ooide, T.; Watanabe, A.; Asada, M. Acquiring peekaboo communication: Early communication model based on reward prediction. In Proceedings of the IEEE 6th International Conference on Development and Learning, London, UK, 11–13 July 2007; pp. 116–121. [CrossRef]

74. Kose-Bagci, H.; Dautenhahn, K.; Syrdal, D.S.; Nehaniv, C.L. Effects of Emodiment and Gestures on Social Interaction in Drumming Games with a Humanoid Robot. *Connect. Sci.* **2010**, *22*, 103–134. [CrossRef]

75. Weinberg, G.; Blosser, B. A leader-follower turn-taking model incorporating beat detection in musical human-robot interaction. In Proceedings of the 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI), La Jolla, CA, USA, 9–13 March 2009; pp. 227–228. [CrossRef]

76. Kuriyama, T.; Shibuya, T.; Harada, T.; Kuniyoshi, Y. Learning Interaction Rules through Compression of Sensori-Motor Causality Space. In Proceedings of the Tenth International Conference on Epigenetic Robotics, Lund, Sweden, 5–7 November 2010; pp. 57–64.

77. Demiris, Y.; Dearden, A. From Motor Babbling to Hierarchical Learning by Imitation: A Robot Developmental Pathway, 2005. Available online: http://cogprints.org/4961/ (accessed on 10 February 2019).

78. Wieser, E.; Cheng, G. A Self-Verifying Cognitive Architecture for Robust Bootstrapping of Sensory-Motor Skills via Multipurpose Predictors. *IEEE Trans. Cogn. Dev. Syst.* **2018**, *10*, 1081–1095. [CrossRef]

79. Crutchfield, J. Information and Its Metric. In *Nonlinear Structures in Physical Systems—Pattern Formation, Chaos, and Waves*; Lam, L., Morris, H.C., Eds.; Springer: Berlin/Heidelberg, Germany, 1990; pp. 119–130.

80. Dautenhahn, K. Getting to know each other—Artificial social intelligence for autonomous robots. *Robot. Auton. Syst.* **1995**, *16*, 333–356. [CrossRef]

81. Mirza, N.A. Grounded Sensorimotor Interaction Histories for Ontogenetic Development in Robots. Ph.D. Thesis, Adaptive Systems Research Group, University of Hertfordshire, Hatfield, UK, 2008.

82. Applied Science Laboratories. *Mobile Eye Gaze Tracking System*; Applied Science Laboratories: Industry, CA, USA, 2012.

83. Wittgenstein, L. *Philsophical Investigations*; Blackwell Publishing: Hoboken, NJ, USA, 2001; First published in 1953.

84. Nehaniv, C.L.; Mirza, N.A.; Olsson, L. Development via information self-structuring of sensorimotor experience and interaction. In *50 Years of Artificial Intelligence*; Lungarella, M., Pfeifer, R., Iida, F., Bongard, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 87–98.

85. Hafner, V.V.; Kaplan, F. Interpersonal maps: How to map affordances for interaction behaviour. In Proceedings of the 2006 International Conference on towards Affordance-Based Robot Control, Dagstuhl Castle, Germany, 5–9 June 2006; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–15.

86. Vygotsky, L.S. *Mind in Society*; Harvard University Press: Cambridge, MA, USA, 1978.