Friendly Machines: Interaction-Oriented Robots Today and Tomorrow

Stephen J Cowley
Takayuki Kanda

1.0 Introduction

Robot engineers have recently built machines whose functions are based on acting jointly with human beings. By participating freely in social life, interaction-oriented robots attract us to establish relationships with them. They may use facial expressions to reproduce aspects of human behaviour (Breazeal & Scassellati, 1999), or visual and auditory data that mimic aspects of personality (Okuno *et al.* 2002) They may carry out tasks like guiding people in a museum (Burgard *et al.* 1998), or use gaze to identify 'intentions' implicit in behaviour (Scassellati, 2000). Below, we report on how an interaction-oriented robot influences schoolchildren who 'get to know' the machine. In describing Robovie, we spell out why it could be of use in, for example, peer tutoring in a foreign language (Kanda *et al.* 2004a). In this paper, our main goal is to show how the friendly machine builds relationships and, thus, shows potential for making a positive contribution to society

Interactions occurred during a two-month experiment in an elementary school. In analyzing human-robot encounters we aim to use the robot's progress to enhance its future performance. Accordingly, we examine what children manifestly value in encounters with 'Robovie'. Focusing on what robots could detect and respond to, we highlight salient human activity and, especially, strategic responses that could be used to extend a robot's interactional powers. Highlighting 'realistic' behaviour, we sketch changes in human-robot encounters over the two months while emphasising how children tried to change

*Alter*nation 12.1a (2005) 79-105 *ISSN 1023-1757*

the interactional context. Thus, we stress moments when children either take the robot to 'mean' something or seek to 'tell' the robot something by using strategic affect. We believe that the 'context-making' function of such interpretations and signals is of considerable importance. Pursuing this, we sketch what context-making implies for software-design.

2.0 Interaction Oriented Robotics and Relationships

In interaction oriented robotics, machines are designed to interact with humans. These robots differ greatly from the task-oriented robots on which research has usually focused. The most successful task-oriented robots serve in factory automation where they typically assemble electronic devices or deal with heavy objects. Other task-oriented robots include space-exploration devices such as the Mars rover, cleaning robots, and so on. These all perform in limited domains because their control problems arise in a human-independent physical world. So, while task-oriented robots rely on physics and mechatronics, social behaviour is paramount in interaction-oriented robotics. Control is no longer separable from how robots influence us and thus participate in human life. In this field, perhaps the most notable success is with Aibo which behaved like a pet (Fujita, 2001). In more psychologically-oriented work, Breazeal and her colleagues explored robot sociality with respect to learning (e.g. Breazeal & Scasselati, 1999). Relatedly, Okuno and his team have a humanoid head that tracks a speaking person's visual and auditory data while altering a controlling parameter to adjust the robot's 'personality' (Okuno, et al 2002). In a practical application, Burgard *et al.* (1998) report on a tour guide robot whose robust navigational skills have been used in orienting people to a museum. Others have focused on specific capacities. For example, recognizing that joint-attention (Moore & Dunham, 1995; Tomasello, 1999) is essential to social life, Scassellati (2000) made a device that follows gaze and, using a different design, another robot develops similar powers (Kozima & Vatikiotis-Bateson, 2001). Plainly, robots can act in line with the 'intentions' that partly constitute human behaviour. Finally, a robot system can estimate human evaluation of robot doings by 'observing' body movements (Kanda *et al.* 2003a)

Reversing this emphasis, we examine how humans engage with robots. Using experimental work, we report how individual schoolchildren relate to robot partners and how the class-as-a-whole is affected by Robovie's presence. This issue arises because, to have a positive social impact, robots must prompt humans to evaluate their responses. Later, such activity can, we hope, be anticipated by the robots themselves. If this is done, instead of relying on canned behaviour to 'fool' humans, robots can act as if they grasped how activity is meant. While this may seem ambitious, no complex understanding is required. Rather, robots can use predictable human display that is inherent to the affective variability of their behaviour. Further, by modelling such processes, our work may serve the study of social learning. For the moment, however, in dealing with what response affords, we focus on behaviour that can enable robots to exploit human attempts at relationship-building.

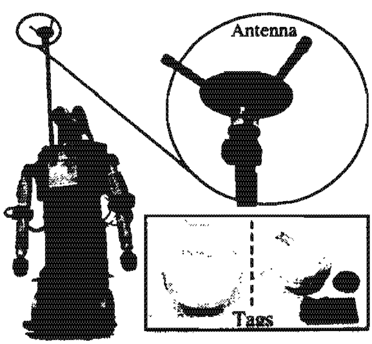

Figure 1: Robovie and Wireless tags

3.0 Robovie: A Behavioral Approach

Robovie is designed to interact at a child's level. For this reason, rather than focus on appearance or learning, the work aims to generate interactionally appropriate behaviour patterns. These rely on human-like expression using robot arms together with its eyes and head. Taken together, these produce gestures that prompt human interaction-oriented behaviour. Robovie also recognizes individuals by using actuators and auditory, tactile, ultrasonic, and vision sensors. The machine's processing and control systems, the computer and motor control hardware, are located within the robot's body (see, Figure 1).

The robot's activities are controlled by software that ensures it performs consistent behaviour. In design, our 'active interaction' approach aims to compensate for the machine's imperfect sensory processing. This is important because sensory-recognition technology is not yet sufficiently advanced to identify much human behaviour. In this approach, therefore, robots proactively initiate interactions that entice humans to respond adaptively to the robot. The robot performs a series of interaction-oriented behaviours so that its embodiment (head, eyes, arms, etc.) entrains humans to its behaviour. This is generated by a situated module whose operations use communicative units (see Kanda *et al.* 2002 for discussion of the mechanism).

Currently, the robot uses 100 interactive patterns including shaking hands, hugging, playing paper-scissors-rock, exercising, greeting, kissing, singing, verbal output, and pointing to a nearby object. In addition, the robot has about 20 idling behaviours such as scratching the head and folding the arms, as well as 10 moving-around behaviours. In total, the robot utters more than 300 constructions and recognizes about 50 word-sounds. These give rise to patterns sequenced in accordance with simple rules. For example, the robot may trigger interaction with, "Let's play. Touch me." Next it exhibits idling or moving-around until the child responds. If this occurs, it performs friendly behaviours to sustain the child's interest. Then, when the child stops reacting, the robot ceases its behaviour, produces "good bye," and returns to idling or moving-around.

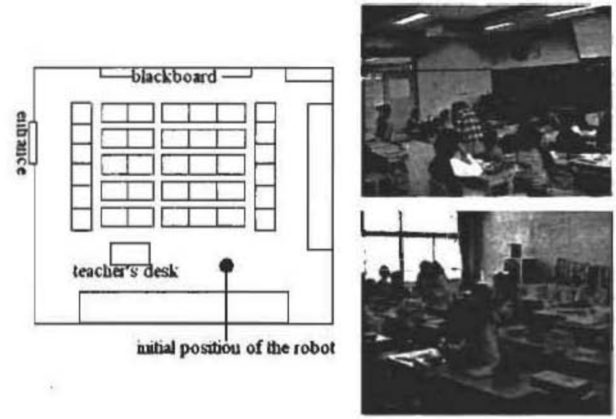

Figure 2: Environment of the elementary school

Given interest in relationships and long-term interaction, the robot has person identification functions based on infrared logo carried by each child. To facilitate this, the child's name is sometimes uttered when he or she is near the robot. For instance, if a child (named Yamada) is near the robot, hearing "Hello, Yamada-kun, let's play together" will strike her as significant. The affective consequences of the robot's action thus prompt both immediate and subsequent interaction. Where patterns emerge across interactions, following Hinde (1979), the resulting behaviour is said to enact a 'relationship'. Further to this, sustained interaction and relationships are also served by pseudo-learning. When a child interacts repeatedly with the device, the machine expands its active repertoire. Thus while a child who interact with the robot for the first time sees at most 10 behaviours, one with 180 minutes of experience may encounter up to 100 interactive patterns. Since actions are coupled with the child's experience this Robovie, appearance of learning is created. Last, the robot confides personal-themed matters to its frequent partners. This time threshold serves to motivate children to spend more time with the robot. Personal themes include the comments, "I like chattering" (said to a child who has played for 120 minutes), "I don't like the cold" (180 minutes), "I like our class teacher" (420 minutes), and "I support the Hanshin-Tigers (a baseball team)" (540 minutes).

To achieve human-like expression and recognize individuals, the robot uses various actuators and sensors. Its arms posses 4 degrees of freedom (DOF), its eyes 2, and its head 3 (Fig. 1, left, above). The sensory equipment includes auditory, tactile, ultrasonic, and vision sensors, which allow the robot to behave autonomously during encounters. For similar reasons, processing and control systems, such as computer and motor control hardware, are inside the robot's body. To identify individuals, it uses a wireless tag system for multi-person recognition. Recent radio frequency identification (RFID) technology permits this to be achieved using contactless identification cards and chips. In this study, children wore nameplates (5 cm in diameter) in which a wireless tag was embedded. A tag (Fig. 1, lower-right, above) periodically transmitted an ID to a reader on the robot. In turn, the reader relayed received IDs to the robot's control system. Software made it possible to adjust the reception range of the receiver's tag in real-time. The wireless tag system provided the robots with a robust means of identifying many children simultaneously. Consequently, Robovie showed adaptation by recalling the interaction history of a relationship (Kanda *et al.* 2003b)

4.0 Methods

Using Robovie, we conducted a field experiment in a school. Next, we sketch what happened before and, in the following sections, turn to how the robot affected the children.

Robovie was maintained in the classroom of an elementary school. It could thus interact easily with 37 subjects (10-11 years old, 18 male and 19 female) who belonged to a fifth-grade class. The experiment lasted for about 2 months or 32 'experiment days'. (Of 40 school days, 8 were omitted for pedagogical reasons.). The children freely interacted with the robot during a 30-minutes recess after lunch (Fig. 2.). To focus on long-term interactions and relationships, the children wore nameplates with an embedded wireless tag. These enabled the robot to record the tags, recognize the children, and calculate how long each spent with the robot. The data were used in analyzing the interaction. Finally, we administered a questionnaire that asked about both the children's friendship with other children and their interest in the robot.

5.0 How Children Relate to Robovie

We classify the nine weeks into three phases (Fig. 3), and explain the interactional transitions between them. Then we focus on who interacted with Robovie and their claimed reasons for so doing.

During the first phase, children crowded around the robot. Initially, they started a queue (Fig. 4) and, on the first day, up to 17 children were simultaneously present. During the first two weeks, its novelty ensured that the robot almost always had children nearby. Although the numbers gradually decreased, at least one child was almost always engaged in interaction. We highlight some scenes:

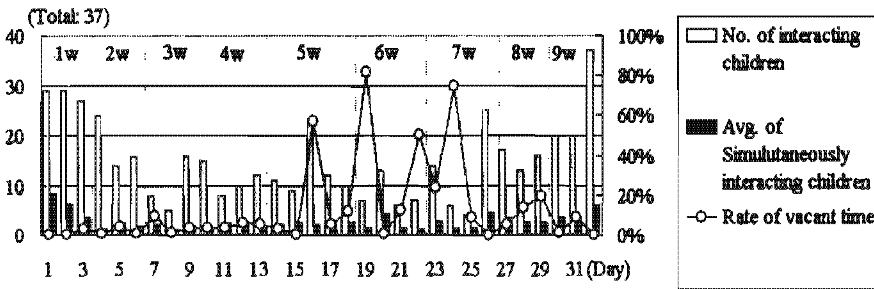

Figure 3: Transitions of the interaction between children and the robot

Many children were attracted by the name calling behaviour.

- Children tried to get the robot to call their names by showing their nameplates to the robot's eye and omnidirectional camera (Fig. 4-b).
- Hugging behaviour was a favourite of the children

(a) Beginning of the first day: Children formed a line (b) showing nameplate (c) "I can't see" behaviour preferred

Figure 4: Scene of the experiment

In the second phase, from the 3rd to the 7th week, interactions tended to stabilize. Generally, Robovie attracted up to ten children and, at any one moment, one or more would interact with the machine. When it was raining, children who often played outside boosted the number of partners who played the machine. During these five weeks, as interest diminished, vacant time increased. Then, the "confiding of personal matters" behaviour first appeared and became popular. In this second phase, we observed the following.

- Child A observed the "confiding of personal matters" and told her friend, "the robot said that if you play with it for a long time, it will tell you a secret."
- Child B said, "Please tell me your secret (personal matters)!"
- Although Child C asked the robot about the personal matters, the robot didn't say anything. Child D was watching the scene and told child C what had previously been confided.

The robot gradually performed new behaviours using a pseudo-learning mechanism, and these behaviours caught their attention.

• When Robovie's eyes were hidden (Fig. 4-c), it brushed off the obstacle and said "I can't see." This new behaviour was so popular that children often tried to hide the robot's eyes.

• The robot started singing, and the children joined in.

In the final phase, although the number of children around the robot increased, the number playing with the robot remained constant.

	Correlation / Statistic test result (*p<.05)
Gender (male/female)	not significant
Friendship motivation (Q.1)	0.35 *
Mechanical interest (Q.2)	-0.40 *
Indoors/ Outdoors (Q.3)	significant *

Many simply came to watch the interaction. On the first day of the 8th week, the class teacher told them that the robot would leave at the end of the next week and, clearly, this affected their behaviour. Further, as "confiding of personal matters" became well-known, children were fascinated. For example, they listed the robot's claims on the blackboard. Of these, the most popular was its statement, "I like the class teacher." When the robot first said this, children ran out of the classroom to tell the teacher. Finally, on the last day, the children held a farewell party for the robot. They formed a queue and played with the robot one by one.

5.1 Who Did What When?

To investigate who did what in detail, we sub-classified the children by how much time they had spent with the robot. Simply, we divided them into a "more than half" category (children who played with Robovie more than 16 out of 32 days) and a "less than half" one (children who played with it on fewer than or equal to 16 days). This enabled us to compare the children's explanations of what they had done with measures of total interaction time. In total, 10 children (4 males and 6 females) fell into the "more than half" group (27 had interacted less than half) and, unsurprisingly, these children had played more consistently with the robot over the period. By contrast, others had tended to play with it, especially, when it was novel and just before it left (in the first and third phases) (Figures 5 and 6).

To get a preliminary idea of how children explained their actions, we used a questionnaire asking whether they were motivated by friendship, mechanical interest, gender and where they usually played. Not surprisingly, explanations in terms of friendship (Q.1) had a significant positive correlation with total interaction, and, strikingly, mechanical interest (Q.2) a significant negative correlation. While the effect of gender was non-significant, children who usually played inside were significantly more likely to fall into the positively correlated group than outdoors types (Q.3: outdoors type and indoors type) ($F(1,35)=4.39$, $p<.05$) (Table 1).

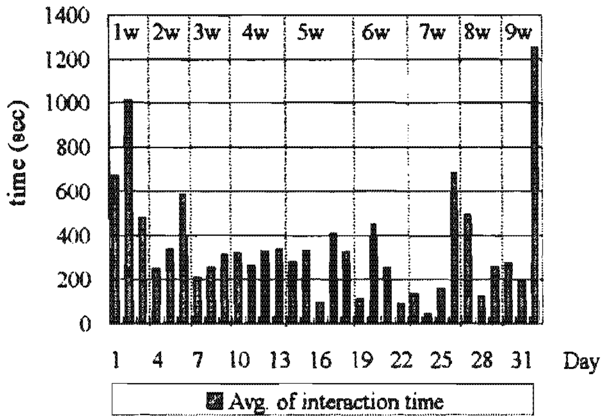

Figure 5: Average interaction time (More than 16 days children)

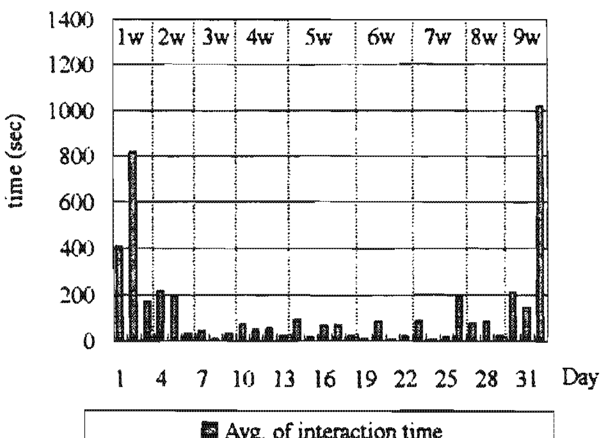

Figure 6: Average interaction time
(Less than or equal to 16 days children)

These findings suggest that when children want to be friends, they do *not* want to know about mechanisms. Less surprisingly, friendship motivation affects the time spent playing with the robot as does whether a child is an outside or an inside type. What is most striking, however, is that a motivation for relationships (as opposed to seeing the robot as a machine) is correlated with the time spent in encounters. Clearly, since much depends on a child's imagination the relationship is asymmetrical and represents what we call social 'partnerhood'.

6.0 Engaging with Interaction-Oriented Robots

Robovie continued friendly interaction for two months using its basic functions, a capacity to identify individuals, pseudo-learning and personal confidences. The children's responses contrasted with what often happens with a novel object. This was because encounters were shaped both by micro-behaviour and relationship typical patterns. The relationships dimension of encounters is marked, above all, in differences in how children respond. While the subject of current research, we next delimit how children's reactions varied across the stages to suggest this signifies for current design. Next, however, we sketch why we focus on relationships.

Primate behaviour has complexity not found in other social mammals (see, Hinde, 1983, 1987). Especially in the great apes, this is reflected in dominance hierarchies, individual recognition, the all-pervasiveness of affect, and interindividual relationships. These serve, above all, in alliances whose hedonic quality influences individual outcomes while contributing to the formation, maintenance and transformation of social roles. With Humphrey (1976), the rise of such relationships probably drove the evolution of primate intelligence. In humans, moreover, another factor is prominent: while other species have culture (e.g. Rendell & Whitehead, 2001), our material and oral institutions permit novel ways of spreading acting, feeling and knowing in space-time. Together with social intelligence, culture transforms lifeworlds by re-organizing activity and, by extension, brains (e.g. Deacon, 1997). Both relationships and cognitive powers, therefore, are crucial to the outcomes that arise as we co-operate and compete with one another.

While all behaviour is multiply caused (Tinbergen, 1952; Hinde, 1987), material and oral culture give human life-forms unique complexity. Our interactions can deviate from common primate patterns by virtue of our use of language. In Dunbar's (1993) terms, we add gossip to relationships. This not only allows for forms of indirect manipulation but gives a role to folk psychology and many social institutions. Taken seriously, this view allows talk to be seen as an extension of expressive behaviour (Cowley & Spurrett, 2003; Spurrett & Cowley, 2004). Since this makes human communication, ontogenetically, no different from that of other animals, the view has advantages. Above all, like that found of other species, human communication depends on expressive activity that serves to assess and manage conspecifics (Owings & Morton, 1998). Accordingly, as Hinde (1983; 1987), shows, it can be described by dialectical linkages between levels of complexity. On this model, even an act of greeting can be informative about a person's internal state, identity, desire to interact, social background, allegiance to sub-groups, inter-group history as well as situation and time-bound cultural process. Simply, human activity exploits wants and beliefs.

Human-robot encounters use how a child's beliefs are manifest in behaviour. How the device responds can be designed, to an extent, around seeking to manage humans by orienting to social norms. Unless goal driven or task-oriented, humans treat messages as inseparable from messengers and create coherence by ensuring that text and context constitute each other (see, Glenn *et al.* 2003). These outcomes, then, use cognitive powers that spread beyond the brain. Agents can rely on physical resources (or programs), material artefacts, and organized activity (Hutchins, 1995). Human cognition is distributed in that artefacts, norms and beliefs are used for cognitive, affective and practical goals. Emphasising the distributed nature of cognition, we stress that robots prompt action based in beliefs and affect-laden impressions. In contrast to human-computer interaction, the human-like nature of a robot prompts humans to experience a range of emotions. Children thus use interaction-oriented robots not merely to optimize goal-directed behaviour (within the constraints of task and competence) but also to gain affective rewards. To co-ordinate activity 'realistically', a robot's behavioural repertoire must include ways of responding to the context-making that we describe in terms of beliefs, feelings, moods and desires.

While human-robot relationships exploit design or hard-wiring, they also use hedonic behaviour, tricks, in-built biases and environmental factors. Just as biological systems soft-assemble new kinds of functionality, robots can use affective behaviour encounters to retool their repertoire. In this respect there is a parallel with a human ability to co-construct higher-order functions by using input-output mechanisms to exploit external features of interactions. For example, if speaking evokes unexpected laughter, a speaker will reorient both what is said and how s/he presents her/himself by producing (what we call) a new 'context'. With robots, opportunities for similar behaviour arise, say, when children cover Robovie's eyes. This 'debugging' behaviour has the potential to engender new forms of interaction. Another example arises with children who relish the 'hugging' that Robovie is hard-wired to perform. In their belief-worlds, this is *more than* concerted movement. Although this sense is unconscious and based on affect and cultural processes, hugging is treated as intrinsic to a relationship. In building interaction-oriented robots, therefore, we focus on behaviour motivated by and aimed at relationships. Accordingly, we now ask how in-built tricks and biases influence the children and, equally, how human activity aims to influence Robovie. In reporting on both human response and context-making, we reflect on events of significance for relationships.

6.1 Human Response to Simple Tricks

Human response to Robovie is not dominated by either motor movements or canned phrases. Far from treating the machine's activity as rule-governed output, children often interpret it as designed to change the context. This semblance of context-making arises, in part, from Robovie's 'simple tricks'. As specified below, these both persuade children that they are party to two-way relationships and, in real-time, motivate strategic context-making. For the software designer this means that, in principle, robot-motivated response could be used to soft-assemble high-level interactional functions. Next, we show how simple tricks prompt context-making interpretation. This enables us, in following section, to show how human propensities to base relationships on affect prompt context-making behaviour that a robot could detect with even current technology.

Since Robovie recognizes individuals, humans react as if the behaviour was interesting (Kanda *et al.* 2003b) and, perhaps, flattering. Thus they respond positively while trying to set up name-calling routines. For example while one child shows a name-plate to a camera, another's strategy is to tell the robot what he is called. Although this can be described at an interaction level, the children are engaged in context-making that, in a human setting, serves at building a two-way relationship. By attempting to "tell" the robot something, they attempt to consolidate their partnerhood. Were robots able to detect and support the relevant beliefs, context-making outcomes could be incorporated in an interaction history to serve 'special' features of relationships. From a child's perspective, the resulting behaviour would show that the robot 'understood' relationship-level patterns. Using something like the self-fulfilling prophesies of child development, they could be used to develop joint routines. Robots and human-partners could use context-making to construct interactional routines that give a relationship a unique quality. For a child, this would be of considerable social significance.

Similar generalizations apply to pseudo-learning. While some children found that the novelty-value of the robot wore off, others showed sensitivity to its repertoire. Thus, one child who played with the robot intermittently found it boring because, in part, "Robovie can talk, but it always speaks about the same thing." In contrast, another said, "Since the vocabulary of the robot increases, it became easy to talk with the robot." The examples show that pseudo-learning contributes to long-term interaction and relationships. Using an ability to remember what Robovie previously said as well as human beliefs (and wants), a child values context-making behaviour that makes talking 'easier'. Given such results, Robovie's potential can be developed by controlling the increase of such behaviour. Ideally, of course, this will link with context-making to allow for the strategic management of a specific relationship.

The most significant effects of the robot's ability to 'confide personal matters' appears, not in relationships, but at the level of the group. Instead of providing a new context for relationships, this enabled a child to "tell" to the group. Given design where a reward for interaction was that of being singled out for confidences, playing with the robot influenced a child's group-status (at least briefly). This goes towards explaining comments like "I played with Robovie to investigate its personal matters." Evidence that canned phrases affected how the group perceived Robovie is found in for example, institutional use of the blackboard to list his confidences. No doubt, this had a positive effect on the person to whom the 'matter' was confided and at the same time, the individual's place in the class. Robovie has potential for using interindividual relationships to influence the class-as-a-whole. Interaction-oriented robots can exert positive effects on human groups.

6.2 Spontaneous Context-Making Signals

Perhaps the most significant finding in how children respond to Robovie is that, as with humans, what they do is less affected by robot behaviour than an imagined relationship. This depends, above all, on interpreting aspects of behaviour as designed to change the context.

Clear evidence is found among children who want to be friends with the machine. Rather than treat Robovie as a computer-like tool, these children experience partnerhood with a device they imagine to be human-like. Not only does this come out in questionnaires but it is also manifest in what they say and do. Thus one child who interacted extensively with the machine reported, "Robovie seems lonely and wants to talk," and, "although Robovie is a robot, I feel it has a human-like presence." She also said "when I interacted with Robovie, I felt as if I had interacted with a human friend. Perhaps, this is because I got accustomed to interacting with it." For some children, Robovie's human-like properties are salient.

While reflected in verbal expression, context-making attributions are also manifest in behaviour. In only minutes, therefore, a human observer can establish which children have the 'best' relationships with Robovie. This is possible because the tone of an encounter is shaped by a child's sense of the context-of-relationship. While ongoing research is needed to clarify 'good' human-robot interaction, an intuitive notion of quality rapidly appears. For example, after Robovie has rewarded her with a song, one child does a dance of delight and then pats the robot's head. This is important because, in principle, such context-making could be predicted and used to trigger robot responses. Indeed, if robots attend to such context-making, they would be experienced as attuning to how children feel. While robotics research must not rely on changing human attributions, context-making can be used to enhance robot-human relationships. In this way it would parallel infant-directed speech which, while 'unrealistic' makes infant affect and bodily action more comprehensible. Further, since this kind of context-making has strategic functions like making things 'interesting', similar events can help a robot drive long-term behavioural change by coupling inner and outer motives. Changes in activity could thus be prompted by context changing behaviour: the robot might, for example, manifestly try to hide its mechanism, or act to make a child more alert to its 'friendliness'.

6.3 Relationships and Human Emotions

We have used examples of context-making that do not use affect. In primates, however, interindividual relationships juxtapose hedonic events with competitive practices and changing social roles. In humans, these use a range of emotions and bodily dynamics that serve, above all, to build, maintain and challenge alliances (Ross & Dumouchel, in press). Of course, in primates, these ensure group cohesion and reflect both an individual's powers and social 'status'. Accordingly, many affective strategies used during interaction exploit a subset of emotions for social ends (allowing us to avoid nasty dilemmas). This is especially so for emotions (and bodily dispositions) like interest, disappointment, pride, shame, respect and guilt as well as expression-kinds associated with behaviour that we describe as showing delight, disappointment and surprise. It is of interest, then, whether human response exploits this emotional profile and the extent to which it is of potential value for robots.

While possible to give only a sketch, it is clear that reactions to Robovie reflect, among other things, children's strategic signals. For example, *Hanako* (invented name) experiences a close relationship with the robot. Not only did she spend a total of 8.44 hours with the machine but, after the experiment, she reports that she thinks of Robovie as a friend. Thus, she is happy when Robovie calls her name and describes herself as chatting with the machine (in fact, she replied to verbal sounds). When she touched his shoulder and Robovie said "What is it?", she felt "Robovie behaves as if it is human." On one occasion, she got him to sing a song. When it carried out this wish, she carried out a dance of delight. Clearly, individual differences affect relationships through the child's 'model' of Robovie. Even statistically, those interested in how the machine works have poorer relationships than those who see Robovie as a friend. Children who interacted frequently reported, "Robovie seems lonely and wants to talk with others as if it is a human," and "with Robovie, I felt as if I am with a human friend." This, moreover, is bound up with preferences about modes of play and, specifically, if they choose to interact with the robot or other objects.

What is striking about such incidents is that, far from reacting to the robot, events that are interpreted within a personal relationship nonetheless motivate the child to share experience. Not only does Hanako feel delight but, she seeks to share this with her classmates and, then, comes back to show gratitude to the robot. Plainly, if robots can be pre-programmed to pick up strategic signals, this would impact on how we conceptualize relationships and the group as a whole. Not only could elicited behaviour set off such responses but, in principle, Robovie could respond to her responding. In principle, such events could be stored in a relationship memory and recycled to increase a child's sense that she was special *to the robot*. By setting up relationship-based norms, powerful affective responses could be provoked. In this way the robot would mimic primate-like intelligence without needing hedonic tone. As goes without saying, this could produce positive (and negative) responses: the routine could thus also sustain reinforcement learning. For the same reason, it could be used in ways that had an impact on an individual's social status.

7. 0 Robotics in Behavioural Science: Future Directions

The complexity of child response to Robovie suggests that, just as the invention of computers boosted the study of cognition, the development of robots may change thinking about complex behaviour. In dealing with robots, children exploit events in ways that are irreducible to behavioural sequences. As seen in around simple tricks, context-making and affect, child activity defies Stimulus-Response or Input-Plan-Action description. While sensitive to Robovie's actions, this serves mainly to background attempts at relationship building. As context-making strategies show, encounters are 'deeply' affected by telling and sharing affect. What happens depends, strikingly, on wants and beliefs. Equally, what we 'think' of robots (are they partners or mechanical tools?) affects encounters as much as a child's feelings, social status, and their lived relationship. While not unexpected, these findings matter. Above all, software can be designed to deal with effects that are only distantly related to 'legal' input. Instead of relying on defining symbol-object connections in advance (MacDorman, 1999), it is possible to play down the use of competence-performance models. Instead, software designers can develop systems that soft-assemble by using what a human individual's behaviour affords. Salient features of encounters thus become resources used in the robot's relationship-building. While much can be gained from robust sensory systems, much also depends on creating software that exploits human context-making.

In robotics, current thinking focuses on using cognitive resources efficiently and realistically. It is often assumed that competence-performance models are a good basis for designing software. Our findings give reason for doubt. In practical terms, focus on efficiency ensures that, as technology and tasks change, designs become obsolete. This, we believe, means that designers of social robots need to conceptualise the software-behaviour relationship strategically. In addition to programs that control what robots do, we need mechanisms that prompt and respond to longitudinal changes in how humans seek to alter the context. In conceptual terms, of course, robot behaviour must be designed to influence social life while also exploiting the constraints of the physical world. In this domain, what is characteristic of humans is not the consistency of behaviour but, rather, that their activity adjusts round norms. Accordingly, when other persons are present, we constantly adjust our doings to context (see, Goffman, 1959, 1974). Indeed, as Watzlawick *et al.* noted (1967), the pattern is so marked that even doing nothing is usually communicative or, in our terms, has strategic, context-changing value. Models that specify 'realistic' human-robot encounters can emphasise context-making to use changes in human behaviour for parameter setting. While tricks can provoke context-making, machines can also be made sensitive to their strategic human-based counterparts. Accordingly, software can be dedicated to self-organizing behaviour that establishes what seem to be two-way relationships. This is why it matters that robot-child encounters use behaviour, beliefs prompted by simple tricks, context-making and affect. To be 'realistic' robots can exploit what people say and do together with feelings, wants and beliefs. This kind of software design will enable a system to develop what can be called 'social strategy management'. In such a model, a robot will exploit both behavioural variability and context-making in routines and relationships.

To specify how people respond to robot doings and set up relationship-oriented initiatives we need to develop coding systems to characterise human-robot interaction. In current work, therefore, we aim to capture what human activity affords a robot (of given specifications) and, in the longer term, to describe 'quality' human-robot interaction. While such models depend on longitudinal observation, they will also use theories of sociocognitive development and constructivist social learning. Since humans respond to Robovie in complex ways, the robot can be eclectic in using resources to optimize performance. On this view, rather than focus on hardware specification based on behaviour, we stress the need for soft-assembled interfaces that, using interactions, make a machine friendly. Given human interest in relationships, human-robot encounters are best seen as complex social behaviour. Rather than treat cognitive resources as exclusively internal, robot response can be trained by children's activity. This is consistent with work on learning in apes and human infants where developmental complexity depends on interactive gearing of environment and brain. During infant development, moreover, people facilitate their adjustments and, in so doing, encounters complexify so that joint behaviour adjusts to shared beliefs. In such terms, encounters are dynamical events where humans exploit the iconic, indexical and symbolic properties of activity (Deacon, 1997; Thibault, 2000; Cowley *et al.* 2004). By attending to the quality of movements, we gradually become participants in joint action. Human-robot encounters can develop along such lines provided that we use imaginative software design. Above all, this must enable human behaviour to serve the robot in reorganizing its actions to 'fit' recognizable human acts.

Human-robot encounters already use built-in functions together with belief-based analysis. Note is made of strategic affect, context-making and how human desires and beliefs about relationships shape their reactions to, say, simple tricks, pseudo-learning and robot appearance. It follows, then, that strategic phenomena ought to be of as much interest to robots as is human behaviour. Further, since relationships motivate much human action, it is especially important to consider the relevant regularities. This, indeed, is the importance of debugging behaviours as well as their context-making counterparts. In seeking to characterise how interaction develops in relationships, we move towards defining what makes some encounters 'good'. Of course, this level of description cannot be separate from either software design or the robot's hardware. In appealing to cognitive efficiency and human-like performance, importance must also be given to behaviour and appearance. Yet, far from focusing single-mindedly on making action 'true-to-life', we must establish how behaviour varies within (predictable) human parameters. This, we believe, is the importance of perceptually salient events like debugging, strategic affect, and context-making. In principle, such child-activity can be detected by Robovie and, thus, used to respond to what children believe. If a robot responds to debugging, a child will feel she has made a difference; in picking up on context-making affect, she will feel that the robot is sensitive to her feelings. Since the child will act as if the robot is responding intelligently, this will allow the machine to mimic two-way relationships. In software design, one goal is to use the robot to concert interaction around relationships that can be described both behaviourally and around strategic affect. One can speculate, moreover, that such behavioural effects can be enhanced by giving robots powers of facial expression. Indeed, given such signalling, children would treat the robot as acting strategically: furthermore, if based on human performance, such dynamic properties would make them easy to recognize and detect.

8.0 Friendly Machines Today

Placing Robovie in an elementary school for two weeks shows clearly that children seek to build relationships with an interaction-oriented robot. More than that, they treat such a machine as inherently friendly. While the evidence is less strong, questionnaires and behavioural evidence confirm that Robovie's presence has positive effects on individuals, relationships and the class. This is achieved, we stress, by a robot which is human-like through the eyes of imaginative children. Currently, of course, there are serious limitations not only on the robot's sensory and motor resources but also software-design that functions by supplementing interactionally appropriate initiatives with output based on sense-data, individual recognition, pseudo-learning and secret-telling.

Given the software's simplicity, the complexity of human-robot encounters is startling. Above all, children value the machine's human-like nature and treat it with affection. Far from perceiving Robovie as a rule governed device, they freely strive to establish a relationship with a human-like friend. Significantly, they perceive some unusual behaviours as context-making and, thus, of potential value in relationships. While many are taken personally and lead to reciprocal attempts at context-making (such in individual recognition and name telling), others have marked social consequences. This is most obviously so in the robot's confiding behaviour. In this classroom, having a relationship that 'persuades' Robovie to sing songs or tell you secrets makes a child partner feel good and, briefly at least, raises her social standing. In dealing with Robovie, therefore, it is misleading to describe events as program derived sequences. Even now, response to Robovie is driven less by what the robot does than by what children believe and feel.

Viewed both longitudinally and around striking incidents, children value moments where the machine is felt to produce behaviour directed at the individual. They value friendliness. In this respect, there is a telling contrast with the 'personalization' important in relationships with computers. Since Robovie cannot be programmed, issues of ownership and personalization give way to what children treat as a complex interface that, potentially, can sustain a two-way relationship. Drawing on a history of encounters, a child tends to treat the machine as human-like and, as described above, may come to feel special 'for' the robot. This way of setting up a relationship, we stress, depends no more on robot behaviour than a child's imaginings or a robot's tricks. Strikingly, the perception underpins both context-making and produce positive group outcomes.

Currently, a two-way relationship is a figment of the child's imagination. Where Robovie is taken to be "telling" things, events depend on hard-wired tricks. Conversely, where the robot is "told" things, it lacks the wherewithal to detect (let alone interpret!) the activity. This has implications for the robot's perceptual and motor systems and, above all, software design. First, it shows that limits to the competence-performance models that serve von Neumann machines. If Robovie and children are to do more than set up interactional routines, software must promote the rise of strategic behaviour. Instead of tying software to function, co-action must become a basis for developing new kinds of functionality. A history of encounters can shape software design which promotes strategic social management. In this way the robot could use distinctive responding to identify human context-making moves. This might be hard-wired by programming where energetic response to pseudo-learning or personal confidence was treated as strategic and thus as carrying special value. Conversely, if robots are designed to detect strategic signalling, this will enhance their capacity for relationships. While current technology is only beginning to get robots to orient to human beliefs and feelings, relationship building can be improved if attention is given to perceptually salient, strategically important context-making.

Examination of how children respond to a friendly robot can tell us much about human nature. For this reason, we are currently developing a model of how behaviour changes over time. Among other things, we hope to program robots that integrate human 'tellings' with their repertoire. This social learning is, we believe, consistent with seeing human intelligence as based in relationships. Since Robovie is unlike us, it is extraordinary that children want to be his friend. It is also striking that they work at relationships and, showing their best side, allow a friendly machine to influence both their self-worth and the dynamics of a class-group. Such outcomes make us confident that Robovie is one of the first in a long line of friendly robots. As we come to understand more about human responses and how each party motivates the other to attune to this adjustment, such machines will change dramatically. We will find many uses for human-like robot partners that have, among other things, the powers of tomorrow's computing technology.

Acknowledgements

We would like to express gratitude to Prof. Hiroshi Ishiguro and Prof. Karl MacDorman of Osaka University for valuable advice. We wish to thank the teachers and students at the school for participation and helpful suggestions. We also thank Rumi Sato and Naoki Saiwaki, who helped with this field trial in the elementary school. This research was supported by the NICT of Japan.

References

Breazeal, C & A Scassellati, 1999. Context-Dependent Attention System for a Social Robot, *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'99)*, pp. 1146-1151.

Burgard, W, AB Cremers, D Fox, D Hahnel, G Lakemeyer, D Schulz, W Steiner & S Thrun, 1998. The Interactive Museum Tour-Guide Robot, *Proceeding. of National Conference on Artificial Intelligence*, pp. 11-18.

Cowley, SJ & D Spurrett 2003. Putting Apes, (Body and Language) Together Again. *Language Sciences*, 25: 289-318.

Cowley, SJ, S Moodley, & A Fiori-Cowley 2004. Grounding Signs of Culture: Primary Intersubjectivity in Social Semiosis. *Mind, Culture and Activity*, 11/2: 109-132.

Deacon, T 1997. *The Symbolic Species: The Co-Evolution of Language and the Human Brain*. Harmondsworth: Penguin.

Dunbar, R 1993. Grooming, Gossip and the Evolution of Language. Harmondsworth: Penguin.

Fujita, M 2001. AIBO; Towards the Era of Digital Creatures, *International Journal of Robotics Research,* 20/10, pp. 781-794.

Glenn, P, CD LeBaron & J Mandelbaum 2003. *Studies in Language and Social Interaction*. New York: Lawrence Erlbaum.

Goffman, E 1959. *The Presentation of Self in Everyday Life*. New York: Doubleday Anchor.

Goffman, E 1974. *Frame Analysis: An Essay on the Organisation of Experience.* New York: Harper and Row.

Hinde, R 1979. *Toward Understanding Relationships*. London: Academic Press.

Hinde, R 1983. *Primate Social Relationships: An Integrated Approach*. Oxford: Blackwell.

Hinde, R 1987. *Individuals, Relationships, Culture.* Cambridge: Cambridge University Press.

Humphrey, N 1976. The Social Function of Intellect. In *Growing Points in Ethology,* P. Bateson & R. Hinde (eds.), Cambridge: Cambridge University Press, pp. 303- 317.

Hutchins, E 1995. *Cognition in the Wild*. Cambridge MA: MIT Press.

Kanda, T, H Ishiguro, M Imai, T Ono, & K Mase 2002. A Constructive Approach for Developing Interactive Humanoid Robots, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2002),* pp. 1265-1270.

Kanda, T, H Ishiguro, M Imai, & T Ono, 2003a. Body Movement Analysis of Human-Robot Interaction, *International Joint Conference on Artificial Intelligence (IJCAI2003),* pp. 177-182

Kanda, T. T Hirano. D Eaton & H Ishiguro 2003b. Person Identification and Interaction of Social Robots by Using Wireless Tags, *IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp.1657-1664.

Kanda, T, T Hirano, D Eaton & H Ishiguro 2004a. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial, *Journal of Human Computer Interaction*, Vol. 19, No. 1-2, pp. 61-84

Kanda, T, R Sato, N Saiwaki, & H Ishiguro 2004b. Friendly Social Robot that Understands Human's Friendly Relationships, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2004)*, 2004. (to appear)

Kozima, H. & E Vatikiotis-Bateson 2001. Communicative Criteria for Processing Time/Space-Varying Information, *Proceedings of the IEEE International Workshop on Robot and Human Communication*.

MacDorman, KF 1999. Grounding Symbols through Sensorimotor Integration. *Journal of the Robotics Society of Japan*, 17(1): 20-24.

Moore C & PJ Dunham, (eds.) 1995. *Joint Attention: Its Origins and Role in Development*. New York: Lawrence Erlbaum.

Okuno, HG, K Nakadai, & H Kitano 2002. Realizing Audio-Visually Triggered ELIZA-Like Non-Verbal Behaviors, *PRICAI2002, LNAI 2417, Lecture Notes in Artificial Intelligence*, Springer-Verlag, pp. 552-562.

Owings, D & E Morton 1998. *The Evolution of Vocal Communication: A New Approach*. Cambridge: Cambridge University Press.

Rendall, L & H Whitehead 2001. Culture in Whales and Dolphins. *Behavioral and Brain Sciences*, 24(2), 309-382.

Ross, D & P Dumouchel, 2004. Emotions as Strategic Signals. *Rationality and Society* 16: 251-286.

Scassellati, B 2000. Investigating Models of Social Development Using a Humanoid Robot, *Biorobotics*, MIT Press.

Spurrett, D & SJ Cowley, 2004. How to Do Things without Words. *Language Sciences*, 26/5; 443-466.

Thibault, P 2000. The Dialogical Integration of The Brain in Social Semiosis: Edelman and the Case for Downward Causation. *Mind, Culture and Activity*, 7/4, 291-311.

Tinbergen, N 1952. Derived Activities: Their Causation, Biological Significance, Origin and Emancipation during Evolution. *Quarterly Review of Biology*, 27: 1-32.

Tomasello, M 1999. *The Cultural Origins of Human Cognition*. Cambridge MA: Harvard University Press.

Watzlawick, P Bavelas, J & D Jackson 1967. *The Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies and Practices*. New York: W.W. Norton.

Authors' Contact Details

Stephen J Cowley (s.j.cowley@herts.ac.uk)
University of Hertfordshire, UK &
University of KwaZulu-Natal, South Africa
Takayuki Kanda (kanda@atr.jp)
ATR Intelligent Robotics and Communication Laboratories,
Kyoto, Japan