PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2007) **362**, 679–704
doi:10.1098/rstb.2006.2004
Published online 13 February 2007

Socially intelligent robots: dimensions of human–robot interaction

Kerstin Dautenhahn*

School of Computer Science, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK

Social intelligence in robots has a quite recent history in artificial intelligence and robotics. However, it has become increasingly apparent that social and interactive skills are necessary requirements in many application areas and contexts where robots need to interact and collaborate with other robots or humans. Research on human–robot interaction (HRI) poses many challenges regarding the nature of interactivity and 'social behaviour' in robot and humans. The first part of this paper addresses dimensions of HRI, discussing requirements on social skills for robots and introducing the conceptual space of HRI studies. In order to illustrate these concepts, two examples of HRI research are presented. First, research is surveyed which investigates the development of a cognitive robot companion. The aim of this work is to develop social rules for robot behaviour (a 'robotiquette') that is comfortable and acceptable to humans. Second, robots are discussed as possible educational or therapeutic toys for children with autism. The concept of interactive emergence in human–child interactions is highlighted. Different types of play among children are discussed in the light of their potential investigation in human–robot experiments. The paper concludes by examining different paradigms regarding 'social relationships' of robots and people interacting with them.

Keywords: social robots; human–robot interaction; robotiquette; robot companion

1. INTRODUCTION: THE NATURE OF ARTIFICIAL (SOCIAL) INTELLIGENCE

Humans seem to have a particular curiosity about understanding and simulating nature in general, and, specifically, human beings. This desire has found its manifestations in a variety of 'simulacra', including moving and 'speaking' statues in Egypt *ca* 2000 years ago. Hero of Alexandria's work is an outstanding example of building highly sophisticated devices using the scientific knowledge available at that time, i.e. exploiting physics, e.g. water or vapour powering movable parts connected via ropes and levers in order to impress people by opening doors of temples, moving statues seemingly autonomously (Richter 1989). Other impressive examples of simulations of humans include the 'androids' built in the sixteenth, seventeenth and eighteenth centuries in Europe, where a variety of machines were constructed simulating human activities, such as writing, dancing or, as shown in figure 1, trumpet playing, based on delicate and sophisticated clockwork mechanisms available at that time. The design of these androids focused on human-like realistic appearance and the simulation of a few human activities, different from later research on artificial intelligence (AI), which was similarly aiming at simulating human activities, but focused on the 'mind'. Thus, instead of a realistic replication of one or very few human activities, replicating the flexibility and adaptability of human intelligence became a big challenge.

*k.dautenhahn@herts.ac.uk

One contribution of 19 to a Discussion Meeting Issue 'Social intelligence: from brain to culture'.

Since its origin, which can be dated back to 1956, AI research has been strongly inspired and motivated by human intelligence; human thinking and problem-solving dominated until the late 1980s, whereby chess-playing, theorem-proving, planning and similar 'cognitive' skills were considered to exemplify human intelligence and were proposed as benchmarks for designing systems that should either simulate human intelligence (weak AI) or become intelligent (strong AI). In this human-centred viewpoint, any creatures other than adult human beings, e.g. elephants, dolphins, non-human primates as well as three-month-old children, were not considered to be relevant subjects for the study or modelling of 'intelligence'. While progress has been made in the domains of what is considered now 'classical AI', e.g. chess-playing programs are able to beat expert human chess players, and AI technology is widely used, e.g. in e-commerce and other applications involving software agents, from the perspective of trying to understand or create human intelligence, it has become apparent that such skills are not necessarily those that 'make us human'. Also, attempts to put 'AI' on wheels, i.e. to design AI robots, illuminated a fundamental problem with the view of intelligence as 'disembodied' and 'symbolic', i.e. getting a robot to do even very 'simple things', e.g. wandering around in an office environment and not to bump into obstacles, turned out to be very surprisingly difficult. Other simple things that humans do with 'little thinking', e.g. recognizing a cup placed on a table behind a vase, grasping and carrying the cup filled with coffee to the dining room without spilling the coffee, turned out to be big scientific challenges. More fundamentally, skills involving sensing and acting and close couplings between these in order to deal with the

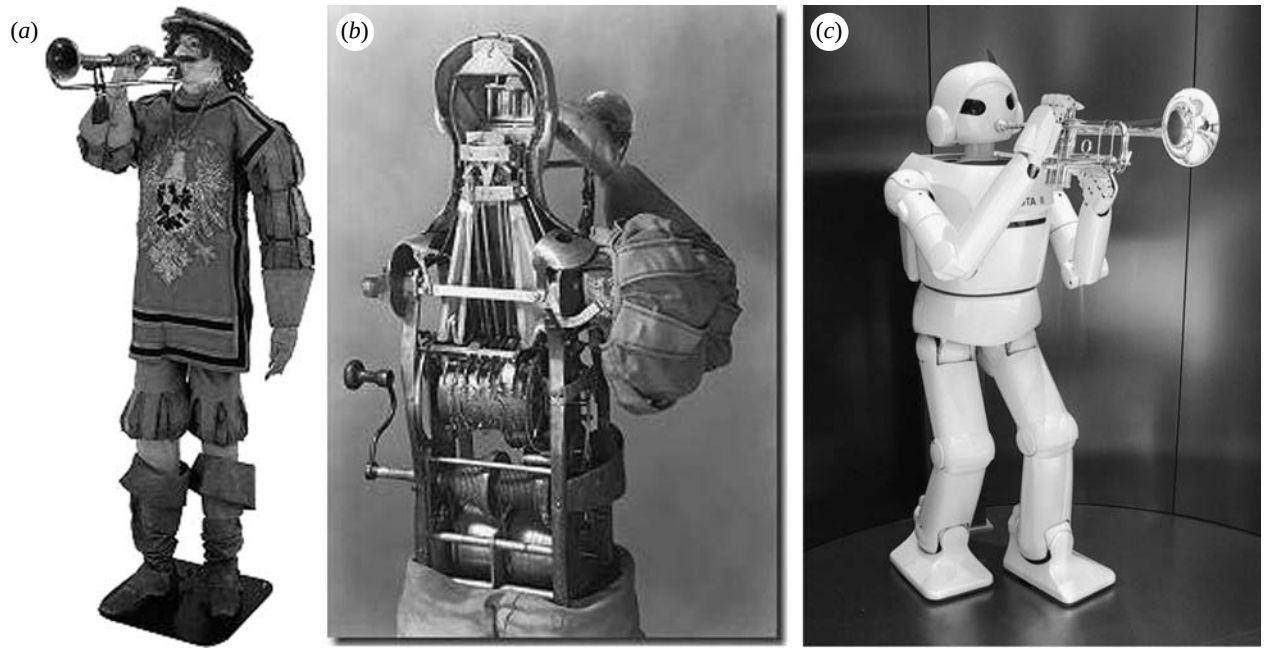

Figure 1. People have long been interested in machines that simulate natural processes, in particular machines that simulate human behaviour and/or appearance. (*a*, *b*) The famous trumpet player designed by Friedrich Kaufmann in Dresden, Germany, (source: http://www.deutsches-museum.de/ausstell/meister/e_tromp.htm; copyright Deutches Museum, Munich). A variety of other androids were created trying to simulate appearance and behaviour of humans (and other animals), based on clockwork technology available at the time. Pierre Jaquet-Droz and Jacques de Vaucanson are among the famous designers of early androids in the eighteenth century. (*c*) A recent example, using the latest twenty-first century robotics technology, to simulate aspects of human behaviour: the Toyota robot at the Toyota Kaikan in Toyota City (This Wikipedia and Wikimedia commons image is from the user Chris 73 and is freely available at http://commons.wikimedia.org/wiki/Image:Toyota_Robot_at_Toyota_Kaikan.jpg under the creative commons cc-bu-sa 2.5 license.)

dynamics and unpredictability of the 'real world' have become the new big challenges. Rather than focusing on the 'problem-solving mind', the 'mind in the body', placed in and part of a surrounding environment, became a focus of attention.

More recently, sensorimotor skills emphasizing the embodied nature of human intelligence (including locomotion, object manipulation, etc.) are considered to be the more fundamental but certainly more biologically and developmentally plausible milestones that researchers are aiming at, highlighting the close relationships between mind, body and environment, work that has been pioneered by Brooks and others since the 1980s (see collection of articles by Steels (1994), Brooks (1999) and Pfeifer & Scheier (1999)). In such a 'nouvelle AI' viewpoint, a robot is more than a 'computer on wheels', as it had been considered in AI for decades. A nouvelle AI robot is embodied, situated, surrounded by, responding to and interacting with its environment. A nouvelle AI robot takes its inspirations not necessarily from humans, i.e. insects, slugs or salamanders can be equally worthwhile behavioural or cognitive models depending on the particular skills or behaviours that are under investigation. This paradigm shift in AI had important consequences for the type of robotics experiments that researchers conducted in the field of nouvelle AI, i.e. an ecological balance between the complexity of 'body', 'mind' and 'environment' was considered highly important. The complexity of a robot's sensor system and the amount of sensory information to be processed need to be balanced and find a correspondence in a creature's ability to interact with and respond to the environment, given its particular internal goals and/or tasks imposed by its designers/experimenters. According to this approach, it is, for example, not advisable to put high-resolution sensor systems on a robot that possesses only two degrees of freedom (i.e. is able to move around on the floor). Most robotic platforms available in the 1990s, either self-built using, for example, Lego construction kits, or commercially available robots such as Kheperas, Koalas (K-Team) or Pioneers (Active Media Robotics), were restricted in their sensorimotor abilities to wander in a purpose-built arena, avoid certain obstacles and sometimes respond to certain gradients in the environments (e.g. light) via specific sensors (figure 2). The simplicity of the sensor and actuator system made many researchers focus on 'internal operations' of the robot, i.e. its control system (whether designed by the experimenter or evolved using evolutionary algorithms). A typical behaviour set of a 1990s nouvelle AI robot consists of {Wander, Avoid-Obstacle, Positive or Negative Phototaxis}. Such robotic test beds have been widely used to investigate the development of machine learning techniques applied to robot controllers, whereby the robot learns to avoid obstacles or 'find' a light source (which was often modelled as a 'food source'). Other more biologically inspired scenarios included robots in a simulated 'ecosystem' where they had to operate self-sufficiently, including recharging their batteries, or experiments inspired by swarm intelligence in social insects (e.g. Bonabeau *et al.* 1999). The kind of intelligence that these robots could demonstrate was clearly far from any behaviour considered human-like: behaviours such as wandering around in the environment and being able to respond to certain stimuli in the environment are

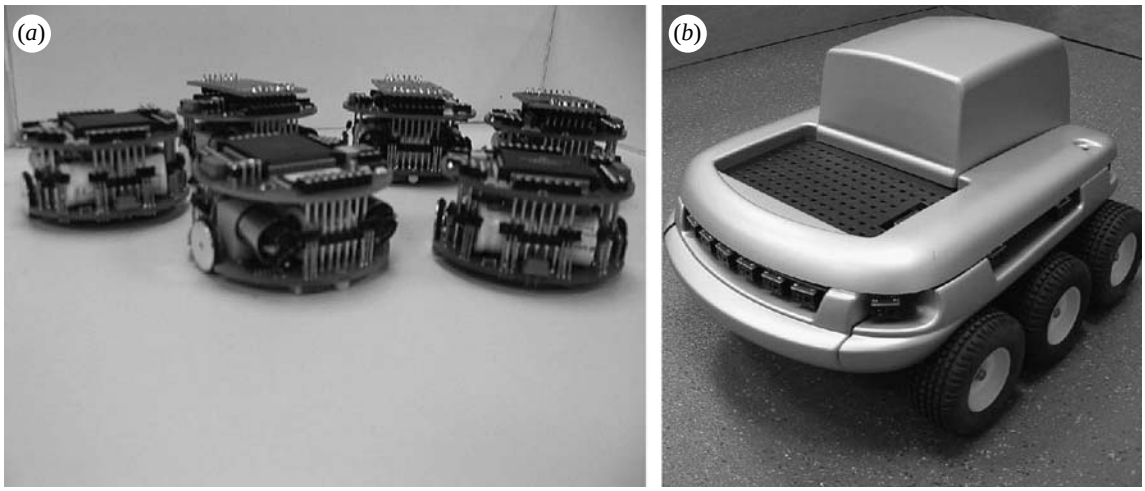

Figure 2. Experimental platforms that have been used widely in 'nouvelle AI' research: (*a*) Khepera and (*b*) Koala, both from K-Team (http://www.k-team.com). Both robots have two degrees of freedom that allow wandering in the environment. Optionally, grippers can be fitted to pick up objects.

exhibited even by bacteria. Insects, far from simple as biological complex systems, but nevertheless showing behaviour as individuals closer in magnitude to the limited scope of behaviour that can be simulated with machines available in the 1990s, became popular models for 'behaviour-based AI', the branch of nouvelle AI concerned with developing behaviour control systems for robots.

By the mid-1990s, new research initiatives took off, following, in principle, the nouvelle AI paradigm, but aiming at robots with human-like bodies and human-like minds, most famously the Massachusetts Institute of Technology robot 'Cog', an upper torso humanoid robot, later accompanied in the same laboratory by 'Kismet', a robot that consisted of an articulated face and expressed 'emotions'. Cog, in particular, was aimed at modelling, if not synthesizing, human-like intelligence, a goal too ambitious to reach within a few years. However, this initiative by Rodney A. Brooks revived an interest in studying human-like intelligence in machines (Brooks *et al.* 1999).

Despite impressive examples of sensorimotor skills in the present-day robots and some examples of social interactions of robots with other robots or people, reaching human-like intelligence remains a big challenge and the cognitive abilities of present-day robots are still limited, while research dominantly focuses on how to instantiate human-like intelligence in machines that can intelligently interact with the environment and solve tasks.

An alternative viewpoint towards AI, for which the author has been arguing since 1994 (Dautenhahn 1994, 1995, 1998, 1999*a*, 2004*a*), is to propose that one particular aspect of human intelligence, namely *social intelligence*, might bring us closer to the goal of making robots smarter (in the sense of more human-like and believable in behaviour); the social environment cannot be subsumed under 'general environmental factors', i.e. humans interact differently with each other than with a chair or a stone. This approach is inspired by the social intelligence hypothesis (also called social brain hypothesis; Dunbar 1993, 1996, 1998, 2003), which suggests that primate intelligence primarily evolved in adaptation to social complexity, i.e. in order to interpret, predict and manipulate conspecifics (e.g. Byrne & Whiten 1988; Byrne 1995, 1997; Whiten & Byrne 1997). The social intelligence hypothesis originated in the studies of non-human primates: the seminal work of Alison Jolly, who studied lemur intelligence and noted that while they lack the intelligence to learn about and manipulate objects, different from monkeys, they show similarly good social skills, led her to conclude that, 'primate society, thus, could develop without the object-learning capacity or manipulative ingenuity of monkeys. This manipulative, object cleverness, however, evolved only in the context of primate social life. Therefore, I would argue that some social life preceded, and determined the nature of, primate intelligence' (Jolly 1966, p. 506).

Thus, there may be two important aspects to human sociality: it served as an evolutionary constraint that led to an increase of brain size in primates, which in turn led to an increased capacity to further develop social complexity. The argument suggests that during the evolution of human intelligence, a transfer took place from social to non-social intelligence, so that hominid primates could transfer their expertise from the social to the non-social domain (Gigerenzer 1997). Note that for the present paper, it is not important whether the social domain was the primary factor in the evolution of primate and human intelligence, it is sufficient to know and accept that it did play an important role, possibly in conjunction with or secondary to other factors, e.g. ecological or social learning capacities (Reader & Laland 2002).

AI since its early days has tried to simulate or replicate human intelligence in computers or robots. Given what has been suggested about the phylogeny of human intelligence, whereby the importance of the social environment also becomes apparent in ontogeny, making robots social might bring us a step further towards our goal of human-style AI.

However, despite a change in viewpoint from the so-called 'classical' to the 'nouvelle' direction of AI, social intelligence has not yet been fully recognized as a key ingredient of AI, although it has been widely investigated in fields where researchers study animal and human minds. Acknowledging the social nature of human intelligence and its implications for AI is an

Figure 3. Experimental humanoid robot platform for the study of synchronization, turn-taking and interaction games inspired by child development. Kaspar, a child-sized humanoid robot developed by the Adaptive Systems Research Group at the University of Hertfordshire. (*a*) Kaspar has a minimally expressive head with eight degrees of freedom in the neck, eyes, eyelids and mouth. The face is a silicon rubber mask, which is supported on an aluminium frame. It has two degrees of freedom in the eyes fitted with video cameras and a mouth capable of opening and smiling. It has six degrees of freedom in the arms and hand and is thus able to show a variety of different expressions. (*b*) Kaspar's expressions: happy; neutral; and surprised (Blow *et al.* 2006). (*c*) Some of Kaspar's expressions using movements in the head and arms.

exciting challenge that requires truly interdisciplinary viewpoints. Such viewpoints can be found in the field of *human–robot interaction* (*HRI*), where researchers are typically addressing robots in a particular service robotics task/application scenario, e.g. robots as assistants, and where social interaction with people is necessarily part of the research agenda. However, it is still not generally accepted that a robot's social skills are more than a necessary 'add-on' to human–robot interfaces in order to make the robot more 'attractive' to people interacting with it, but form an important part of a robot's cognitive skills and the degree to which it exhibits intelligence.

Applying the 'social intelligence hypothesis' to AI implies that social intelligence is a key ingredient of human intelligence and, as such, a candidate prerequisite for any artificially intelligent robot. Research on intelligent robots usually focuses first on making robots cognitive by equipping them with planning, reasoning, navigation, manipulation and other related skills necessary to interact with and operate in the non-social environment, and then later adding 'social skills' and other aspects of social cognition. Alternatively, inspired by findings from research into social intelligence in humans and other social animals, social intelligence should be viewed as a fundamental ingredient of intelligent and social robots. To phrase it differently, developing an intelligent robot means developing first a socially intelligent robot. Particularly promising to reach the goal of (social) intelligence in robots is the research direction of 'developmental robotics' (Lungarella *et al.* 2004). Figure 3 shows a robotic platform used for the study of interaction games between robots and humans, work carried out within a European project in developmental robotics called Robotcub (http://www.robotcub.org).

In the rest of this paper, we shall illustrate work on robots that have the beginnings of rudimentary social skills and interact with people, research that is carried out in the field of HRI.

First, we discuss the dimensions of HRI, investigating requirements on social skills for robots and introducing

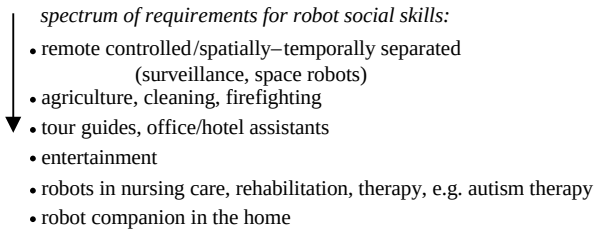

Figure 4. Increasing requirements for social skills in different robot application domains.

the conceptual space of HRI studies. Definitions of 'social robots' are discussed. In order to illustrate these concepts, two examples of research in two current projects will be presented. First, research into the design of robot companions, work conducted within the Cogniron project, will be surveyed. A robot companion in a home environment needs to 'do the right things', i.e. it has to be useful and perform tasks around the house, but it also has to 'do the things right', i.e. in a manner that is believable and acceptable to humans. Second, HRIs in the context of the Aurora project, which investigates the possible use of robots as therapeutic or educational toys for children with autism, will be discussed. The emergent nature of interactions between the children and a simple mobile robot will be discussed, emphasizing that the behaviour that might appear 'social' from an observer's point of view does not necessarily involve specific internal modelling of interaction or 'social intelligence'. The paper concludes by examining different paradigms regarding 'social relationships' of robots and people interacting with them.

2. WHAT SOCIAL SKILLS DOES A ROBOT NEED?

Investigating social skills in robots can be a worthwhile endeavour for the study of mechanisms of social intelligence, or other aspects regarding the nature of social cognition in animals and artefacts. While here the inspiration is drawn from basic research questions, in robotics and computer science, many research projects aim at developing interactive robots that are suitable for certain application domains. The classification and evaluation of HRIs with respect to the application area is an active area of research (e.g. Yanco & Drury 2002, 2004; Scholtz 2003; Steinfeld *et al*. 2006).

However, given the variety of different application domains envisaged or already occupied by robots, why should such robots, where their usefulness and functionality are a primary concern, possess social skills, given that the development of social skills for robots is costly and thus needs to provide an 'added value'? The answer to this question depends on the specific requirements of a particular application domain (see Dautenhahn 2003). Figure 4 shows a list of different application domains, where increasing social skills are required. At one end of the spectrum, we find that robots, e.g. when operating in space, do not need to be social, unless they need to cooperate with other robots. In contrast, a robot delivering the mail in an office environment has regular encounters with customers, so within this well-defined domain, social skills contribute to making the interactions with the robot more convenient for people. At the other end

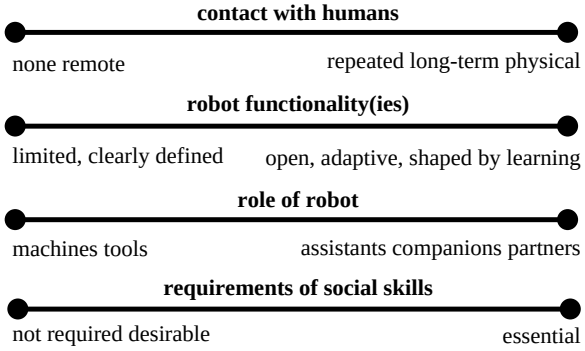

Figure 5. Evaluation criteria to identify requirements on social skills for robots in different application domains. Contact with humans ranges from none, remote contact (e.g. for robots operating in deep-sea environments) to long-term, repeated contact potentially involving physical contact, as is the case, for example, in assistive robotics. The functionality of robots ranges from limited, clearly defined functionalities (e.g. as vacuum cleaning robots) to open, adaptive functions that might require robot learning skills (e.g. applications such as robot partners, companions or assistants). Depending on the application, domain requirements for social skills vary from not required (e.g. robots designed to operate in areas spatially or temporally separated from humans, e.g. on Mars or patrolling warehouses at night) to possibly desirable (even vacuum cleaning robots need interfaces for human operation) to essential for performance/acceptance (service or assistive robotics applications).

of the spectrum, a robot that serves as a companion in the home for the elderly or assists people with disabilities needs to possess a wide range of social skills which will make it acceptable for humans. Without these skills, such robots might not be 'used' and thus fail in their role as an assistant.

In order to decide which social skills are required, the application domain and the nature and frequency of contact with humans need to be analysed in great detail, according to a set of evaluation criteria (Dautenhahn 2003), each representing a spectrum (figure 5).

3. HUMAN–ROBOT INTERACTION: THE CONCEPTUAL SPACE OF HRI APPROACHES

The field of HRI is still relatively young. The annual IEEE RO-MAN conference series that originated in 1992 in Japan and has since then travelled across the world reflects this emerging new field. HRI is a highly interdisciplinary area, at the intersection of robotics, engineering, computer science, psychology, linguistics, ethology and other disciplines, investigating social behaviour, communication and intelligence in natural and artificial systems. Different from traditional engineering and robotics, *interaction with people* is a defining core ingredient of HRI. Such interaction can comprise verbal and/or non-verbal interactions.

(a) *Approaches to social interactions with robots*

HRI research can be categorized into three, not mutually exclusive, directions, which are as follows.

— *Robot-centred HRI* emphasizes the view of a robot as a *creature*, i.e. an *autonomous entity* that is pursuing its own goals based on its motivations, drives and emotions, whereby interaction with

people serves to fulfil some of its 'needs' (as identified by the robot designer and modelled by the internal control architecture), e.g. *social needs* are fulfilled in the interaction, even if the interaction does not involve any particular task. Skills that enable the robot to 'survive in the environment' or otherwise 'fulfil internal needs' (motivations, drives, emotions, etc.) are a primary concern in this approach. Research questions involve, for example, the development of sensorimotor control and models and architectures of emotion and motivation that regulate interactions with the (social) environment.

— *Human-centred HRI* is primarily concerned with how a robot can fulfil its task specification in a manner that is *acceptable and comfortable to humans*. Here, research studies how people react to and interpret a robot's appearance and/or behaviour, regardless of its behavioural robot architecture and the cognitive processes that might happen inside the robot. Challenges include the following: finding a balanced and consistent design of robot behaviour and appearance; designing socially acceptable behaviour; developing new methods and methodologies for HRI studies and evaluation of HRIs; identifying the needs of individuals and groups of subjects to which a robot could adapt and respond; or avoiding the so-called 'uncanny valley' (Mori 1970; Dautenhahn 2002; MacDorman & Ishiguro 2006), where more and more human-like robots might appear 'unnatural' and evoke feelings of repulsion in humans. The perception of machines is influenced by anthropomorphism and the tendency of people to treat machines socially: see studies by Reeves & Nass (1996), which showed that humans tend to treat computers (and media in general) in certain ways as people, applying social rules and heuristics from the domain of people to the domain of machines. This 'media equation' they proposed (media equals real life) is particularly relevant for robotics research with the 'human in the loop', namely where people interact with robots in the role of designers, users, observers, assistants, collaborators, competitors, customers, patients or friends.

— *Robot cognition-centred HRI* emphasizes the robot as an *intelligent* system (in a traditional AI sense, see §1), i.e. a machine that makes decisions on its own and solves problems it faces as part of the tasks it needs to perform in a particular application domain. Specific research questions in this domain include the development of cognitive robot architectures, machine learning and problem solving.

Often we find an approach of decomposition of responsibilities for aspects of HRI research investigated in single disciplines and only at a later stage brought together, e.g. development of the robot's body separately from the development of the robot's 'behaviour' as it appears to humans and its 'mind'. This bears the risk of arriving at an unbalanced robot design, a 'patchwork' system with no overall integration. A synthetic approach requires collaboration during the whole life cycle of the robot (specification, design, implementation, etc.), which remains a big challenge considering traditional boundaries between disciplines and funding structures. However, only a truly interdisciplinary perspective, encompassing a *synthesis* of robot-centred, human-centred and robot cognition-centred HRIs, is likely to fulfil the forecast that more and more robots will in the future inhabit our living environments.

Defining socially acceptable behaviour, implemented, for example, as social rules guiding a robot's behaviour in its interactions with people, as well as taking into account the individual nature of humans, could lead to machines that are able to adapt to a user's preferences, likes and dislikes, i.e. an individualized, personalized robot companion. Such a robot would be able to treat people as individuals, not as machines (Dautenhahn 1998, 2004*b*). In §4, this notion of a robot companion is elaborated in more detail.

(b) *What are social robots?*

Various definitions of social robots or related concepts have been used in the literature, including the ones that are as follows.

(i) *Socially evocative*. Robots that rely on the human tendency to anthropomorphize and capitalize on feelings evoked, when humans nurture, care or involve with their 'creation' (Breazeal 2002, 2003).

(ii) *Socially situated*. Robots that are surrounded by a social environment which they perceive and react to. Socially situated robots are able to distinguish between other social agents and various objects in the environment (Fong *et al.* 2003).

(iii) *Sociable*. Robots that proactively engage with humans in order to satisfy internal social aims (drives, emotions, etc.). These robots require deep models of social cognition (Breazeal 2002, 2003).

(iv) *Socially intelligent*. Robots that show aspects of human-style social intelligence, based on possibly deep models of human cognition and social competence (Dautenhahn 1998).

Fong *et al.* (2003) propose the term 'socially interactive robot', which they define as follows.

(v) *Socially interactive robots*. Robots for which social interaction plays a key role in peer-to-peer HRI, different from other robots that involve 'conventional' HRI, such as those used in teleoperation scenarios.

Socially interactive robots exhibit the following characteristics: express and/or perceive emotions; communicate with high-level dialogue; learn models of or recognize other agents; establish and/or maintain social relationships; use natural cues (gaze, gestures, etc.); exhibit distinctive personality and character; and may learn and/or develop social competencies.

As can be seen from the above lists, the notion of social robots and the associated degree of robot social intelligence is diverse and depends on the particular research emphasis.

(c) *Relationships between HRI approaches*

Let us consider the range from a robot cognition viewpoint that stresses the particular cognitive and social skills a robot possesses, to the human-centred perspective on how people experience interaction and view the robot and its behaviour from an observer's perspective. Here, socially evocative robots are placed at one extreme end of the spectrum where they are defined by the responses they elicit in humans. In this sense, it would not matter much how the robot looked or behaved (like a cockroach, human or toaster), as long as it were to elicit certain human responses. At the other end of the spectrum, we find socially interactive robots that possesses a variety of skills to interact and communicate, guided by an appropriate robot control and/or cognitive architecture. For socially interactive robots, while internal motivations and how people respond to them are important, the main emphasis lies on the robot's ability to engage in interactions. Towards the robot-centred view, we find sociable machines, the robot-as-creature view, where a robot engages in interactions for the purpose of fulfilling its own internal needs, while cognitive skills and responses of humans towards it will be determined by the robot's needs and goals (see Breazeal 2004). Sociable robots are similar to socially intelligent robots in terms of requiring possibly deep models of cognition; however, the *emphasis* here is on the robot engaging in interactions in order to satisfy its internal needs. Socially situated robots are similarly related to the viewpoint of a robot-as-creature, but less so. Here, robots are able to interact with their social environment and distinguish between people and other agents (not as a symbolic distinction, but, for example, based on sensor information able to distinguish between humans and objects). A socially situated robot does not need to possess any model of 'social intelligence', 'social interactions' emerge from the robot being situated in and responding to its environment. Socially situated robots do not need to have human appearance or behaviour. Section 5 gives an example of a socially situated robot and the emergence of HRI games involving a robot that is not using any explicit 'social rules'. Finally, socially intelligent robots possess explicit models of social cognition and interaction and communication competence inspired by humans. Such a robot is simulating, if not instantiating, human social intelligence. It behaves similarly to a human, shows similar communicative and interactive competences, and thus is likely also to match human appearance to some degree, in order to keep behaviour and appearance consistent. The way in which humans perceive and respond to a socially intelligent robot is similarly important, since its interactions with humans model human–human interactions. Consequently, for a socially intelligent robot, robot-centred, human-centred and robot cognition-centred HRI is required. Figure 6 shows the three different views on HRI discussed in this section, highlighting the emphasis used in different approaches using different definitions of robot social behaviour and forming a conceptual space of HRI approaches where certain definitions are appropriate, as indicated.

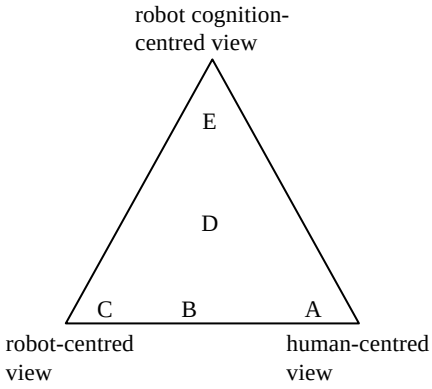

Figure 6. The conceptual space of HRI approaches. A, socially evocative; B, socially situated; C, sociable; D, socially intelligent; E, socially interactive (see text for explanations). Note: any robotic approach that can possibly be located in this framework also involves a more or less strong robotics component, i.e. the robot needs to be able to perform behaviours and tasks which can involve substantial challenges in the cases of a robot that possesses a variety of skills, as required, for example, for robots in service applications. This is less so in the cases where, for example, HRI research can be carried out with simple toy-like robots, such as Lego robots.

4. A CASE STUDY OF HUMAN–ROBOT INTERACTION: ROBOT COMPANIONS

Service robots are estimated to become increasingly prevalent in our lives. Typical tasks developed for domestic robots include vacuum cleaning, lawn-mowing and window cleaning. In robotics research, community service robotics, where robots perform tasks for people and/or in interaction with people, has become an interesting challenge: see Thrun (2004) for a discussion of past- and present-day robotics in the context of HRI.

As part of the European project Cogniron (cognitive robot companion), we investigate the scenario of a robot companion in the home, i.e. in a domestic environment shared with people. In this context, we define a robot companion as follows:

> A robot companion is a robot that (i) makes itself 'useful', i.e. is able to carry out a variety of tasks in order to assist humans, e.g. in a domestic home environment, and (ii) behaves socially, i.e. possesses social skills in order to be able to interact with people in a socially acceptable manner.

The concept of a robot companion comprises both the 'human-centred view' (it needs to perform these tasks that are believable, comfortable and acceptable to the humans it is sharing the environment with) and the robot cognition point of view: a variety of tasks need to be performed in a flexible and adaptive manner; the robot needs to adapt to and learn new and dynamically changing environments; and the overall behaviour of the robot needs to be 'consistent' (figure 7). The robot-as-creature viewpoint, e.g. how the robot can satisfy its needs, only plays a minor role.

A truly *personalized robot companion* takes into consideration an individual human's likes, dislikes and preferences and adapts its behaviour accordingly (Dautenhahn 2004*b*). Also, different people might have different preferences in terms of what tasks the robot

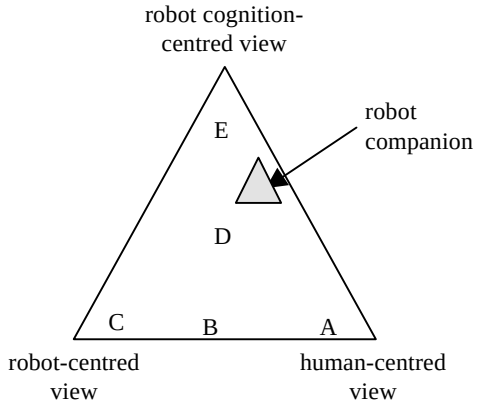

Figure 7. Challenges for a robot companion at the intersection of human-centred and robot cognition-centred views. The right balance needs to be found between how the robot performs its tasks as far as they are perceived by humans (human point of view) and its cognitive abilities that will determine, e.g. decision-making and learning (robot cognition view).

should perform or what its appearance should be like. A variety of products are on the market, which differ in appearance, usability and range of features, even for devices where the functionality seems clearly defined, e.g. cars or mobile phones. However, there is no 'one car for all drivers' and similarly we hypothesize that 'one robot for all' will not exist, i.e. will not be accepted by consumers.

What social skills does a robot companion need? Using the evaluation criteria proposed in §2, we arrive at the following characterization.

— *Contact with humans* is repeated and long term, possibly 'lifelong'. The concept of a robot companion is a machine that will share our homes with us over an extended period of time. The owner should be able to tailor certain aspects of the robot's appearance and behaviour, and likewise the robot should become personalized, recognizing and adapting to its owner's preferences. The attitude towards and opinion of such a machine will be biased by 'first impressions', but will change during long-term experiences.
— *A robot's functionality* can be limited, e.g. vacuuming or window cleaning; however, different from such single-purpose machines, a robot companion will possess a variety of skills, e.g. in addition to performing typical household tasks it will be able to communicate and interact with its users to negotiate tasks and preferences, or even to provide 'companionship'. Ideally, the machine is able to adapt, learn and expand its skills, e.g. by being taught new skills by its 'owner', and, possibly, occasional software updates. Thus, its functionality will be open, adaptive and shaped by learning.
— *The role of a companion* is less machine-like and more human-like in terms of its interaction capabilities. Rather than a machine that, if broken, is replaced, people living in the household might develop a relationship with the robot, i.e. view the companion robot as part of the household, possibly, similarly to pets.
— *Social skills* are essential for a robot companion. Without these, it will not be accepted on a long-term basis. For example, a robot that says in the morning 'good morning, would you like me to prepare breakfast for you' is interesting, but would we want this robot to say the same phrase *every* morning? Likewise, a robot that approaches and says 'would you like me to bring a cup of tea' is appealing, but would we want the robot to ask this question while we are watching our favourite television programme? Thus, social skills, the development of a robotic etiquette (Ogden & Dautenhahn 2000), or robotiquette, as a set of heuristics and guidelines on how a robot should behave and communicate in its owner's home are not only desirable, but also essential for the acceptance of a robot companion.

Within work in Cogniron on social behaviour and embodiment, the University of Hertfordshire team adopts a human-centred perspective and investigates robot behaviour that is acceptable to humans in a series of user studies, i.e. experiments where human subjects are exposed to and/or interact with a robot. The studies take place in simulated or real living rooms; experiments include laboratory studies in simulated living rooms (transformed lecture or meeting rooms) or in a more naturalistic environment, the University of Hertfordshire Robot House (a more naturalistic and ecologically valid environment, which has been found to be more suitable in order to make subjects comfortable and feel less 'assessed' and 'monitored' during the experiments). The studies were exploratory since no comparative data or theories were available which could be applied directly to our experiments. Other research groups are typically studying different scenarios and tasks, using different robot platforms with different kinds of HRI, and their results can thus not be compared directly (e.g. Thrun 1998; Nakauchi & Simmons 2000; Goetz & Kiesler 2002; Severinson-Eklundh *et al.* 2003; Kanda *et al.* 2004; Robins *et al.* 2004*a*; Kahn *et al.* 2006).

Within Cogniron, we have performed a series of HRI studies since the start of the project in January 2004. In this paper, we focus on a particular HRI study carried out in summer 2004. The robots used in the study are commercially available, human-scaled, PeopleBot robots. Details of the experimental set-up are described elsewhere (e.g. Walters *et al.* 2005*a*, 2006). Here, we briefly outline the main rationale for this work and briefly summarize the results.

In our first study in a simulated living room, we investigated two scenarios involving different tasks: a negotiated space task (NST) and an assistance task (AT). In both scenarios, a single subject and the robot shared the living room. The NST involved a robot and a human moving in the same area, which resulted in 'spatial encounters', e.g. when the robot and human were on collision course. The AT involved the subject sitting at a table being assisted by the robot, which notices that a pen is missing and fetches it (figure 8). Figure 9 shows the layout of the simulated living room. The dashed lines indicate the movement directions of the subjects and the robot. The study

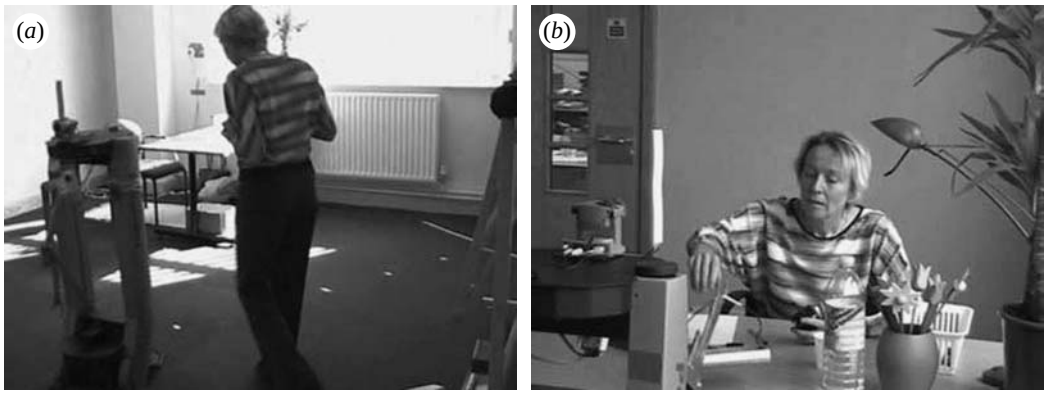

Figure 8. (*a*) Negotiated space task and (*b*) assistance task.

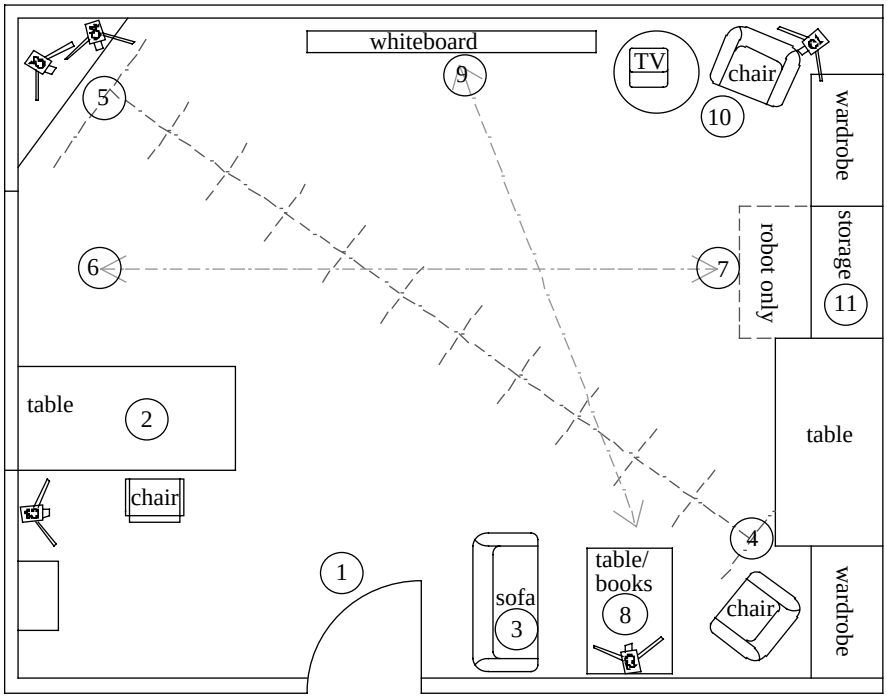

Figure 9. Layout of the experimental room for the negotiated space and assistance tasks. The room was provided with a whiteboard (9) and two tables. One table was furnished with a number of domestic items—coffee cups, tray, water bottle, kettle, etc. The other table (2) was placed by the window to act as a desk for the subject to work at while performing the assistance task, a vase with flowers, a desk light, and a bottle and glass of water were placed on the table. The room also included a relaxing area, with a sofa (3), a small chair and a low rectangular coffee table. Directly opposite, next to the whiteboard, was another low round coffee table, with a television placed on it. A second small chair stood in the corner. Five network video cameras were mounted on the walls in the positions indicated, recording multiple views of the experiment.

included 28 subjects balanced for age-, gender- and technology-related background. The robot's behaviour was partially autonomous and partially remote controlled (Wizard-of-Oz, WoZ technique; see Gould *et al*. 1983; Dahlback *et al*. 1993; Maulsby *et al*. 1993), whereby the illusion is given to the subjects during the experiment that the robot operates fully autonomously.

Each subject performed both tasks twice. The behaviour of the robot was either 'socially interactive' or 'socially ignorant'. These two robot behaviour styles were designed by an interdisciplinary research team. The selection and classification of behaviours into these two categories was done, for the purposes of this experiment, purely on the basis of what changes the robot would make to its behaviour if no human were present. If the robot performed in an 'optimal way' (from a robotics perspective, e.g. taking the shortest path between two locations), and made little or no change to its behaviour in the presence of a human, then the behaviour was classified as *socially ignorant*. If the robot took account of the human's presence, by modifying its optimum behaviour in some way, this was classified as *socially interactive* behaviour. As little was known about how the robot should actually behave in order to be seen to be socially interactive or socially ignorant, this assumption was chosen as it was in accord with what would be seen as social behaviour by the robot from a robotics perspective.

The following behaviours were classified as *socially ignorant.*

(i) When moving in the same area as the human, the robot always took the direct path. If a human was in the way, the robot simply stopped and said 'excuse me' until the obstacle was removed.

(ii) The robot did not take an interest in what the human was doing. If the human was working at a task, the robot interrupted at any point and fetched what was required, but did not give any indication that it was actively involved, or was taking any initiative to complete the task.
(iii) The robot did not move its camera, and hence its apparent 'gaze', while moving or stationary unless it was necessary to accomplish the immediate task.

These behaviours were classified as *socially interactive*, which are as follows.

(iv) When moving in the same area as a human, the robot always modified its path to avoid getting very close to the human. Especially, if the human's back was turned, the robot moved slowly when closer than 2 m to the human and took a circuitous route.
(v) The robot took an interest in what the human was doing. It gave the appearance of looking actively at the human and the task being performed. It kept a close eye on the human and anticipated, by interpreting the human's movements, if it could help by fetching items. If it talked, it waited for an opportune moment to interrupt.
(vi) When either moving or stationary, the robot moved its camera in a meaningful way to indicate by its gaze that it was looking around in order to participate or anticipate what was happening in the living room area.

During the trials, the subjects used a comfort level device, a hand-held device that was developed specifically for this experiment and used to assess their subjective discomfort in the vicinity of the robot. Comfort level data were later matched with video observations of subjects' and robot's behaviour during the experiments. Also, a variety of questionnaires were used before the experiment, after the experiment and between the sessions, with distinct robot behaviour styles, i.e. socially ignorant and socially interactive. These included questionnaires on subjects' and robot's personality as well as general questions about attitudes towards robots and potential applications. In the same experiment, other issues were investigated, including human-to-robot and robot-to-human approach distances, documented elsewhere (Walters *et al.* 2005*b*, 2006).

In this exploratory study, we addressed a number of specific research questions. These concerned the relationship between subjects' personality characteristics and their attribution of personality characteristics to the robot, including the effect of gender, age, occupation and educational background. We also investigated whether subjects were able to recognize differences in robots' behaviour styles (socially ignorant and socially interactive) as different 'robot personalities'. In the NST we were interested in which robot behaviours made subjects most uncomfortable and how robot and subjects dynamically negotiated space. In the AT we investigated which approach (robot behaviour style) subjects found most suitable. Moreover, we assessed which robot tasks and roles people would envisage for a robot companion.

The following provides a summary of some of the main results.

— *Subject and robot personality*. For individual personality traits, subjects perceived themselves as having stronger personality characteristics compared to robots with both socially ignorant and socially interactive behaviour, regarding positive as well as negative traits. Overall, subjects did not view their own personality as similar to that of the robot's, whereby factors such as subject gender, age and level of technological experience were important in the extent to how subjects viewed their personality as being dissimilar/similar to the robot personality. Overall, subjects did not distinguish between the two different robot behaviour styles (socially ignorant and socially interactive) in terms of individual personality traits (for further details, see Woods *et al.* 2005).
— *Negotiated space task*. Results show that the majority of the subjects disliked the robot moving behind them, blocking their path or on collision path towards them, especially when the robot was within 3 m proximity. The majority of subjects experienced discomfort when the robot was closer than 3 m, within the social zone reserved for human–human face-to-face conversation between strangers, while they were performing a task. The majority of subjects were uncomfortable when the robot approached them when they were writing on the whiteboard (i.e. robot was moving behind them) or trying to move across the experimental area between the whiteboard and the desk, where the books were located (figure 9; for further details, see Koay *et al.* (2005, 2006)). Note that the results from this study need to be interpreted in the context of this particular task. In other studies where the robot approached a person or a person approached a robot, most people were comfortable with approach distances characteristic of social interaction involving friends (Walters *et al.* 2005*b*, 2006). In these situations, the subjects were not interrupted by the robot and thus were probably more tolerant of closer approach distances. This issue highlights the problem of generalizing results from HRI studies to different scenarios, robots, tasks, robot users and application areas.
— *Attitudes towards robots*. The questionnaire data showed that 40% of participants in the current study were in favour of the idea of having a robot companion in the home, compared to 80% who stated that they liked having computer technology in the home. Most subjects saw the potential role of a robot companion in the home as being an assistant, machine or servant. Few were open to the idea of having a robot as a friend or mate. Ninety per cent stated that it would be useful for the robot to do the vacuuming, compared to only 10% who would want the robot to assist with childcare duties. Subjects wanted a future robot companion to be predictable, controllable,

considerate and polite. Human-like communication was desired for a robot companion. Human-like behaviour and appearance were less important (for details, see Dautenhahn *et al.* (2005)).

As can be seen from the brief description of the experiments here, the tasks for the subjects and the robot's behaviour, and the overall approach, were highly exploratory but can lay the foundation for 'robotiquette', i.e. a set of rules or heuristics guiding the robot's behaviour. However, do (aspects of) social intelligence necessarily need to be implemented in terms of specific social rules for a robot? How much of the social aspects of the behaviour are emergent and only become social in the eyes of a human observer without any corresponding, dedicated mechanisms located inside the robot?

In order to illustrate these issues, §5 investigates the case of a socially situated robot, a 'social robot without any (social) rules', whereby social behaviour emerges from simple sensorimotor rules situated in a human–robot play context.

5. EMERGING SOCIAL INTERACTION GAMES

In order to highlight the point that social behaviour can emerge without necessarily any specific 'social' processes being involved in creating that behaviour, we describe in the following how turn-taking behaviour was achieved in trials where a mobile robot interacted with children, in this particular case children with autism, as part of the Aurora project (described fully in §6). The principle of interactive emergence can also be found in other robotics works, including Grey Walter's biologically inspired famous robots, the 'tortoises', built in the late 1940s (Walter 1950, 1951). The robots called Elsie and Elmer could 'dance' with each other due to phototaxis, leading to mutual attraction with a light source attached to each robot, without any specific perception of 'the other robot' and with no special social rules implemented. More recent examples of emerging social behaviour between robots include, for example, experiments using simple robot–robot following behaviour that resulted from sensorimotor coordination and gave rise to imitation learning (Billard & Dautenhahn 1998). A similar principle of 'social behaviour without (social) rule' is embodied in Simon Penny's robot 'Petit Mal', built 45 years after the tortoises, with the specific purpose to interact with museum visitors as an artistic installation (Penny 2000).

Figure 10 shows the mobile robot used in this research. Describing the robot's control architecture goes beyond the scope of this paper, but for the purpose of this paper, it is relevant to note that the robot's behaviour was guided by the following two basic implemented behaviours:

— *obstacle avoidance*: using its infrared sensors, it causes the robot to avoid the object and move away and
— *approaching heat source*: using input from its heat sensor about directions of heat sources, it causes the robot to turn and move towards the heat source.

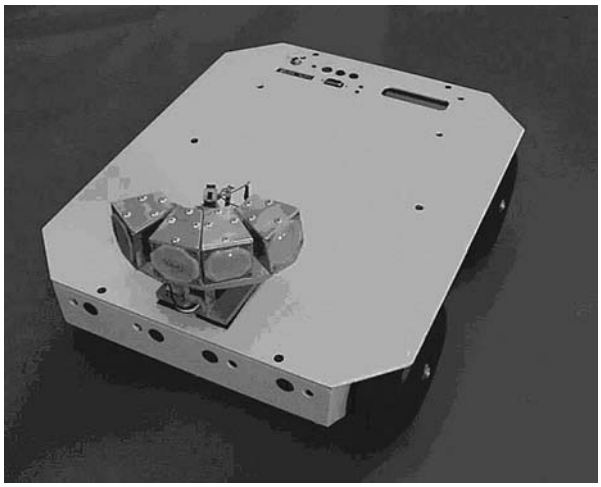

Figure 10. The Labo-1 robot used in the trials on playful interaction games with children with autism. The robot is 38× 28 cm large, 21 cm high, mass 6.5 kg and has four wheels. Its four-wheel differential drive allows smooth turning. The robot has eight active infrared sensors positioned at the front (four sensors), rear (two) and one sensor on each side. A pyroelectric heat sensor was mounted on the front end of the robot and enabled it to detect head sources. This sensor was used to detect children. A voice generation device was used optionally to create speech phrases such as 'hello there', 'where are you', 'can't see you', depending on its sensory input (e.g. whether a child was detected or not). The speech was used purely to add variety to the robot's behaviour.

Both behaviours are active at the same time and triggered by their respective sensor systems. The robot's behaviour was purely reactive, without any internal representations of the environment. At the beginning of the trials with children, the robot is placed in the centre of the room with some open space. Thus, with no obstacles or heat sources within the robot's range, it will remain motionless until it perceives either an obstacle or a heat source. Note that the heat sensor could similarly respond to a warm radiator, since nowhere within the robot's control system were the children 'recognized'. The child could interact with the robot in any position they liked, e.g. standing or kneeling in front of the robot or lying on the floor. As long as the child was within the robot's sensor range, interaction games could emerge.

Since a child, from the perspective of the robot, is perceived as an obstacle and at the same time as a heat source, these two simultaneously active processes gave rise to a variety of situations. Once the robot perceives a heat source, it will turn towards it and approach as closely as possible. While it approaches closely, the infrared sensors activate the obstacle-avoidance behaviour, so that the robot will move away from the heat source. From a distance, it can again detect the heat source and approach. This interplay of two behaviours resulted in the following situations.

(i) If the child remains stationary and immobile, the robot will approach and then remain at a certain distance from the child, the particular distance being determined by internal variables set in the computer program as well as properties of the robot's sensorimotor system. From an interaction perspective, the observed behaviour can be called 'approach'.

Figure 11. (*a*) A child with autism playing a chasing game. The boy went down on his knees, which gives him a better position facing the robot. (*b*) Another child decided to lie down on the floor and let the robot approach, resulting in turn-taking games. (*c*) A third child playing chasing games with a different mobile robot (Pekee, produced by Wany Robotics).

Figure 12. Playing turn-taking games with the robot. See text for a detailed description of this game.

(ii) If the child moves around the room without paying attention to the robot, the robot will approach and seemingly 'try to keep a certain distance' from the child. As long as the child stays within the robot's sensor range, this will result in a 'following' behaviour. From an interaction perspective, the behaviour can be called 'keeping contact'.

(iii) If the child moves around the room but pays attention to the robot in a 'playful' manner, the child might run away from the robot, waiting for it to approach again, upon approach running away again. Here, the child can be said to 'play' with the robot a game where they are being chased.

(iv) If the child approaches the robot, the robot will move away. If done repeatedly, this can cause the robot to be 'chased across the room' or even 'cornered'. While being 'chased away', the robot will remain 'focused' on the child, due to its heat sensors that cause it to turn towards the child. Here, the child plays a chasing game with the robot, whereby roles are reversed compared to (iii).

(v) Alternating phases of (iii) and (iv) can lead to the emergence of interaction games involving turn-taking (see example in figure 11*b* showing a child lying on the floor in front of the robot). The child stretches his arm out towards the robot and moves his hand towards the robot's front (where the infrared sensors are located), which causes the robot to back-up (i.e. triggering obstacle-avoidance behaviour). The robot then moves backwards, but only up to a certain distance where it again starts to approach the child (guided by its heat sensor). It approaches the child up to a point where the infrared sensors (triggered by the child's body or stretched-out hand held at the same height as the infrared sensor) cause obstacle avoidance once again. As far as approach and avoidance behaviours are concerned, we observe turn-taking in the interaction (figure 12). In this situation, the child very quickly 'discovered' how to make the robot back-up and the interaction continued for approximately 20 min until the child had to leave and go back to class.

The interactive situations (iii)–(v) described above are robust in the sense that any movements of the child that bring parts of his body closer to the robot can trigger the heat or infrared sensors: the system does not depend on the precise perception of the child's body position, location and movements. However, this robustness requires the interactive context of a child playing with the robot, or, in other words, the robot's environment must provide salient stimuli in the 'appropriate' order and with appropriate timing to which the robot can respond appropriately (according to its design). The lack of this context and the corresponding stimuli can result in 'non-social' behaviour, e.g. if the robot were placed in front of a radiator, it would approach up to a certain distance, stop and remain immobile. Also, for example, in situation (ii) described above, if the child moves around in the room too quickly so that the robot loses contact, then the robot will stop unless any other obstacles or heat sources are perceived. Exactly the same two behaviours are responsible for these non-social as well as the other socially interactive behaviours.

Situations (iii)–(v) above exemplify interactive games played between the child and the robot, representing a case of interactive emergence, defined by Hendriks-Jansen (1996), whereby 'patterns of activity whose high-level structure cannot be reduced to specific sequences of movements may emerge from the interactions between simple reflexes and the particular environment to which they are adapted'. The adaptation of the robot to the 'interactive

environment', i.e. the tuning of its sensorimotor behaviour, was done by the robot's programmer. All the situations described above depend on a variety of parameters (e.g. the robot's speed) that had to be determined in preliminary experiments. Thus, the robot's behaviour has been carefully tuned to afford playful interactions when placed in a context involving children. The robot's reflex-like programming based on the two behaviours controlling approach and avoidance was complemented by the child discovering how to interact with the robot via its two sensor systems (heat and infrared sensors) located at its front. The patterns of turn-taking and/or chasing emerged without an explicit representation in the robot's control program, no internal or external clock drove the turns and no internal goal or representation of the environment was used. The timing of the turns and chasing games emerged from the embodied sensorimotor coupling of the two interaction partners, i.e. the child and the robot in the environment. This aspect of mutual activity in interaction is reflected in Ogden *et al.*'s (2002) definition of interaction as a reciprocal activity in which the actions of each agent influence the actions of the other agents engaged in the same activities resulting in a mutually constructed pattern of complementary behaviour.

Note that turn-taking is a widely studied phenomenon in psychology as well as computer science and robotics, whereby various research questions are studied, such as the evolution of turn-taking (e.g. Iizuka & Ikegami 2004), or the design of a psychologically and neurobiologically plausible control architecture for a robot that can give rise to turn-taking and imitative behaviour (e.g. Nadel *et al.* 2004). In the above example, the robot's control program is non-adaptive; it does not learn but simply responds reactively to certain environmental stimuli. However, the very simple example above shows that very few (in this case two) carefully designed behaviours for a simple robot (simple compared to the state of the art in robotics in 2006) can result in interesting and (from the point of view of the children) enjoyable interaction games played with children. Such a bottom-up perspective on socially interactive behaviour demonstrates that for the study of certain kind of social behaviour, assumptions about the robot's required level of social intelligence need to be considered carefully. Rather than modelling the social environment explicitly in the robot's control program, placing the robot in such an environment where it is equipped with simple behaviours responding to this particular environment serves its purpose ('the social world is its own best model'). From an observer's point of view, the robot played interaction games with the children, without any explicit knowledge about turn-taking or the 'meaning' of interactions. However, as long as the robot is involved in interactions with a child, numerous hypotheses might be created about the robot's (social) intelligence. Only when taken out of the interactive context which it had been designed for and adapted to (e.g. when placed in front of a radiator) can different hypotheses be tested, in this case illuminating the basic processes driving the robot's behaviour that do not entail any social dimension.

Now, let us extrapolate this work, assuming a sophisticated robot that has been carefully designed to afford a variety of interactions. With a large number of sensors and actuators, a simple parallel execution of behaviour will not be adequate, so more sophisticated behaviour-arbitration mechanisms need to be used (e.g. as described by Arkin (1998)), and internal states may regulate its interactions with the environment. The robot's movements, timing of behaviours, etc. have been carefully designed for the specific purpose of interacting with people. Thus, it can not only approach and avoid, but also interact verbally and non-verbally in a variety of ways inspired by human behaviour (body language, speech, gestures, etc., its interaction kinesics mimicking humans). Now, let us assume that the robot is indistinguishable in its appearance from humans, it is an android (MacDorman & Ishiguro 2006). We observe the robot in different situations where it meets and interacts with people. How can we find out about the robot's level of social intelligence, whether it is 'purely a clever collection of stimulus-response behaviours' or whether it has an internal representation of 'social behaviour'? Similar to putting our small mobile robot in front of a radiator, we might test the android by exposing it to various types of social situation, attempting to see it failing, so that the nature of the failure might illuminate its (lack of) assumptions and knowledge about the world. We might design a rigorous experimental schedule, but for such a sophisticated robot, we might spend a lifetime going through all possible (combinations of) social situations. But if we are lucky, then we might see the robot failing, i.e. behaving inappropriately in a social situation. It might fail disastrously, e.g. getting stuck in front of a radiator. However, it might fail similarly to how humans might fail in certain social situations, e.g. showing signs of 'claustrophobia' on a packed underground train or expressing anxiety when being monitored and assessed in experiments. If it fails in a human-like manner, we probably consider it as a candidate machine with human-like social intelligence, or even consider that these failures or flaws might merit it to be treated as human. But we will still not know exactly what mechanisms are driving its social behaviour. However, does it matter?

6. PLAYING WITH ROBOTS: THE AURORA PROJECT

The interactions described in §5 were observed as part of research carried out in the Aurora project, which investigates the usage of robotic playmates in autism therapy. The aspect of play is a core part of the project.

Play therapy can play an important part in increasing quality of life, learning skills and social inclusion.[1] According to the National Autistic Society (NAS; http://www.nas.org.uk), the following argument can be put forward in favour of teaching children with autism to play. Play allows children to learn and practise new skills in safe and supportive environments (Boucher 1999), providing a medium through which they develop skills, experiment with roles and interact with others (Restal & Magill-Evans 1994). Children with autism are disadvantaged in their use of play for these

Figure 13. (*a*) The Robota robot and (*b*, *c*) its modified appearances used in the trials with children with autism. The robot's main body contains the electronic boards and the motors that drive the arms, legs and head (Billard 2003). A pilot study showed the use of the robot's sensing abilities and autonomous behaviour was not suitable for our trials, thus the robot has since then been used as a remote-controlled puppet (controlled by the experimenter). A summary of the use of this robot in autism therapy and developmental psychology is provided by Billard *et al.* (in press). Experiments describing how children with autism react to different robot appearances are reported by Robins *et al.* (2004*b*,*d*).

purposes. Play also matters for children with autism, because playing is the norm in early childhood, and a lack of play skills can aggravate children's social isolation and underline their difference from other children (Boucher 1999). Boucher emphasizes that play should be fun. Improving the play skills of children with autism gives them a sense of mastery and increases their pleasure and their motivation to play (that is a justifiable aim in itself). Play gives children with autism, who may have difficulty in expressing feelings and thoughts in words, chances to express themselves, and offers opportunities to engage within mutually satisfying social play, which can be used as a vehicle for developing the social skills that they so often lack. These opportunities are created by a shared understanding of pleasure experienced in play episodes (Sherratt & Peter 2002). It also prevents secondary disabilities by enabling participation in social and cultural events (Jordan & Libby 1997). Sherratt & Peter (2002) suggest that teaching children with autism to play may increase a fluidity of thought and reduce conceptual fragmentation. In particular, if play is taught to young children, it may assist them in reducing repetitive and rigid behavioural patterns and encourage communication development. Wolfberg's (1999) review of the intervention literature discovered that play (particularly with peers) has had a relatively small role in the education and treatment of children with autism.

The Aurora project investigates the use of robots in autism therapy, trying to engage children in therapeutically relevant playful interactions with a robot (involving turn-taking, imitation, joint attention and proactive behaviour), as well as using robotic toys as mediators to the social environment.

In Dautenhahn & Werry (2004), the basic motivation, starting points, related work, as well as the psychological background of this work are discussed in detail. Here, we can only provide a brief summary of the main rationale underlying the project. Literature suggests that people with autism enjoy interacting with computers (e.g. Powell 1996; Murray 1997; Moore 1998), which provides a starting point for our investigations. Robots are different from computers, since interacting with them is embodied and situated in the real world, and requires the child to involve their body in a more extensive way (compared to just operating a mouse or keyboard). Thus, the starting point of our work is the assumption that autistic children enjoy *playing* with robots. The enjoyment of the children is an important element in our work, based on our belief that interaction itself could be rewarding to the child. Typically developing children can experience social interaction as rewarding and enjoyable. Consequently, they are not just responding to other people, but they actively seek contact. It is unclear to what extent robots can teach autistic children the 'fun' of play and interaction, but it seems that a playful context is a good starting point for our investigations, similar to an approach put forward by Ferrara and Hill for language therapy:

> …A more appropriate starting place for therapeutic intervention with autistic children might be to focus on their development of social play. Social objects with low intensity should first be presented in a game that has a highly predictive and repetitive sequence of activities. Complexity of social stimuli and game activities should gradually increase in intensity. When the child begins to show pleasure in these games and to initiate them, the introduction of language and cognitive tasks matching the complexity of the game would be appropriate.
>
> (Ferrara & Hill 1980, p. 56)

Robots have been proposed to be used for the study of child development (Michaud & Caron 2002; Michaud *et al.* 2005) or rehabilitation (Plaisant *et al.* 2000), autism therapy (Weir & Emanuel 1976; Dautenhahn 1999*b*; Werry & Dautenhahn 1999; Kozima 2002; Michaud & Théberge-Turmel 2002; Davis *et al.* 2005; Kozima *et al.* 2005) and autism diagnosis (Scassellati 2005). For a critical discussion of using robots in autism therapy, see Robins *et al.* (2005*a*).

The use of robots in autism therapy poses many challenges. Potentially different solutions might prove suitable for different groups of children with different

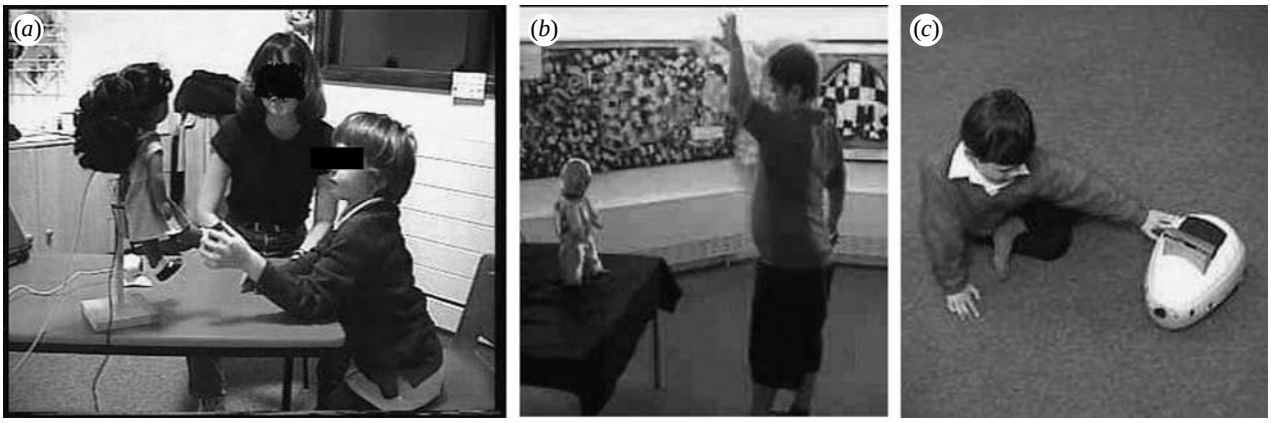

Figure 14. Children with autism playing with a robot. (*a*) The picture shows a child interacting with Robota, the humanoid robot doll, playing an imitation game, whereby the robot can imitate arm movements when the child is sitting opposite the robot facing it and moving its arms in explicit ways that can be recognized by the robot (Dautenhahn & Billard 2002). Shown are the robot, the child and a carer providing encouragement. This was the first trial using Robota for playing with children with autism. Owing to the constrained nature of the set-up, required by the limitations of current robotic technology, we decided not to use this robot any longer as an autonomously operating machine, and later only used the robot remotely controlled by the experimenter (out of the children's sight). In one of the experiments, we varied the appearance of the robot and found that the children's initial response towards a 'plain-looking' robot is more interactive than towards the robot with its doll face (Robins *et al.* 2004*b*; figure 13). (*b*) The picture shows a child with autism playing imitation games with the robot. Note the completely unconstrained nature of the interactions, i.e. the child himself had decided to move towards and face the robot, on his own terms, after he had become familiar with the robot (as part of a longitudinal study; Robins *et al.* 2004*c*, 2005*b*). (*c*) The picture shows an autistic child playing with a mobile robot (Pekee, produced by Wany Robotics). The advantage of small mobile robots is that it allows the children to move around freely and adopt different positions, e.g. lying on the floor, kneeling, walking, even stepping over the robot, etc., exploring the three-dimensional space of potential interactions. In experiments with typically developing children, we developed a technique that can classify interactions of the children with the robot, using clustering techniques on the robot's sensor data. We were able to identify different play patterns that could be linked to some general activity profiles of the children (e.g. bold, shy, etc.; Salter *et al.* 2004). We will use this technique in the future to allow the robot to adapt to the child during the interactions (Salter *et al.* 2006; François *et al.* in press). In principle, this technique might also be used to assess the children's play levels or, possibly, for diagnostic purposes.

abilities and personal interests. The particular therapeutic issues that should be addressed are also likely to influence the choice of robotic designs used.

So far, we have been using two types of robots: a small humanoid robotic doll and mobile robots (figures 13 and 14). The robots are small and safe for children to use. Our initial trials confirmed that autistic children generally take great interest in the robots and enjoy playing with them. Many children smiled, laughed or showed other signs of enjoyment during the interactions with the robot. We also observed vocalizations and verbalizations addressed to the robots. We use this playful scenario as a context where the children can be engaged in therapeutically or educationally useful behaviour. Encouraging proactiveness of behaviour in autistic children is one of the major goals of the Aurora project. Addressing deficits in turn-taking and imitation skills are additional goals. Importantly, our motivation is not to develop the robot as a replacement for teachers, other caretakers or people in general. The Aurora viewpoint is that robots should mediate between the (from an autistic child's perspective) widely unpredictable world of 'people' and the much more predictable world of machines. However, due to their situatedness and embodiment, the robots never behave completely predictably. This is an important issue, since otherwise the robots would only perpetuate repetitive or stereotypical behaviour. The purpose of our robots is to help autistic children to better understand and interact with other people. So far, no clinical trials have been carried out regarding the therapeutic impact of interaction with the robot. However, results are encouraging, in particular those gained in a longitudinal study, which exposed children with autism to the robot repeatedly over several months.

We study robots with different appearances and capabilities in order to facilitate, investigate and compare different types of interactions. We adopted an approach where the children can freely interact with the mobile robots, e.g. while playing with the robot, the children can sit on the floor, move around the robot, touch the robot or simply stand in a corner of the room and watch the robot. Figure 15 shows that children thus use their whole body in interactions with the robot.

The mobile robots we have been using are *Labo-1* (donated by Takashi Gomi from Applied AI Systems) and *Pekee* (produced by Wany Robotics). The mobile robots are programmed so that the children can play simple interaction games with them, such as chasing, following and other simple turn-taking games. In a comparative study involving children with autism playing with Labo-1 as well as a non-robotic passive toy of the same size (figure 16), we found that children with autism pay significantly more attention to the robot and direct more eye gaze at the robot (Werry & Dautenhahn, in press).

The small humanoid robot Robota (Billard 2003) is a doll shaped versatile robot that can move its arms, legs and head, in addition to having facilities for vision, speech and for producing music. However, interactions with Robota, e.g. whereby the robot can imitate children's arm movements, are more constrained,

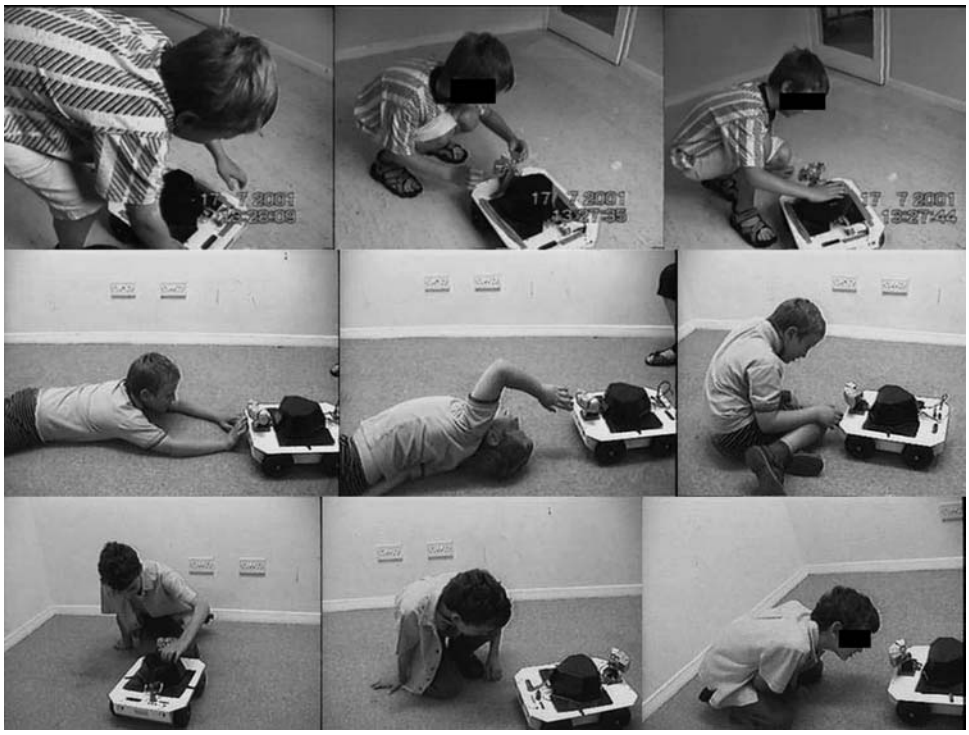

Figure 15. Varieties of interactions in playful encounters of children with autism and a mobile robot. Note the embodied nature of the interactions: children are using a variety of different postures and movements in order to play with the robot.

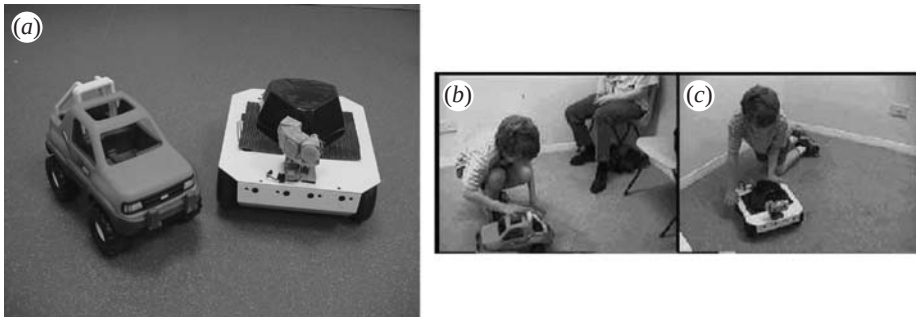

Figure 16. (*a*) Toy truck on the left and Labo-1 robot on the right. (*b*) Child interacting with the toy truck and (*c*) the robot. This comparative study involved 17 children with autism between 6 and 9 years old. Trials lasted approximately 10 min, i.e. the children interacted with the robot for 4 min. Next, both toy truck and robot were present for 2 min (robot switched off). Then, children played with the toy truck for 4 min. The order of presenting the toy truck and the robot was randomized. The toy truck (robot) was hidden during interactions with the robot (toy truck). Owing to the nature of our approach, we stopped a trial when a child seemed to become bored, distressed or wanted to leave the room. Interactions with the toy truck involved a lot of repetitive behaviour, e.g. spinning the wheels, pushing it against a wall; all children very quickly lost interest. Interactions with the robot were much more 'lively', the children were more engaged and played with the robot longer than with the toy truck. Note that any statements we make about the engagement of the children have been confirmed by teachers, carers and autism experts watching the videos with us. Differences were confirmed in behavioural coding of attention and eye gaze in both conditions. Details of this work are reported by Werry (2003) and Werry & Dautenhahn (in press).

i.e. they require the children to sit at a table and face the robot (Dautenhahn & Billard 2002). Consequently, we performed a series of trials using the robot as a puppet controlled by an experimenter. This approach turned out to be very successful and resulted in a series of trials where children demonstrated interactive and communicative competencies, using the robot as *mediator* in order to interact with the experimenter or other children (Werry *et al*. 2001; Dautenhahn 2003; Robins *et al*. 2004*a*, 2005*c*), as shown in figure 17. Generally, the experimenter played an important part in these trials, very different from other experimental set-ups, which try to remove the experimenter as much as possible. Instead, the experimenter, who in our case is an experienced therapist as well as computer scientist, played a crucial part in how he remotely controlled the robot, being sensitive to the children's behaviour and the overall context (Robins & Dautenhahn 2006; figure 17, top row).

Our general set-up of the trials is very playful; the children are not required to solve any tasks other than playing, and the only purpose of the robot is to engage children with autism in therapeutically relevant behaviours, such as turn-taking and imitation. A key issue is that the children proactively initiate interactions rather than merely respond to particular stimuli. Additionally, the chosen set-up is social, i.e. it involves not only the robot and the autistic child present, but

Figure 17. The robot as a mediator. Top: examples of children with autism interacting with the experimenter, with the robot acting as a mediator (Robins *et al.* 2005*b*). Middle and bottom: the robot as a mediator facilitating interactions between children with autism. These examples emphasize the ultimate goal of the Aurora project, namely to help children with autism to connect to the social world of humans, and not necessarily to bond with robots.

Figure 18. A pair of children with autism simultaneously playing with the Labo-1 robot.

also can include other children, the teacher or other adults. This social scenario is used by some children in a very constructive manner, demonstrating their communicative competence, i.e. they use the robot as a focus of attention in order to interact and/or communicate with other people in the room. Trials with a mobile robot included pair-trials, where pairs of children were simultaneously exposed to the robot (Werry *et al.* 2001; figure 18). Table 1 presents different play styles that could be observed, ranging from social

Table 1. Summary of trials where pairs of children played with the robot (see details in Werry *et al.* (2001)). (The children were paired up by the teachers according to mutual familiarity and social/communication abilities. Although only three pairs were studied, interesting differences in how the children are playing with each other and the robot can be identified. This case study clearly shows how a robot can mediate play and interactions among children with autism. Further studies need to investigate whether the robot can improve children's social play skills. The table is modified from Dautenhahn (2003).)

social abilities in social interactions with people	level of interest in robot	play style	child behaviours directed at robot
highest (interacting with people)	high interest in robot—a very strong focus of attention and curiosity; good verbal and communication skills used to interact with adults (robot-centred communication, e.g. asking questions about what robot can do)	social play: children were playing and communicating with each other, the robot, and the adults present; social learning/teaching: in one instance, the experimenter instructed one child how to chase the robot, the child then instructed the other child	exploratory and interactive play with robot (touching, operating, etc.), interest in chasing, following a robot's speech, questioning experimenters about robot's skills, great interest in 'what robot should do', etc.
medium (few interactions with people)	varied interest in the robot; examples: one child was more interested in the car park visible through the room's window than in the robot, and another child interacted with the robot, but treated other people present in the room as an 'audience'	non-social play: children were playing with robot simultaneously, but not playing with each other, 'accidental' interaction when both competed for robot's 'attention', the teacher occasionally had to give guidance/calm children down when children grew bored	interest in robot's destructive skills, cornering robot, shouting at robot to get it to move, operating robot (touching, manipulating)
lowest (withdrawn, very little interaction with people)	both children showed great interest in playing with the robot	non-social play: one child dominates the interaction, an open competition for robot emerges, occasional use of communication skills in trying to obstruct the other child, e.g. 'leave him alone'	giving vocal commands and directions to robot, operating robot

to non-social and competitive play. Similarly, trials involving the small robot doll and pairs of children elicited a variety of interactions among the children (Robins *et al.* 2005*c*).

An important part of our work is the development of appropriate scenarios and techniques in order to evaluate details of robot–child interactions. We developed a technique that can be used to quantitatively evaluate video data on robot–child interactions. The technique is based on micro-behaviours (Tardiff *et al.* 1995), including eye gaze and attention. We used this technique in the comparative study mentioned above, where we studied how autistic children interact with a mobile robot as opposed to a non-robotic toy (Werry 2003; Werry & Dautenhahn, in press). The same technique has also been used in analysing interactions of children with autism with the small robotic doll (e.g. Robins *et al.* 2004*c*, 2005*b*). A range of different qualitative as well as quantitative evaluation techniques are likely to be needed in order to reveal not only statistical regularities and patterns, but also meaningful events of behaviour in context. An application of conversation analysis has revealed interesting aspects of how the robot can elicit communication and interaction competencies of children with autism (Dautenhahn *et al.* 2002; Robins *et al.* 2004*a*). Thus, using robots in autism therapy and education poses many challenges. Our work, grounded in assistive technology and computer science, is highly exploratory in nature: new robot designs, novel experimental set-ups and scenarios and new experimental paradigms are investigated; also, various evaluation techniques, known from ethology and psychology, need to be adapted to the specific context of this work. Aurora is a long-term project the author has been pursuing since 1998. Results do not come easily, but the work remains challenging and rewarding; the enjoyment of the children is the best reward.

From a conceptual viewpoint, which types of play could any type of robot encourage with children with autism? Studies of play and social participation lead to conceptual distinctions suggested by Parten (1932; table 2). With respect to these categories, in the context of using robots as assistive technology, what type of social participation can occur and/or can be encouraged? Table 2 provides suggestions on possible studies involving robots and children. The range of possible play scenarios involving robots and children is potentially huge. For children with autism, as well as other children, robots are fun to play with. Thus, robots seem a promising tool to teach children with autism how to play in a way that might integrate them better in groups of typically developing and autistic children.

The Aurora project is an example of HRI research, where human–robot as well as human–human relationships matter. Possible relationships between robots and people are discussed in more detail in §7.

Table 2. Types of play and social participation according to Parten (1932) and applied to robot–child play.

types of play and social participation	characterizations for activities in children	possible studies involving robots and children with autism
unoccupied behaviour	child is apparently not playing, observing any activities that might be exciting, otherwise moves their body around or glances around the room	will be observed if a child is not interested in the robotic set-up at all and/or when other stimuli present are stronger and more attractive to the child; we can expect their will look at the experimenters or other adults/children present, out of the window, etc; in one of the trials where pairs of children played with the Labo-1 robot (see table 1 and figure 18), we observed this behaviour with an autistic boy who loved cars, in a situation where the experimental room had a window facing the car park; instead of playing with the robot, let alone the other child, he decided to watch the car park during the trials, despite encouragement from the teacher present to investigate the robot
onlooker	child watches other children play from close distance; often talks to the children who are playing, but does not enter the game	if a situation involves a robot and two or more children, this situation might arise when one or more children play with the robot and another child or other children are 'onlookers'; depending on the interest of the children, the robot might be used as a motivational object to 'join in'.
solitary independent play	child plays alone and independently with toys different from those used by other children who can be in the vicinity, child pursues their own activity independently of what others are doing	this is the typical situation in a scenario involving one child and the robot, unless the set-up facilitates 'mediation', i.e. allowing contact to be made to the experimenter or other children, as discussed above; it might also happen in a set-up involving one robot and two children; in this case one child withdraws (or is made to withdraw by the other child; see table 1, which mentions an instance of competitive play) and the other child is left alone playing with the robot; note: in trials with pairs of children described above (Werry *et al.* 2001), we also identified non-social, competitive play, i.e. children competed with each other for access to the (robotic) toy; this can be expected, in particular, for autistic children with low social skills
parallel activity	child plays independently but with similar toys as children in the vicinity use, plays beside other children, no attempt to influence or modify the other children's behaviour, does not try to engage with them	let us assume a hypothetical experimental set-up involving two robots and two children; here, both children might then play 'on their own' with the robot, manipulating objects, etc., beside the other child, without any connection between their play; let us now consider a modified scenario involving one mobile robot whereby parts of the robot can be controlled independently by two children, e.g. one child can control the robot's movements on the ground and the other can control the robot's arm that can pick up objects; we might observe parallel play; however, the situation might afford coordination of the children's activities and thus lead to cooperative play; see below
associative play	child plays with other children, borrowing and loaning material, but no coordination or organization of activities around 'goals' or materials, each child acts as they like; all children in a group will be engaged in very similar or identical activity, which is the focus of the conversation	in the case of two children playing with two robots, the children might coordinate to play the same game with the robot, e.g. a chasing game as the common goal, but each child is playing it on his own

(*Continued.*)

Table 2. (*Continued.*)

types of play and social participation	characterizations for activities in children	possible studies involving robots and children with autism
cooperative or organized supplementary play	child plays in group organized for some purpose: achieving some competitive goal, making a product, playing a formal game, etc; strong sense of belonging to the group; one or two members direct activities of others, roles are assigned, division of labour so that children's activities complement each other to achieve a common goal	children play truly 'together', e.g. assign roles to each other, divide the labour into different jobs, try to achieve the goal together; based on the scenario involving a robot controlled by two children, in this situation, the children might pay attention to what the other is doing, thus coordinating the robot's movements on the ground and its arm movements; a common goal could be, for example, to arrange objects and build structures together, etc; through coordination, they could achieve results that none of them could achieve alone: e.g. picking up objects and building a tower, or collecting objects in the play area and arranging them into a pile; if competition is considered to be useful, a game could be introduced, e.g. in the case of robots that can manipulate objects building a tower as high as possible; in the case of a mobile robot, 'herding' the robot together could be common goal; investigating how the robot responds or 'plays' with each child could be another goal the children could explore

7. DIFFERENT PARADIGMS REGARDING HUMAN–ROBOT RELATIONSHIPS IN HRI RESEARCH

In this section, two paradigms regarding the relationships between humans and robots in HRI research are distinguished as follows:

— the caretaker paradigm and
— the assistant/companion paradigm

(a) *The caretaker paradigm in human–robot interaction*

> This paradigm considers humans as caretakers of robots: the role of the human is to identify and respond to the robot's emotional and social 'needs'. The human needs to keep the robot 'happy' which implies showing behaviours towards the robot characteristic of behaviour towards infants or baby animals.

In this approach, humans interacting with the robot are expected to adopt the role of a 'caretaker' for the robot, which is considered an 'artificial creature'. In this robot-centred view, the human needs to identify and respond to the robot's internal needs, e.g. by satisfying its 'social drives'. This approach is clearly demonstrated in Breazeal's (2002) work on Kismet, a robotic head with facial features. The robot is treated as a 'baby infant' or 'puppy robot' with characteristic specific and exaggerated child-like features satisfying the 'Kindchenschema' (baby pattern, baby scheme, *schéma* ≪*bébé*≫). The Kindchenschema is a combination of features that are characteristic of infants, babies or baby animals, which appeals to the nurturing instinct in people (and many other mammals) and trigger respective behaviours. The concept of the Kindchenschema goes back to the ethologist Lorenz (1971), who claimed that when confronted with a child, certain social behaviour patterns involved when 'caring for the young' are released by an innate response to certain cues typically characterizing babies. These cues include, for example, a proportionally large head with a protruding forehead, large ears and eyes below the midline of the head, small nose, and, generally, a rounded body shape. Such a set of characteristics has been exploited widely in the toy market, for comics, and also recently for computer animated characters. The young ones of many animals, most notably mammals, and even the adults of certain species (e.g. bear, squirrel, dolphin) show certain of these features that make them more attractive to people than other animals that do not show these features.

Note that eliciting social responses towards artefacts does not necessarily require implementation of features of the Kindchenschema. Braitenberg (1984) discussed in a series of thought experiments that people would attribute goals, intentions and even emotions to his vehicles, whose behaviour was guided by simple sensorimotor couplings, e.g. a robot that would go towards a light source whenever it could perceive it might be interpreted as 'liking' the light source, a different robot avoiding the light source and driving away at high speed might be said to 'be afraid' of it. Seminal experiments by Heider & Simmel (1944) had already demonstrated the tendency of subjects to interpret the behaviour of moving geometric shapes on a screen in terms of intentionality. With regard to computers, Reeves & Nass (1996) provided a powerful argument for what they labelled the media equation: media = real life—'People treat new media like real people…. People confuse media and real life…. People's responses to media are fundamentally social and natural…. Media are full participants in our social and natural world'. Indeed, experiments show that computers even elicit cultural and gender stereotypes; they evoke emotional and other responses as part of our

natural social pattern of interaction (see a fuller discussion of these issues by (Dautenhahn 2004*a*)).

However, similarly to a child who might pretend that the wooden stick in his hand is a sword but clearly knows that the object is not a real sword, we *know* that computers are not people: we easily dispose of them, we clearly do not treat them as family members or friends, etc. Work by Kahn *et al.* (2006) has shown that while children treat an AIBO robot in many ways like a real dog and interact with it socially, they do not perceive it exactly like a real dog, e.g. do not attribute moral standing to it. Thus, while we are building (some kind of) social relationship with technological devices, including computers and robots, do we *really* want to *bond* with computers?

According to Dunbar and the social brain hypothesis (Dunbar 1993, 1996, 1998, 2003), one important factor in the evolution of primate brains and primate intelligence was the need to deal with increasingly complex social dynamics (primate politics), which justifies the 'expensive' nature of large brains, i.e. approximately 20% of our resting energy being used to keep the brain operational. Also, Dunbar argues that human language has evolved as an efficient means of social bonding (2.8 times more efficient than physical grooming used by non-human primates as a bonding mechanism). However, humans need to be selective regarding how many 'friends' they have: according to Dunbar there is an evolutionary constraint, a cognitive limit of 150 on the number of members of our social networks (individualized relationships, not counting 'anonymous' contacts we could potentially build, e.g. via email with strangers). Thus, as I have argued in more detail elsewhere (Dautenhahn 2004*a*), robots (or virtual 'pets') trying to be our friends, and requiring us to treat them like friends, might overload our cognitive capacities.

Moreover, maintaining good relationships with family and friends does not come 'for free', it involves certain efforts, which include the following:

— *Emotional investment* in our children, other family members and friends (implies not only fun in interaction, but also entails expectations, commitments, concerns, disappointments, etc.).
— *Psychological investment.* Although for most people applying a theory of mind, perceiving and expressing empathy, paying close attention to others' needs, reading others' behaviour and identifying subtle cues that might be important in regulating interactions, listening and interpreting language, coordinating turns in interaction, managing cooperation, handling arguments and discussions in dyadic and group situations, memorizing interaction histories, etc., seem to come 'naturally', many of these skills require (modification by) learning during early socialization and development; thus they come at a psychological 'cost'. For example, how much empathy can we express in one day? A 'mechanical', psychologically/emotionally detached response could be given easily. However, empathizing in the sense of re-experiencing and relating to own and other's experiences requires far more effort. Medical staff can adopt a 'professional attitude', but with the real danger of looking at the 'patient', rather than the 'person'; they have been criticized for this lack of empathy. Recently, state-of-the-art virtual environment technology has been used to foster empathy with patients, allowing medical doctors to see and experience the world through their patients' eyes and body, continuing other activities which have been ongoing for several decades aimed at fostering empathy using interpersonal communication, and also empathy training for medical students and doctors (Kramer *et al.* 1989; Evans *et al.* 1991). The need for such types of training highlights the problem of empathizing with strangers very frequently, even after short-term contact, and on a daily basis. If empathy came for free, it would not be considered demanding or exhausting to communicate in an 'empathic' way with dozens of patients each day.
— *Physiological investment.* Communication and interaction are energy consuming, and speaking, gesturing, an extended period of firm concentration, etc., are physical activities. Giving a keynote speech can be as exhausting as chasing behind a bus.

Elaborating the above points in great detail would go beyond the scope of this paper. The main purpose is to indicate that social interaction involves emotional, physical and physiological activities that have a cost. What do we get in return from interactions with humans or other social animals (e.g. dogs)? The answer is that usually we gain *a lot*, e.g. emotional support from family, friendship, love and companionship, apart from the fact that cooperation is a core ingredient of, and makes possible, human culture. However, humans, dogs and other biological organisms, which we might consider our friends are sensitive beings, i.e. we 'invest' in them and the return is 'real' (as far as one can tell that emotions, love and friendship really exist).

How about possible investments towards robots? Social interaction and communication with robots are costly too. If humans are expected to interact with robots similarly to human friends or children, then these costs will also occur in HRI. Do we want to make the same investments in robots that we make, for example, in our friends or children? Do we want to worry about how to fulfil our robots' emotional and social needs? Do we get the same 'reward' from an infant robot smiling at us compared to a child (assuming that for the time being, we are still able to clearly distinguish between robots and humans in face-to-face interaction; cf. discussion of robot/human indistinguishability by Dautenhahn (2004*a*) and MacDorman & Ishiguro (2006)). Is a robot really 'happy' when it smiles, or are robot emotions simulated or real? Can mechanical interactions be as rewarding as those with biological organisms? Do we get the same pay-off from HRI as from human–human interaction in terms of emotional support, friendship and love? Answers to these questions are likely to be culturally dependent, as well as specific to certain application areas (e.g. medical benefits might outweigh other concerns). Is it ethically justifiable to aim to create robots that people bond with, e.g. in the case of elderly people or people with special needs?

(b) *The companion paradigm in human–robot interaction*

> This paradigm considers robots as caretakers or assistants of humans: the role of the robot is to identify and respond to the human's needs, primarily in the sense of assisting in certain tasks. The robot needs to ensure the human is satisfied and happy (with its behaviour), which implies showing behaviours towards the human that are comfortable and socially acceptable considering a particular user.

In §4 the concept of a robot companion was discussed in more detail. The companion paradigm emphasizes the assistant role of a robot, i.e. a useful machine, able to recognize and respond to a human's needs, trying to be useful. A companion robot assisting a person in everyday environments and tasks adopts a role similar to that of personal assistants or butlers, consistent with our results reported in §4. Important characteristics for such a robot are to be considerate, proactive and non-intrusive, to work towards a relationship of trust and confidentiality with the human, to possess 'smooth' communicative skills, to be flexible, willing to learn and adapt, and be competent.

Note that this is different from a 'master–slave' metaphor of human–robot relationships. Relationships with robots were also an issue in Karel Čapek's famous play RUR (Rossum's Universal Robots), which premiered in 1921 and introduced the word 'robot'. Here, robots were 'artificial people', machines that could be mistaken for humans (thus more closely related to present-day work on androids; MacDorman & Ishiguro 2006). The play introduces a robot factory that sells these human-like robots as a cheap labour force, while later the robots revolt against their human masters, a favourite scenario in the science fiction literature and movies, but a highly unlikely scenario from a robotics point of view.

The notion of a 'robot companion' emphasizes primarily its usefulness for people, as well as the robot's 'benign' behaviour. In this way, the approach pursued in the Cogniron (§4) and Aurora projects (§6) is consistent with a companion approach.

8. CONCLUSION

This paper provides an introduction to HRI research in the context of human and robot social intelligence, developing a conceptual framework and using two concrete HRI projects as case studies in order to illustrate the framework. Different definitions of social robots and viewpoints have been discussed, emphasizing different aspects of robot cognition and human responses and attitudes towards robots. The discussion highlighted that HRI studies and experiments on social robots address fundamental issues on the nature of social behaviour and people's (experimenters' as well as users') view of robots. Any particular project in the area of HRI could identify its fundamental research goals and aims in the context of this framework.

Two examples of HRI studies have been presented. Research into a robot companion, meant to become a service robot in the home, aims at developing explicit social rules (a robotiquette) which should allow people to interact with robots comfortably. This approach is different from developing robots as therapeutic 'playmates' for children with autism. Here, the concept of interactive emergence has been highlighted, whereby turn-taking games emerge in play between the children and a simple robot that only possesses very basic behavioural 'rules', but appears social when situated in a play context. In the latter case, the social rules are implicit, emerging from the interactions, while based on the careful design of the robot's sensory and behavioural repertoire.

HRI is a growing but still young research field. The future will tell whether it can develop into a scientific field that will have its long-lasting place in the scientific landscape. Several challenges need to be faced, most prominently, those that follow.

(i) Future research in HRI needs to build a foundation of theories, models, methods, tools and methodologies which can advance our understanding of HRI and allow experiments to be replicated by other research groups. At present, results are difficult to compare across experiments due to the impact of a robot's behaviour, appearance and task, as well as the interaction scenarios studied, as mentioned in §4. Any particular HRI study can only investigate a small fraction in the huge design space of possible HRI experiments. But without a scientific culture of being able to replicate and confirm or refute other researchers' findings, results will remain on the level of case studies.

(ii) New methodological approaches are needed. Many useful inspirations can be derived from the study of animal–animal or human–human interactions in ethology, psychology and social sciences. Similarly, the field of human–computer interaction can provide starting points for the design and analysis of HRI experiments. However, *robots are not people*. In interactions with machines, humans use heuristics derived from human–human interaction (Reeves & Nass 1996), which gives us interesting insights into the 'social heritage' of our intelligence. However, people do not treat machines identically to human beings (e.g. we do not hesitate to replace our broken or insufficient laptop with a new one). Thus, care needs to be taken when adopting methodologies, for example from social sciences, and apply them unchanged to HRI studies. Also, *robots are not computers*, either. Interacting with physically embodied and socially situated machines is different from interaction via computer interfaces. Other fields can provide important input to HRI methodologies, but a range of novel methodologies are necessary in order to advance the field, and researchers in HRI have indeed started to take the first steps (e.g. Robins *et al.* 2004*d* or Woods *et al.* 2006*a*,*b*).

HRI is a highly challenging area that requires interdisciplinary collaboration between AI researchers, computer scientists, engineers, psychologists and others, where new methods and methodologies need to be created in order to develop, study and evaluate

interactions with a social robot. While it promises to result in social robots that can behave adequately in a human-inhabited (social) environment, it also raises many fundamental issues on the nature of social intelligence in humans and robots.

Humphrey (1988), in a famous paper (originally published in 1976), which discusses primate intelligence, argues for the necessity of developing a laboratory test of 'social skill'. His suggestion is as follows. 'The essential feature of such a test would be that it places the subject in a transactional situation where he can achieve a desired goal only by adapting his strategy to conditions which are continually changing as a consequence partly, but not wholly of his own behaviour. The 'social partner' in the test need not be animate (though my guess is that the subject would regard it in an 'animistic' way); possibly it could be a kind of 'social robot', a mechanical device which is programmed on-line from a computer to behave in a pseudo-social way'.

Now, 30 years after the original publication of Humphrey's idea, it is within our grasp to have robots, humanoid or non-humanoid, taking the role of a social partner in such a social intelligence test. However, 40 years after Alison Jolly's original article indicating that it is the social domain that defines us as human primates, it is still open as to what social intelligence for robots could or should mean from the perspective of humans. Despite the potential usefulness of social robots as scientific tools for understanding the nature of social intelligence on the one hand, and for the design of robotic assistants, companions or playmates that will have their places in society on the other hand, it is unclear whether the 'social–emotional' dimension in human–human interaction can be fulfilled by robots, i.e. whether the inherently 'mechanical nature' of HRIs can be replaced by truly meaningful social exchanges. While I doubt that robots can overcome their 'robotic heritage', viewing them as part of a social environment where meaning in interactions is provided by the richness and depth of human experiences might be a more realistic and more 'humane' vision for social robots than viewing them as 'selfish' machines.

Part of the survey on the Aurora project in this article formed the basis of Ben Robins's and Iain Werry's Ph.D. theses. I acknowledge their contribution to the work and the photo material. The particular work summarized in the context of the Cogniron project was carried out by the following researchers: Michael L. Walters, Kheng Lee Koay, Sarah Woods and Christina Kaouri. We are grateful to Takashi Gomi who donated the Labo-1 robot, and Aude Billard who designed the humanoid doll Robota and made it available to our studies. I would like to thank Gernot Kronreif for discussions on robotic toys for children. The work described in this paper was partially conducted within two EU Integrated Projects: COGNIRON (The Cognitive Robot Companion) funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020, and RobotCub (Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours) funded by the European Commission through Unit E5 (Cognition) of FP6-IST under Contract FP6-004370.

ENDNOTE

[1]According to the PTUK organization (Play Therapy in UK, http://www.playtherapy.org.uk), studies indicate that 20% of children have some form of psychological problem and that 70% of these can be helped through therapies, including play therapy.

REFERENCES

Arkin, R. C. 1998 *Behavior-based robotics*. Cambridge, MA: MIT Press.

Billard, A. 2003 Robota: clever toy and educational tool. *Robot. Auton. Syst.* **42**, 259–269. (doi:10.1016/S0921-8890(02)00380-9)

Billard, A., Robins, B., Dautenhahn, K. & Nadel, J. In press. Building Robota, a mini-humanoid robot for the rehabilitation of children with autism. *RESNA Assistive Technol. J.*

Billard, A. & Dautenhahn, K. 1998 Grounding communication in autonomous robots: an experimental study. *Robot. Auton. Syst.* **24**, 71–79. (doi:10.1016/S0921-8890(98)00023-2)

Blow, M. P., Dautenhahn, K., Appleby, A., Nehaniv, C. L. & Lee, D. 2006 Perception of robot smiles and dimensions for human–robot interaction design. In *Proc. 15th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN06), Hatfield, UK, 6–8 September 2006*, pp. 469–474.

Bonabeau, E., Dorigo, M. & Theraulaz, G. 1999 *Swarm intelligence–from natural to artificial systems*. Oxford, UK: Oxford University Press.

Boucher, J. 1999 Editorial: interventions with children with autism methods based on play. *Child Lang. Teaching Therapy* **15**.

Braitenberg, V. 1984 *Vehicles—experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Breazeal, C. 2002 *Designing sociable robots*. Cambridge, MA: MIT Press.

Breazeal, C. 2003 Towards sociable robots. *Robot. Auton. Syst.* **42**, 167–175. (doi:10.1016/S0921-8890(02)00373-1)

Breazeal, C. 2004 Social interaction in HRI: the robot view. *IEEE Trans. Syst. Man Cyber.: Part C* **34**, 181–186. (doi:10.1109/TSMCC.2004.826268)

Brooks, R. A. 1999 *Cambrian intelligence*. Cambridge, MA: MIT Press.

Brooks, R. A., Breazeal, C., Marjanović, M., Scassellati, B. & Williamson, M. M. 1999 The Cog project: building a humanoid robot. In *Computation for metaphors, analogy, and agents* (ed. C. L. Nehaniv) Springer Lecture Notes in Artificial Intelligence, no. 1562, pp. 52–87. New York, NY: Springer.

Byrne, R. W. 1995 *The thinking ape*. Oxford: Oxford University Press.

Byrne, R. W. 1997 Machiavellian intelligence. *Evol. Anthropol.* **5**, 172–180. (doi:10.1002/(SICI)1520-6505(1996)5:5<172::AID-EVAN6>3.0.CO;2-H)

Byrne, R. W. & Whiten, A. (eds) 1988 *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford, UK: Oxford University Press.

Dahlback, A., Jonsson, L. & Ahrenberg. 1993 Wizard of Oz studies—why and how. In *Proc. First Int. Conf. on Intelligent User Interfaces, Orlando, Florida*, USA, pp. 193–200.

Dautenhahn, K. 1994 Trying to imitate—a step towards releasing robots from social isolation. In *Proc. 'From Perception to Action' Conference, Lausanne, Switzerland, 7–9 September 1994* (eds P. Gaussier & J.-D. Nicoud), pp. 290–301. Los Alamitos, CA: IEEE Computer Society Press.

Dautenhahn, K. 1995 Getting to know each other—artificial social intelligence for autonomous robots. *Robot. Auton. Syst.* **16**, 333–356. (doi:10.1016/0921-8890(95)00054-2)

Dautenhahn, K. 1998 The art of designing socially intelligent agents—science, fiction, and the human in the loop. *Appl. Artif. Intell.* **12**, 573–617. (doi:10.1080/08839518117550)

Dautenhahn, K. 1999*a* Embodiment and interaction in socially intelligent life-like agents. In *Computation for metaphors,*

analogy and agents (ed. C. L. Nehaniv). Springer Lecture Notes in Artificial Intelligence, no. 1562, pp. 102–142. New York, NY: Springer.

Dautenhahn, K. 1999*b* Robots as social actors: Aurora and the case of autism. In *Proc. CT99, The Third International Cognitive Technology Conference, August, San-Francisco*, pp. 359–374. East Lansing, MI: Michigan State University.

Dautenhahn, K. 2002 Design spaces and niche spaces of believable social robots. In *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN 2002), 25–27 September, Berlin, Germany*, pp. 192–197. Piscataway, NJ: IEEE Press.

Dautenhahn, K. 2003 Roles and functions of robots in human society: implications from research in autism therapy. *Robotica* **21**, 443–452. (doi:10.1017/S0263574703004922)

Dautenhahn, K. 2004*a* Socially intelligent agents in human primate culture. In *Agent culture: human–agent interaction in a multicultural world* (eds R. Trappl & S. Payr), pp. 45–71. New Jersey, NJ: Lawrence Erlbaum Associates.

Dautenhahn, K. 2004*b* Robots we like to live with?!—A developmental perspective on a personalized, life-long robot companion. In *Proc. 3rd IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN 2004), 20–22 September 2004, Kurashiki, Japan*, pp. 17–22. Piscataway, NJ: IEEE Press.

Dautenhahn, K. & Billard, A. 2002 Games children with autism can play with Robota, a humanoid robotic doll. In *Universal access and assistive technology* (eds S. Keates, P. M. Langdon, P. J. Clarkson & P. Robinson), pp. 179–190. London, UK: Springer.

Dautenhahn, K. & Werry, I. 2004 Towards interactive robots in autism therapy: background, motivation and challenges. *Pragmatics Cogn.* **12**, 1–35.

Dautenhahn, K., Werry, I., Rae, J., Dickerson, P., Stribling, P. & Ogden, B. 2002 Robotic playmates: analysing interactive competencies of children with autism playing with a mobile robot. In *Socially intelligent agents—creating relationships with computers and robots* (eds K. Dautenhahn, A. Bond, L. Cañamero & B. Edmonds), pp. 117–124. Dordrecht, The Netherlands: Kluwer Academic.

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M., Koay, K. L. & Werry, I. 2005 What is a robot companion—friend, assistant or butler? In *Proc. IEEE IRS/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2005), 2–6 August 2005, Edmonton, Alberta, Canada*, pp. 1488–1493. Piscataway, NJ: IEEE Press.

Davis, M., Robins, B., Dautenhahn, K., Nehaniv, C. L. & Powell, S. 2005 A comparison of interactive and robotic systems in therapy and education for children with autism. In *Proc. 'Assistive Technology from Virtuality to Reality', 8th European Conference for the Advancement of Assistive Technology in Europe (AAATE'05), Lille, France, 6–9 September 2005*, pp. 353–357. Amsterdam, The Netherlands: IOS Press.

Dunbar, R. I. M. 1993 Coevolution of neocortical size, group size and language in humans. *Behav. Brain Sci.* **16**, 681–735.

Dunbar, R. I. M. 1996 *Grooming, gossip and the evolution of language*. Cambridge, MA: Faber and Harvard University Press.

Dunbar, R. I. M. 1998 The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190. (doi:10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8)

Dunbar, R. I. M. 2003 The social brain: mind, language and society in evolutionary perspective. *Annu. Rev. Anthropol.* **32**, 163–181. (doi:10.1146/annurev.anthro.32.061002.093158)

Evans, B. J., Stanley, R. O., Mestrovic, R. & Rose, L. 1991 Effects of communication skills training on students' diagnostic efficiency. *Med. Educ.* **25**, 517–526.

Ferrara, C. & Hill, S. D. 1980 The responsiveness of autistic children to the predictability of social and non-social toys. *J. Autism Dev. Disorders* **10**, 51–57. (doi:10.1007/BF02408432)

Fong, T., Nourbakhsh, I. & Dautenhahn, K. 2003 A survey of socially interactive robots. *Robot. Auton. Syst.* **42**, 143–166. (doi:10.1016/S0921-8890(02)00372-X)

François, D., Polani, D. & Dautenhahn, K. In press. Online behaviour classification and adaption to human–robot interaction styles. In *Proc. Second ACM/IEEE Int. Conf. on Human Robot Interaction (HRI07), Washington DC, USA*.

Gigerenzer, G. 1997 The modularity of social intelligence. In *Machiavellian intelligence II: extensions and evaluations* (eds A. Whiten & R. W. Byrne), pp. 264–288. Cambridge, UK: Cambridge University Press.

Goetz, J. & Kiesler, S. 2002 Cooperation with a robotic assistant. In *Proc. SIGCHI'02 Conf. on Human Factors in Computing Systems, Minneapolis, MN, 20–25 April 2002*. New York, NY: ACM Press.

Gould, J. D., Conti, J. & Hovanyecz, T. 1983 Composing letters with a simulated listening typewriter. *Commun. ACM.* **26**, 295–308. (doi:10.1145/2163.358100)

Heider, F. & Simmel, H. 1944 An experimental study of apparent behavior. *Am. J. Psychol.* **57**, 243–259. (doi:10.2307/1416950)

Hendriks-Jansen, H. 1996 *Catching ourselves in the act: situated activity, interactive emergence, evolution and human thought*. Cambridge, MA: MIT Press.

Humphrey, N. K. 1988 The social function of intellect. In *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes and humans* (eds R. W. Byrne & A. Whiten), pp. 13–26. Oxford, UK: Clarendon Press.

Iizuka, H. & Ikegami, T. 2004 Adaptability and diversity in simulated turn-taking behavior source. *Artif. Life* **10**, 361–378. (doi:10.1162/1064546041766442)

Jolly, A. 1966 Lemur social behavior and primate intelligence. *Science* **153**, 501–506. (doi:10.1126/science.153.3735.501)

Jordan, R. & Libby, S. 1997 Developing and using play in the curriculum. In *Autism and learning: a guide to good practice* (eds S. Powell & R. Jordan), London, UK: David Fulton.

Kahn Jr, P. H., Friedman, B., Perez-Granados, D. R. & Freier, N. G. 2006 Robotic pets in the lives of preschool children. *Interact. Studies.* 7, 405–436.

Kanda, T., Hirano, T. & Eaton, D. 2004 Interactive robots as social partners and peer tutors for children: a field trial. *Human Comp. Interact.* **19**, 61–84. (doi:10.1207/s15327051hci1901&2_4)

Koay, K. L., Walters, M. L. & Dautenhahn, K. 2005 Methodological issues using a comfort level device in human–robot interactions. In *Proc. IEEE Workshop on Robots and Human Interactive Communication (RO-MAN 2005), Nashville, Tennessee, 13–15 August 2005*, pp. 359–364. Piscataway, NJ: IEEE Press.

Koay, K. L., Dautenhahn, K., Woods, S. N. & Walters, M. L. 2006 Empirical results from using a comfort level device in human–robot interaction studies. In *Proc. 1st Int. Conf. on Human Robot Interaction (HRI'06), Salt Lake City, Utah, 2–3 March 2006*, pp. 194–201. New York, NY: ACM Press.

Kozima, H. 2002 Infanoid: a babybot that explores the social environment. In *Socially intelligent agents: creating relationships with computers and robots* (eds K. Dautenhahn, A. H. Bond, L. Cañamero & B. Edmonds), pp. 157–164. Amsterdam, The Netherlands: Kluwer Academic.

Kozima, H., Nakagawa, C. & Yasuda, Y. 2005 Interactive robots for communication-care: A case-study in autism therapy. In *Proc. IEEE Workshop on Robots and Human*

Interactive Communication (RO-MAN 2005), Nashville, Tennessee, 13–15 August 2005, pp. 341–346. Piscataway, NJ: IEEE Press.

Kramer, D., Ber, R. & Moores, M. 1989 Increasing empathy among medical students. *Med. Educ.* **23**, 168–173.

Lorenz, K. 1971 Part and parcel in animal and human societies. *Studies in animal and human behavior*, pp. 115–195. Cambridge, MA: Harvard University Press (originally published 1950).

Lungarella, M., Metta, G., Pfeifer, R. & Sandini, G. 2004 Developmental robotics: a survey. *Connect. Sci.* **15**, 151–190. (doi:10.1080/09540090310001655110)

MacDorman, K. & Ishiguro, H. 2006 The uncanny advantage of using androids in cognitive and social science research. *Interact. Studies.* **7**, 297–337.

Maulsby, D., Greenberg, S. & Mander, R. 1993 Prototyping an intelligent agent through Wizard of Oz. In *Proc. ACM SIGCHI Conf. on Human Factors in Computing Systems*, pp. 277–284. Amsterdam, The Netherlands: ACM Press.

Michaud, F. & Caron, S. 2002 Roball, the rolling robot. *Auton. Robots* **12**, 211–222. (doi:10.1023/A:1014005728519)

Michaud, F. & Théberge-Turmel, C. 2002 Mobile robotic toys and autism. In *Socially intelligent agents—creating relationships with computers and robots* (eds K. Dautenhahn, A. Bond, L. Cañamero & B. Edmonds), pp. 125–132. Dordrecht, The Netherlands: Kluwer Academic.

Michaud, F., Laplante, J.-F., Larouche, H., Duquette, A., Caron, S. & Masson, P. 2005 Autonomous spherical mobile robot to study child development. *IEEE Trans. Syst. Man Cybern.* **35**, 471–480. (doi:10.1109/TSMCA.2005.850596)

Moore, D. 1998 Computers and people with autism. *Communications* Summer, pp. 20–21.

Mori, M. 1970 Bukimi no tani [the uncanny valley]. *Energy* **7**, 33–35.

Murray, D. 1997 Autism and information technology: therapy with computers. In *Autism and learning: a guide to good practice* (eds S. Powell & R. Jordan), pp. 100–117. London, UK: David Fulton.

Nadel, J., Revel, A., Andry, P. & Gaussier, P. 2004 Toward communication: first imitations in infants, low-functioning children with autism and robots. *Interact. Studies* **5**, 45–74.

Nakauchi, Y. & Simmons, R. 2000 A social robot that stands in line. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2000), Takamatsu, Japan, 31 October–5 November 2000*, pp. 357–364. Piscataway, NJ: IEEE Press.

Ogden, B. & Dautenhahn, K. 2000 Robotic etiquette: Structured interaction in humans and robots. In *Proc. 8th Symp. on Intelligent Robotic Systems (SIRS 2000), The University of Reading, England, 18–20 July 2000*.

Ogden, B., Dautenhahn, K. Stribling, P. 2002 Interactional structure applied to the identification and generation of visual interactive behaviour: Robots that (usually) follow the rules. In: *Gesture and sign languages in human–computer interaction* (eds I. Wachsmuth, & T. Sowa), Springer Lecture Notes LNAI 2298, pp. 254–267. Berlin, Germany: Springer.

Parten, M. B. 1932 Social participation among pre-school children. *J. Abnormal. Social Psychol.* **27**, 243–268. (doi:10.1037/h0074524)

Penny, S. 2000 Agents as artworks and agent design as artistic practice. In *Human cognition and social agent technology* (ed. K. Dautenhahn), pp. 395–414. Amsterdam, The Netherlands: John Benjamins

Pfeifer, R. & Scheier, C. 1999 *Understanding intelligence*. Cambridge, MA: Bradford Books.

Plaisant, C., Druin, A., Lathan, C., Dakhane, K., Edwards, K., Vice, J. M. & Montemayor, J. 2000 A storytelling robot for pediatric rehabilitation. In *Proc. ASSETS'00, Washington, Nov. 2000*, pp. 50–55. New York, NY: ACM Press.

Powell, S. 1996 The use of computers in teaching people with autism. *Autism on the agenda: papers from the National Autistic Society conference*. London, UK: National Autistic Society.

Reader, S. M. & Laland, K. N. 2002 Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl Acad. Sci. USA* **99**, 4436–4441. (doi:10.1073/pnas.062041299)

Reeves, B. & Nass, C. 1996 *The media equation—how people treat computers, television, and new media like real people and places*. Cambridge, UK: Cambridge University Press.

Restall, G. & Magill-Evans, J. 1994 Play and preschool children with autism. *Am. J. Occup. Therapy.* **48**, 113–120.

Richter, S. 1989 *Wunderbares Menschenwerk: Aus der Geschichte der mechanischen Automaten*. Leipzig, Germany: Edition Leipzig.

Robins, B. & Dautenhahn, K. 2006 The role of the experimenter in HRI research—a case study evaluation of children with autism interacting with a robotic toy. In *Proc. 15th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN06), Hatfield, UK, 6–8 September 2006*, pp. 646–651. New York, NY: IEEE Press.

Robins, B., Dickerson, P., Stribling, P. & Dautenhahn, K. 2004*a* Robot-mediated joint attention in children with autism: A case study in a robot–human interaction. *Interact. Studies* **5**, 161–198.

Robins, B., Dautenhahn, K., te Boelhorst, R., Billard, A. 2004*b* Robots as assistive technology—does appearance matter? In *Proc. 13th IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN 2004), 20–22 September, Kurashiki, Japan*, pp. 277–282. IEEE Press.

Robins, B., Dautenhahn, K., te Boekhorst, R. & Billard, A. 2004*c* Effects of repeated exposure of a humanoid robot on children with autism. In *Designing a more inclusive world* (eds S. Keates, J. Clarkson, P. Langdon & P. Robinson), pp. 225–236. London, UK: Springer.

Robins, B., Dautenhahn, K. & Dubowski, J. 2004*d* Investigating autistic children's attitudes towards strangers with the theatrical robot—a new experimental paradigm in human–robot interaction studies? In *Proc. 13th IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN 2004), Kurashiki, Japan, 20–22 September 2004*. Piscataway, NJ: IEEE Press.

Robins, B., Dautenhahn, K. & Dubowski, J. 2005*a* Robots as isolators or mediators for children with autism? A cautionary tale. In *Proc. AISB'05 Symp. Robot Companions Hard Problems and Open Challenges in Human–Robot Interaction, 14–15 April 2005, University of Hertfordshire, UK*, pp. 82–88. Hatfield, UK: SSAISB. See http://www.aisb.org.uk/publications/proceedings.shtml.

Robins, B., Dautenhahn, K., te Boekhorst, R. & Billard, A. 2005*b* *Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Universal Access in the Information Society (UAIS)*. New York, NY: Springer.

Robins, B., Dickerson, P. & Dautenhahn, K. 2005*c* Robots as embodied beings—interactionally sensitive body movements in interactions among autistic children and a robot. In *Proc. 14th IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN 2005), Nashville, Tennessee, USA, 13–15 August 2005*, pp. 54–59. Piscataway, NJ: IEEE Press.

Salter, T., te Boekhorst, R. & Dautenhahn, K. 2004 Detecting and analysing children's play styles with autonomous mobile robots: a case study comparing observational data

with sensor readings. In *Proc. 8th Conf. on Intelligent Autonomous Systems (IAS-8), 10–13 March 2004*, p. 61. Amsterdam, The Netherlands: IOS Press.

Salter, T., Dautenhahn, K. & te Boekhorst, R. 2006 Learning about natural human–robot interaction. *Robot. Auton. Syst.* **54**, 127–134. (doi:10.1016/j.robot.2005.09.022)

Scassellati, B. 2005 Quantitative metrics of social response for autism diagnosis. In *Proc. IEEE Workshop on Robots and Human Interactive Communication* (*RO-MAN 2005, Nashville, Tennessee*), pp. 585–590. Piscataway, NJ: IEEE Press.

Scholtz, J. 2003 Theory and evaluation of human robot interactions. In *Proc. Hawaii Int. Conf. on System Science 36, Jan 2003*.

Severinson-Eklundh, K., Green, A. & Hüttenrauch, H. 2003 Social and collaborative aspects of interaction with a service robot. *Robot. Auton. Syst.* **42**, 223–234. (doi:10.1016/S0921-8890(02)00377-9)

Sherratt, D. & Peter, M. 2002 *Developing play and drama in children with autistic spectrum disorders*. London, UK: David Fulton.

Steels, L. 1994 The artificial life roots of artificial intelligence. *Artif. Life* **1**, 89–125.

Steinfeld, A. M., Fong, T. W., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. & Goodrich, M. 2006 Common metrics for human–robot interaction. In *Proc. 2006 Human–Robot Interaction Conference (HRI 2006)*, *March 2006*, pp. 33–40. New York, NY: ACM Press.

Tardiff, C., Plumet, M. H., Beaudichon, J., Waller, D., Bouvard, M. & Leboyer, M. 1995 Micro-analysis of social interactions between autistic children and normal adults in semi-structured play situations. *Int. J. Behav. Dev.* **18**, 727–747.

Thrun, S. 1998 When robots meet people. *IEEE Intell. Syst.* **27**, 29.

Thrun, S. 2004 Towards a framework of human–robot interaction. *Human Comput. Interact.* **19**, 9–24. (doi:10.1207/s15327051hci1901&2_2)

Walter, W. G. 1950 An imitation of life. *Sci. Am.* **182**, 42–45.

Walter, W. G. 1951 A machine that learns. *Sci. Am.* **185**, 60–63.

Walters, M. L., Woods, S. N., Koay, K. L. & Dautenhahn, K. 2005*a* Practical and methodological challenges in designing and conducting interaction studies with human subjects. In *Robot companions: hard problems and open challenges in human–robot interaction, Symposium at AISB'05, University of Hertfordshire, UK*, pp. 110–119. Hatfield, UK: SSAISB.

Walters, M. L., Dautenhahn, K., Koay, K. L., Kaouri, C., te Boekhorst, R., Nehaniv, C. L.,. Werry, I. & Lee, D. 2005 Close encounters: Spatial distances between people and a robot of mechanistic appearance. In *Proc. IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids2005), December 5–7, Tsukuba, Japan*, pp. 450–455. New York, NY: IEEE Press.

Walters, M. L., Dautenhahn, K., Woods, S. N., Koay, K. L., te Boekhorst, R. & Lee, D. 2006 Exploratory studies on social spaces between humans and a mechanical-looking robot. *Connect. Sci.* **18**, 429–442.

Weir, S., & Emanuel, R. 1976 *Using Logo to catalyse communication in an autistic child*. DAI research report No. 15, University of Edinburgh.

Werry, I. 2003 *Development and evaluation of a mobile robotic platform as a therapy device for children with autism*. Unpublished Ph.D. thesis, University of Reading, UK.

Werry, I. & Dautenhahn, K. In press. The emergent nature of human–robot interaction: an experimental study with children with autism. In: *Modeling biology: structures, behavior* (eds L. da Fontoura Costa, M. Laubichler & G. Müller), Vienna Series in Theoretical Biology. Cambridge, MA: MIT Press.

Werry, I. & Dautenhahn, K. 1999 Applying robot technology to the rehabilitation of autistic children. In *Proc. 7th Int. Symp. on Intelligent Robotic Systems (SIRS99)*, pp. 265–272.

Werry, I., Dautenhahn, K., Ogden, B. & Harwin, W. 2001 Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy. In *Proc. Fourth Int. Conf. on Cognitive Technology: Instruments of Mind* (eds M. Beynon, C. L. Nehaniv, & K. Dautenhahn), pp. 57–74. LNAI 2117. Cambridge, UK: Springer.

Whiten, A. & Byrne, R. W. (eds) 1997 *Machiavellian intelligence II: extensions and evaluations*, Cambridge, UK: Cambridge University Press.

Wolfberg, P. J. 1999 *Play and imagination in children with autism*. New York, NY: Teachers College Press.

Woods, S., Dautenhahn, K., Kaouri, C., te Boekhorst, R. & Koay, K. L. 2005 Is this robot like me? Links between human and robot personality traits. In *Proc. IEEE-RAS International Conference on Humanoid Robots (Humanoids2005), December 5–7, Tsukuba International Congress Center, Tsukuba, Japan*, pp. 375–380. New York, NY: IEEE Press.

Woods, S. N., Walters, M. L., Koay, K. L. & Dautenhahn, K. 2006*a* Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *Proc. 9th IEEE Int. Workshop on Advanced Motion Control (AMC'06), March 27–29, Istanbul, Turkey*, pp. 750–755. New York, NY: IEEE Press.

Woods, S. N., Walters, M. L., Koay, K. L. & Dautenhahn, K. 2006*b* Methodological issues in HRI: A comparison of live and video-based methods in robot to human approach direction trials. In *Proc. 15th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN06), Hatfield, UK, September 6–8, 2006*, pp. 51–58. New York, NY: IEEE Press.

Yanco, H. A. & Drury, J. 2002 A taxonomy for human–robot interaction. In *Proc. AAAI Fall Symp. on Human–Robot Interaction, Falmouth, Massachusetts, November 2002*, AAAI Technical Report FS-02-03, pp. 111–119. AAAI Press.

Yanco, H. A. & Drury, J. 2004 Classifying human–robot interaction: an updated taxonomy. In *Proc. IEEE Conf. on Systems, Man and Cybernetics, The Hague, The Netherlands*, vol. 3, pp. 2841–2846. New York, NY: IEEE Press.