Proceedings of the 2002 IEEE
Int. Workshop on Robot and Human Interactive Communication
Berlin, Germany, Sept. 25-27, 2002

Design Spaces and Niche Spaces of Believable Social Robots

Kerstin Dautenhahn
Adaptive Systems Research Group
University of Hertfordshire
College Lane, Hatfield AL10 9AB, United Kingdom
E-mail K.Dautenhahn@herts.ac.uk

Abstract

This paper discusses the design space of believable social robots. We synthesise ideas and concepts from areas as diverse as comics design and rehabilitation robotics. First, we revisit the work of the Japanese researcher Masahiro Mori in the context of recent developments in social robots. Next, we discuss work in the arts into comics design, an area which has dealt for decades with the problem of creating believable characters. Finally, in order to illustrate some of the important issues involved we focus on a particular application area: the use of interactive robots in autism therapy, work that is carried out in the Aurora project. We discuss design issues of social robots in the context of 'design spaces' and 'niche spaces', concepts that have been defined originally for intelligent agent architectures [24] but which, we propose, can be highly valuable for social robotics design. This paper is meant to open up a discussion towards a systematic exploration of design spaces and niche spaces of social robots.

I. Introduction

Aaron Sloman discusses different possible mappings between the design space and niche space of intelligent agent architectures, in particular those architectures necessary to yield 'human-like' agents.

> "...we need to explore a space of possible designs for behaving systems (design space) and a space of possible sets of requirements (niche space) and the mappings between the two. It is not to be expected that there is any one "right" architecture. As biological diversity demonstrates, many different architectures can be successful, and in different ways. There are different "niches" (sets of requirements and constraints) for which architectures can be evaluated and compared, and such evaluations will not generally yield a Yes/No decision, but rather an analysis of trade-offs, often involving several dimensions of comparison." [24]

This article puts forward the suggestion that the same concepts of *design space* and *nice space*, and relationships between the two, are not only relevant for robot control architectures, but also hold for appearance and behaviour of believable social robots, i.e. robots whose important (if not primary) function is to appear believable to human observers and interaction partners.

Social robots that exist in human-inhabited environments have become increasingly popular, e.g. in application areas such as service robotics [26], entertainment (cf. Sony's robot dog Aibo), or education [8]. How should robots be designed so that they are acceptable to humans in their particular application domains? Is it advantageous to imitate life, i.e. to give the robots shapes and behaviours that resemble humans or other animals? Or are minimal designs preferable, is 'simpler better', and if it is, how application specific is it? This article discusses some of these issues in more detail, drawing on ideas from the Japanese roboticist Masahiro Mori Mori and artistic work on comics. In order to exemplify some of the issues discussed, the last part of the paper focusses on the particular application domain of autism therapy where robotic design is a central issue.

This paper is structured as follows. Section II introduces the *Life-Like Agents Hypothesis* which seemingly is often taken for granted in the autonomous agents community. One of the probably earliest explicit discussions of design spaces of 'life-like' (believable) robots is summarised in section III. Section IV discusses lessons that roboticists could learn from artists who are designing comics. Section V introduces the application background of using robots in autism therapy. Particular robot design issues important in this field are discussed in section VI, before section VII concludes this paper.

II. Life-Like Agents Hypothesis

Robots are often designed to 'imitate' life. The large number of current *humanoid robotics* projects exemplifies this tendency. Based on what we call the 'Life- Like Agents Hypothesis' this approach can be characterised as follows [6]:

> "Artificial social agents (robotic or software) which are supposed to interact with humans are most successfully designed by imitating life, i.e. making the agents mimic as closely as possible animals, in particular humans. This comprises both 'shallow' approaches focussing on the presentation and believability of the agents, as well as 'deep' architectures which attempt to model faithfully animal cognition and intelligence".

Generally it is argued that such life-like agents are desirable because:
(1) the agents are supposed to act on behalf of or in collaboration with humans; they adopt roles and fulfil tasks normally done by humans, thus, it is argued that they require human forms of (social) intelligence,
(2) users prefer to interact ideally with other humans and less ideally with human-like agents. Thus, it is hoped that life-like agents can naturally be integrated in human work and entertainment environment, e.g. as assistants or pets,
(3) Life-like agents can serve as models for the scientific investigation of animal behaviour and animal minds.

Argument (3) is certainly valid and need not be discussed here. However, as we will point out in this article, arguments (1) and (2) are not as straightforward as they seem. Life-like agents that closely mimic human appearance or behaviour can unnecessarily restrict and narrow the apparent and actual functionality of an agent by evoking expectations that the agent cannot fulfil. For example, a humanoid robot elicits strong expectations about the robot's social and cognitive competencies. If such expectations are not being met then the user is likely to experience confusion, frustration and disappointment. This effect is highly context-dependent: in play-like entertainment scenarios such breaches of expectations are likely to be more acceptable than e.g. in application areas where robots serve as assistants. To give an example, let us consider a social robot with humanoid appearance and behaviour which operates in a department store. Humanoid features might have the advantage of evoking an initial feeling of 'familiarity' in a human customer and eliciting anthropomorphic tendencies, but 1) human customers are then likely to also expect the robot to exhibit other human characteristics and functionalities, e.g. extensive social skills, personality and other traits of humans in general, and sales assistants in particular (including that it understands jokes and possesses common sense knowledge), [21], and 2) new or different functionalities that the real agent does not possess need to be integrated in a plausible way in the agent's behaviour, without breaking the suspense of disbelief [20].

Thus, when faced with a human in a department store we might ask ourselves 'who is this?' (a customer? a sales assistant? a security guard? the manager?), but we know clearly *what* the person is, namely a member of the human species, which already allows us to make quite strong assumptions about his abilities, skills and capacities. On the other hand, a robot in the role of a sales assistant leaves us widely in the open about what its skills and capacities are. Can it talk? Can it understand English? Does it know what the colour blue is?, etc.

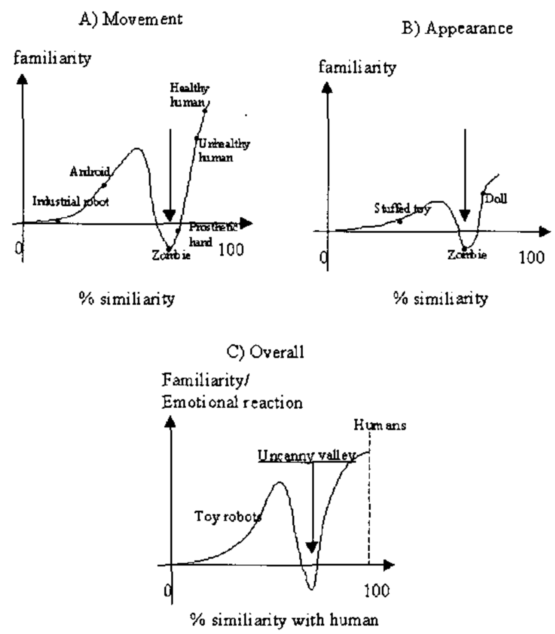

Fig. 1. How life-like should a machine be? Masahiro Mori's uncanny valley, redrawn and modified from [23] and [3].

III. The Uncanny Valley

An interesting analysis of how life-like robots should be manufactured was conducted by the Japanese researcher Masahiro Mori cf. [23], [3]. Mori tried to predict the psychological effects that different robotic designs and other human-like artifacts have on humans. His important contribution to the present discussion is his proposal of the *uncanny valley* for life-like robots, see figure 1. He predicts that the more life-like we make robots (i.e. the more similar they become to us), the more familiar they become (today we would use the term 'believable' instead of 'familiar'), until ultimately, in the case of 100 percent similarity with healthy human beings, familiarity levels reach a maximum. However, the transition has a local minimum, characterised by a sharp drop in familiarity when robots appear very life-like and might sometimes be mistaken for 'real'. In this case the robots can cause an uncanny and unpleasant feeling where still existing (but possibly very small) differences suddenly make us realise that the robots are not real, thus violating our expectations. This is also called the 'Zombie-effect': moving corpses that are (strangely looking but) similar to us, until we realise that they are not alive. Similarly, an arm prosthesis might at first glance appear very real, until we touch the arm and it becomes clear that it is made of plastic and metal. Not surprisingly a number of horror and science fiction stories are based on this effect where familiar people that are 'like-us' are suddenly identified as aliens, zombies or the like. The 'uncanny valley' effect, which is well known in the animation and movie literature, has only recently been acknowledged in robotics [25].

Interestingly, in addition to the overall graph shown in figure 1 (c), Mori distinguished two separate graphs reflecting two different components of similarity, namely movement (a) and appearance (b). For both criteria he assumed curves of similar shape but different amplitudes; the movement curve is considered to be more dominant than the appearance curve. Mori's suggestions are supported by psychological studies on anthropomorphism, cf. [18], [13], which suggest that behaviour and movements are more effective than appearance in eliciting anthropomorphic projections. For robots that operate in human-inhabited this has important consequences.

The next section discusses lessons on believability that can be learnt from comics design.

IV. The Design Spaces of Comics

As comics designers have known for decades, the particular representation used to portray characters in a comic can influence dramatically the way people identify and sympathise with its characters. We are more likely to identify with Lucy or Charlie Brown than with Marilyn Monroe. Many people, children and adults, can identify with the former, they represent iconic 'universal' characters. A Marilyn Monroe representation stands for Marilyn Monroe, a unique individual with a particular biography and personality. We might be interested in her life stories and events that happened to her. And we might feel that she could stand for a whole generation of blonde female actors, but the scope for identifying with her is very different from the kind of 'universal' identification that is possible with simpler, less concrete, more generic characters such as Lucy or Charlie Brown which we can project a variety of subjective experiences onto.

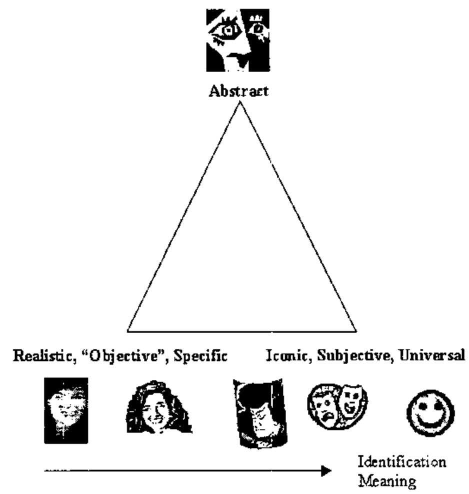

Fig. 2. The design space of comics, inspired by [16].

Figure 2 shows Scott McCloud's design space of comics. In addition to the 'realistic' versus 'iconic' dimension (horizontal) he identifies a third dimension (vertical) along which representations become less and less concrete, where the representation *itself* becomes the focus of attention. What lessons can designers of believable robots learn from this? Firstly, simple designs might be better than realistic anthropo- or zoomorphic designs that try imitating life. Secondly, a 'new' design that is not imitating any naturally existing agent might better suit its role in a robotics context so that the robot's *behaviour* and its *functionalities* predominantly determine the user's attitude towards the machine, and only to a much lesser extent any preconceptions, expectations or anthropomorphic projections that can bias the user's attitude even before any interactions have occurred.

The next section gives an example of robots used in human-inhabited environments, the specific application area is autism therapy.

V. The Aurora Project: Interactive Robots in Autism Therapy

Increasingly researchers study the application of interactive software and robotic systems in autism therapy. Common characteristic of people diagnosed along the autistic spectrum are impairments in communication and social interaction [15], comprising some deficits e.g. in imitation, turn-taking and social play. Seminal work done in the early 1970ies in Edinburgh with one autistic child controlling a teleoperated (non-interactive) robot gave first encouraging ideas [29] about a potential therapeutic role of robots. More recently interactive and robotic systems are being studied [6], [17], [22].

The work discussed in this paper is part of the long-term project Aurora [1] which studies systematically since 1998 how robots with simple interaction skills can be used as a teaching device ('toys') in autism therapy. Aurora stands for 'Autonomous robotic platform as a remedial tool for children with autism'. Particular therapeutically relevant issues that we investigate include imitation and turn-taking games.

Robot-human interactions in the Aurora project are widely unconstrained and unstructured, children are encouraged to explore and 'discover' their interaction skills rather than being taught explicitly, see figure 3. In the general context of robot-human interactions these conditions are much different from other projects on robot-human interaction (e.g. work at the MIT AI-Lab with the social robot Kismet [4]) where the human is expected to interact with the robot while adopting a particular position and orientation towards the robot (e.g. sitting face-to-face in close distance to an interactive robot that is not moving in space). The mobile robot that we are using in the Aurora project allows full-body interactions. The children who are interacting with the robot are between 8-12 years of age. For more general background on the project's robotic and therapeutic issues, and both qualitative and quantitative results see e.g. ([9], [27], [28], [10]).

Fig. 3. An autistic child interacting with a mobile robot that was explicitly not given any life-like features such as fur, tail, legs, facial expressions. With respect to the design triangle shown in fig. 2 the robot's iconic machine-like appearance and appearance is located near the bottom right corner, in comparison to other robots such as Aibo or clearly humanoid robots which are closer to the bottom left corner of the design triangle.

The normal tendency to anthropomorphise, to use fantasy and imagination for interpreting inanimate objects as 'agents' and being 'alive', a tendency that so many entertainment and service robotics projects build on, cannot be applied easily to children with autism. They tend to strongly focus on the *literal* meaning of things, rather than their holistic interpretation. They are more likely to focus on details of toys and other objects (colour, fur, mechanical details etc.) rather than perceiving the overall shape or 'cuteness'. Also, autistic people are often overwhelmed by the multitude of stimuli and different modalities of communication and interaction as they occur in human social interaction. Therefore, confirmed by studies with autistic children and non-autistic control groups [11], with respect to robot design, we suggest that 'simpler is better'. The Aurora project utilises robots with simple designs and small behavioural repertoires. The aim is to slowly guide an autistic child through increasingly complex and therapeutically relevant interaction patterns. In addition to using mobile robots, we have begun studying interactions of autistic children of 5-10 years of age with a small, humanoid robotic doll called Robota, developed by Aude Billard [7].

VI. Niche Spaces: Some Design Issues in the Autism Therapy with Robots

Based on our experience in the Aurora project with different robots including non-humanoid, mobile robots and humanoid robotic dolls, the following advantages and disadvantages are suggested. In some contexts remote-controlled robotic machines might be suitable in autism therapy, i.e. when it is very important that the child is completely in control and when a very constrained interface is appropriate. Small, autonomous, mobile robots on the other hand support unconstrained full-body interactions in space, but they pose high demands on the robot's robustness and (usually a major bottleneck) its perception of the interactions. Zoomorphic robots with human or animal appearance might appeal to children in general ("cuteness factor") but to autistic children the number and complexity of features (appearance, behaviour) can be overwhelming or scary since people with autism have great problems with a 'holistic' perception (integrating different stimuli and modalities into a coherent 'whole', such as a face). Autonomous, human-size, humanoid robots can possibly allow a wide range of 'humanoid' behaviour and can be used to practise specific aspects of human-human interaction (e.g. joint attention, facial expressions). However, interactions with humanoid robots are usually highly constrained and safety issues are a major concern. Small humanoid dolls are safer and can support multi-modal interactions, such as the 'special purpose' robot Robota that can play imitative interaction games, [2]. Robota is small, robust, easily transportable and possesses only a few humanoid features that can be varied. The robot is relatively inexpensive and has a large repertoire of behaviour expressiveness. The major drawback is that Robota (in the set up we used so far) requires children to sit on a chair facing the robot [7].

The issue of life-like versus non-life-like is an open and challenging issue. Future research needs to explore further the design spaces of behaviour, appearance and interactional competencies of robots in autism therapy. The predictability of behaviour is an important issue. Autistic children can often hardly adapt to novel and dynamically changing situations. Although our studies show that children with autism can cope very well with the new situations presented in the trials (being exposed to a robot that they have not seen before), our robots are with respect to behaviour and appearance relatively simple and therefore much more predictable than human beings in natural social interactions.

Studies by Ferrara and Hill [11] confirm our approach taken in the Aurora project. Their study with autistic children and non-robotic toys shows that, in contrast to control groups with typically developing children, autistic children prefer simple designs in relatively predictable environments. They conclude that those form an excellent starting point for therapeutic intervention where one could slowly increase the complexity of the therapeutic toys.

Our own studies with two different robotic platforms (a machine-type mobile robot and a small humanoid doll robot) indicate that the question 'which robot to use in autism therapy?' does not have a simple answer: it is more likely that for particular therapeutic goals and for different groups of children with particular social and cognitive needs we might need specialised designs: the design space needs to map onto the niche space, the space of requirements posed by the particular application domain, taking into account specific needs of groups of users as well as individuals.

Thus, the usage of robots in autism therapy poses many challenges. Potentially different solutions might prove suitable for different groups of children with different abilities and personal interests. The particular therapeutic issues that should be addressed are also likely to influence the choice of robotic design.

VII. Conclusion

As we argued in this paper[1], believable design of robots is a matter of balance: finding the appropriate level of similarity with humans, and taking into account movement and appearance, and possibly many other factors. Various aspects of how the agent looks and behaves need to be consistent. In the related area of virtual agents, e.g. embodied conversational agents ([5]), researchers are intensively studying how to design agents that appear believable to humans, and consistency and balance of design have been identified as key issues [14]. For example, a synthetic face does not necessarily enhance the appeal of a speech interface, a positive effect is only achieved when face and speech are of similar type [12], i.e. synthetic or natural.

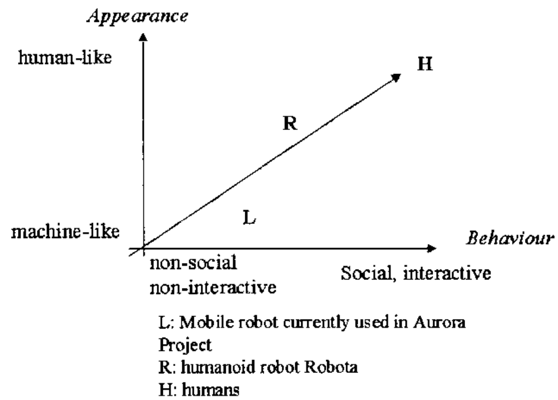

Fig. 4. Variations of behaviour and appearance for robots in autism therapy. Two different robotic systems used in the Aurora project are shown. This figure addresses the dimensions of appearance and behaviour. Ultimately a multi-dimensional design space needs to be investigated. For a variety of other robot designs engineered for playful interactions with children see [19].

Play-like learning scenarios have a great potential in autism therapy [11]. Generally, interactive, social robotic toys seem particularly suited to encourage play, cf. [8]. By systematically exploring the design space along the spectrum of life-like robots (see figure 4) one might find designs suitable for specific user groups. In the context of autism therapy one can develop designs that have therapeutic value and meet the individual, social and cognitive needs of children with autism.

[1] Earlier versions of sections II, III and V appeared in K. Dautenhahn: The Design Space of Life-Like Robots. In D. Polani, J. Kim, T. Martinetz (Eds.) *Fifth German Workshop on Artificial Life: Abstracting and Synthesizing the Principles of Living Systems*, IOS Press, pp. 135-142, 2002.

Acknowledgements

This project is supported by an EPSRC grant (GR/M62648). The mobile robotic platform used is kindly donated by Applied AI Inc. Thanks to Iain Werry, Bernard Ogden, Penny Stribling, Paul Dickerson, John Rae, Stuart Powell and Aude Billard for discussions on the Aurora project, and to Lola Cañamero for giving me pointers to M. Mori's work.

References

[1] AURORA. Url: http://www.aurora-project.com/. Last referenced on 30th of July, 2002.

[2] A. Billard. Play, dreams and imitation in Robota. In K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, editors, *Socially Intelligent Agents – Creating Relationships with Computers and Robots*, pages 165–172. Kluwer Academic Publishers, 2002.

[3] D. Bryant. http://www.arclight.net/~pdb/glimpses/valley.html. Last referenced on 30th of July, 2002.

[4] C. Breazeal and B. Scassellati. Infant-like social interactions between a robot and a human caretaker. *Adaptive Behavior*, 8(1):49–74, 2000.

[5] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. *Embodied conversational agents*. MIT Press, 2000.

[6] K. Dautenhahn. Robots as social actors: Aurora and the case of autism. In *Proc. CT99, The Third International Cognitive Technology Conference, August, San Francisco*, pages 359–374, 1999.

[7] K. Dautenhahn and A. Billard. Games children with autism can play with Robota, a humanoid robotic doll. In S. Keates, P. M. Langdon, P. J. Clarkson, and P. Robinson, editors, *Universal Access and Assistive Technology*, pages 179–190. Springer-Verlag London, 2002.

[8] A. Druin and J. Hendler, editors. *Robots for Kids – Exploring new technologies for learning.* Morgan Kaufmann Publishers, 2000.

[9] K. Dautenhahn and I. Werry. Issues of robot-human interaction dynamics in the rehabilitation of children with autism. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, and S. W. Wilson, editors, *Proc. The Sixth International Conference on the Simulation of Adaptive Behavior (SAB2000)*, pages 519–528, 2000.

[10] K. Dautenhahn, I. Werry, J. Rae, P. Dickerson, P. Stribling, and B. Ogden. Robotic playmates – analysing interactive competencies of children with autism playing with a mobile robot. In K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, editors, *Socially Intelligent Agents – Creating Relationships with Computers and Robots*, pages 117–124. Kluwer Academic Publishers, 2002.

[11] C. Ferrara and S. D. Hill. The responsiveness of autistic children to the predictability of social and non-social toys. *Journal of Autism and Developmental Disorders*, 10(1):51–57, 1980.

[12] L. Gong and C. Nass. Does adding a synthetic face always enhance speech interfaces. Stanford University Class Project downloaded from http://www.stanford.edu/~nass/comm369/.

[13] F. Heider and M. Simmel. An experimental study of apparent behavior. *American Journal of Psychology*, 57:243–259, 1944.

[14] K. Isbister and C. Nass. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *Int. Journal of Human-Computer Studies*, 53:251–267, 2000.

[15] R. Jordan. *Autistic Spectrum Disorders – An Introductory Handbook for Practitioners.* David Fulton Publishers, London, 1999.

[16] S. McCloud. *Understanding Comics: The Invisible Art.* HarperCollins Publishers, Inc., New York, 1993.

[17] F. Michaud, A. Clavet, G. Lachiver, and M. Lucas. Designing toy robots to help autistic children - an open design project for electrical and computer engineering education. *Proc. American Society for Engineering Education*, 2000.

[18] R. W. Mitchell and M. Hamm. The interpretation of animal psychology: Anthropomorphism or behavior reading? *Behaviour*, 134:173–204, 1997.

[19] F. Michaud and C. Théberge-Turmel. Mobile robotic toys and autism. In K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, editors, *Socially Intelligent Agents – Creating Relationships with Computers and Robots*, pages 125–132. Kluwer Academic Publishers, 2002.

[20] C. L. Nehaniv and K. Dautenhahn. Living with socially intelligent agents: a cognitive technology view. In K. Dautenhahn, editor, *Human Cognition and Social Agent Technology*, chapter 16, pages 415–426. John Benjamins Publishing Company, 2000.

[21] D. A. Norman. How might people interact with agents. In Jeffrey M. Bradshaw, editor, *Software Agents*, pages 49–55. AAAI Press/The MIT Press, 1997.

[22] C. Plaisant, A. Druin, C. Lathan, K. Dakhane, K. Edwards, J. M. Vice, and J. Montemayor. A storytelling robot for pediatric rehabilitation. Proc. ASSETS '00, Washington, Nov. 2000, ACM, New York, pp. 50-55, 2000.

[23] J. Reichardt. *Robots: Fact, Fiction and Prediction.* Thames and Hudson Ltd., London, 1978.

[24] A. Sloman. What sort of architecture is required for a human-like agent? Invited talk at Cognitive Modeling Workshop, AAAI96, Portland Oregon, Aug 1996, 1996.

[25] M. Scheeff, J. Pinto, K. Rahardja, S. Snibbe, and R. Tow. Experiences with sparky, a social robot. In K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, editors, *Socially Intelligent Agents – Creating Relationships with Computers and Robots*, pages 173–180. Kluwer Academic Publishers, 2002.

[26] R. D. Schraft and G. Schmierer. *Service Robotics.* AK Peters, Ltd., Natick, MA, 2000.

[27] I. Werry, K. Dautenhahn, and W. Harwin. Investigating a robot as a therapy partner for children with autism. Proc. 6th European Conference for the Advancement of Assistive Technology (AAATE 2001), 3-6 September 2001 in Ljubljana / Slovenia., 2001.

[28] I. Werry, K. Dautenhahn, B. Ogden, and W. Harwin. Can social interaction skills be taught by a social agent? the role of a robotic mediator in autism therapy. In M. Beynon, C. L. Nehaniv, and K. Dautenhahn, editors, *Proc. The Fourth International Conference on Cognitive Technology: Instruments of Mind, LNAI 2117*, pages 57–74, Berlin Heidelberg, 2001. Springer-Verlag.

[29] S. Weir and R. Emanuel. Using LOGO to catalyse communication in an autistic child. Technical report, DAI Research Report No. 15, University of Edinburgh, 1976.