Connection Science
Vol. 18, No. 4, December 2006, 313–317

Taylor & Francis
Taylor & Francis Group

Editorial

Introduction to the special issue on android science

KARL F. MACDORMAN*

School of Informatics, Indiana University, Indianapolis, Indiana 46202, USA

Android science is a new interdisciplinary framework for studying human cognition and interaction based on the premise that a very humanlike robot can elicit the sort of responses people typically direct toward each other (Ishiguro 2006, MacDorman and Ishiguro 2006b). If correct, this would allow for an android to be used as a stand-in for a human being in social, psychological, cognitive and neuroscientific experiments with human participants. Figure 1, adapted from Ishiguro (2006), shows the potential for disciplines from the social and cognitive sciences to engage in a process of hypothesis formation and verification through human–android interaction. For example, from the standpoint of cognitive neuroscience, we can study the effect of appearance on brain activity during the perception of intentional action, and from the standpoint of developmental and social psychology, we can study the effect of contingency and timing on interactions with human infants or autistic children.

One advantage of using an android as an experimental apparatus is that it can be more precisely controlled than a human actor. It also has physical presence, which is lacked by a video or computer simulation of a human being. Moreover, in comparing human–human and human–robot interaction, an android controls better for the effects of appearance than a mechanical-looking robot (MacDorman and Ishiguro 2006c). Thus, for many kinds of experiments, an android offers a good balance between experimental control and ecological validity because it looks more human than other devices and can support more humanlike interaction while still being precisely controllable. Lindblom and Ziemke (2006) note in this issue that 'androids can be used in more realistic experiments on social interaction with people, because they elicit more natural responses, that is, responses more like those of human–human social interaction' (section 3.1). However, experiments with present day androids show mixed results, eliciting human-directed social responses more than mechanical-looking robots but less than human beings. This is not surprising because androids are still only indistinguishable from human beings for short, one- or two-second presentations, as demonstrated by Ishiguro (2006) in this issue; but most experiments can last several minutes (e.g. MacDorman *et al.* 2005). Thus, android robots are still a 'work in progress', defined by a goal that has yet to be attained—and that may be unattainable—namely, perfect human likeness (MacDorman and Ishiguro 2006a).

Connection Science
ISSN 0954-0091 print/ISSN 1360-0494 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/09540090600906258

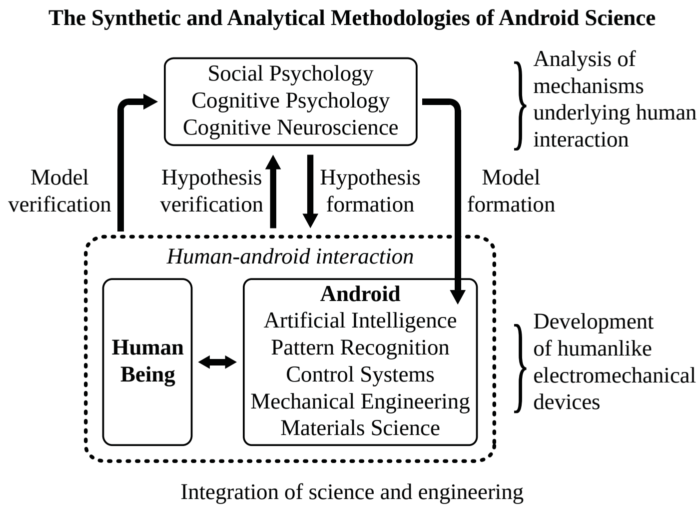

Figure 1. Android science provides both synthetic and analytical methods of understanding human interaction and the mechanisms that underlie it.

In addition to using an android as an analytical tool for hypothesis testing, an experimental setting for human–android interaction provides a testing ground for developing models concerning how cognitive or neural processing influence human interaction (MacDorman and Ishiguro 2006b). As indicated in figure 1, cognitive models can be implemented in the android with the benefit of progress in such disciplines as artificial intelligence and mechanical engineering. These models can then be tested in interaction with human participants and, based on the results, either replaced or incrementally improved. Given the complexity of human behaviour, this process may continue indefinitely. Once again, an android's human likeness affords a degree of ecological validity to human–robot studies that goes beyond the general benefits of embodying and testing cognitive models in robots (e.g. Cowley and MacDorman 1995, Asada *et al.* 2001). Mori's (1970) 'uncanny valley', that is, the heightened sensitivity of human beings to deviations from human norms of appearance and behaviour in near-human forms, provides a diagnostic tool for improving the realism of an android. In this way, the uncanny valley benefits android science.

The analysis of the human perception of androids and human–android interaction can also drive improvements in hardware technologies. It has revealed a need for motor actuators and joints that can come closer to the range of movement and stiffness and flexibility of human limbs and for artificial skin that looks and moves more like human skin for the effective expression of emotion (e.g. Hanson *et al.* 2005).

Ziemke and Lindblom (2006) identify a methodological difficulty with the holistic evaluation of android systems with respect to pinpointing which among a potentially large number of subsystems could have caused a difference in behaviour. However, this is a general problem for the study of human beings in cognitive neuroscience, and is perhaps less likely to impede progress in android science because the workings of an android 'brain' are more accessible than those of a human brain.

From the standpoint of human–robot interaction, androids have a number of advantages compared to mechanical-looking robots. As MacDorman and Ishiguro (2006c) note, the goal of designing sociable robots is problematic because of a lack of consensus concerning how such a robot should look or behave. However, a good point about an android is that it is modelled on a human being: we know how a human being looks and we have a rough idea how one ought to behave—and finding out what we do not know is a fascinating pursuit in itself, with the potential to contribute to disciplines other than human–robot interaction. Lindblom and Ziemke (2006) also note the enormous growth expected in the personal service robot sector, where a humanlike interface is understood as being ideal because it requires the least training of users, it is capable of transparent intentional and emotional expression, and it has morphology best adapted to human surroundings. However, humanlike forms might also seem eerie, if not designed with care (Hanson 2006).

As shown in figure 1, android science sets up a reciprocal relationship between the social and cognitive sciences, which analyse human interaction and its mechanisms, and engineering and computer science, which have provided methods and technologies for the development of electromechanical devices. Ishiguro (2006) emphasizes how android science can lead to a convergence of the cognitive sciences and engineering; but just as importantly, android science sets up a reciprocal relationship between social psychology and cognitive neuroscience by providing a platform for exploring the relationship between neural mechanism and interpersonal activity. Thus, an android could provide a vehicle for unifying sciences that study human beings at widely varying levels of description (MacDorman and Ishiguro 2006b). Some researchers construe android science even more broadly than this, taking it to include the study of all the effects of engineered human likeness, such as the relationship between anthropomorphism and human perception (Hanson 2006, MacDorman and Ishiguro 2006b) or the impact of androids on society (Ishii 2006), law and rights (Calverley 2006), and notions about what it means to be a person (Kahn *et al.* 2006, MacDorman and Cowley 2006, Ramey 2005a, b). The papers in this special issue are to be understood in this broader sense of android science.

In this issue, Turkle *et al.* (2006) take an intimate look at seniors and children and their relationships with Paro, Aibo and other robots that demand nurturance and inspire love. Nursing home residents who cannot even care for themselves are encouraged to partake in the fantasy of 'caring' for Paro, a robot built in the image of a baby seal. While this care-giving can reinforce the fiction that the seniors have autonomy, it can also disrupt it by casting them in the role of children left to be amused by a toy. Nevertheless, Paro can prompt reflection on what it means to be alive. According to Turkle *et al.*, Paro and other *relational artifacts* seem to exist at the boundary between life and death, which can be a cause for sympathy among nursing home residents whose predicament is similar. Such ambiguities might be magnified for androids. Unlike pet robots, they link human and machine in ways that could prove threatening to our personal and human identity (Ramey 2005b).

Relational artifacts also raise a number of ethical concerns voiced by Turkle *et al.* (2006). Given staffing shortages at nursing homes, will the care for the elderly, who are among the most vulnerable members of society, be left in the hands of machines? This question is no longer just hypothetical, but a matter for public policy debate in Japan (Barry 2005). Further questions concern the social and psychological impact of 'sociable' robots: Is it ethical to design robots to exploit parental instincts, robots that may compete with worthier and needier targets of affection, such as real children and pets? And if androids and other robots are designed to satisfy human needs without causing disappointment, does this not encourage narcissism among their users? And is it right for seniors and children to form relationships with robots that, despite their therapeutic effects, lack authenticity—relationships that encourage and depend on human projection in which the robot's understanding is but a pretence sustained by simple mechanisms, such as eye contact and the expression of emotion?

In this issue, Calverley (2006) explores the ethics of building conscious, humanlike machines from the standpoint of the courts. Could not one of the aims of android science, the construction of an artificial human being, be immoral from its inception, given that human beings are legally protected from being experimented on without informed consent? Arguments that would extend protection to androids sometimes parallel and sometimes diverge from those of the animal rights and right to life (anti-abortion and anti-euthanasia) movements, and much hinges on whether androids could be determined to be conscious or suffer pain (Dennett 1994, MacDorman 2004). However, a court or legislative body could declare androids to be legal persons for no other reason than its need to regulate them; but this may have the unforeseen effect of extending constitutional rights to androids.

While human laws codify our notions of how things *ought* to be (Calverley 2006), norms of another kind play a crucial role in child development. In this issue, Cowley and MacDorman (2006) propose that 'through bottom-up processes, norms enable individuals to engage with recurrent patterns that constitute higher-levels of organization beyond the body' (section 3). Norms serve as an evaluative standard, giving meaning to events from the standpoints of the actors involved. Thus, they allow the dynamics of an interaction to be interpreted with reference to individual, interpersonal, group, inter-group and socio-cultural systems. Cowley and MacDorman examine how infants and adults can meet, flout and surpass norm-based expectations by orienting their behaviour to cultural values (Cowley *et al.* 2004, MacDorman *et al.* 2005). They compare the norm-based aspect of how Tetris players develop epistemic actions to how a 'baby's brain self-organizes under the dual control of mother and infant' (section 1): in Tetris norms link a player's affectively attuned cognitive systems to the higher-level system of player-in-the-game. Likewise, norm-based epistemic actions allow the infant to alter its cognitive state by exploiting the mother as a cognitive resource within the higher-level system of the mother–infant dyad. Lindblom and Ziemke (2006) then describe how, at about 9 months of age, the self and joint attention may emerge through social scaffolding during self-induced locomotion.

Given the importance of norms in orchestrating human relationships, Cowley and MacDorman (2006) consider how androids and other robots could exploit them to relate better to people. Also in this issue, Sugiyama *et al.* (2006) analyse cultural norms in human–robot interaction with respect to drawing attention, while Walters *et al.* (2006) analyse them with respect to keeping distance. Lee (2006) considers a number of factors including mimicry, responsiveness and disclosure that could enhance empathy, intimacy and trust in human–android interaction (cf. Cowley and Kanda 2005). 'Women appear to be the gatekeepers of intimacy', writes Lee (2006, section 4.1), being much more likely to receive comfort from others and to have others receive their comfort. Other studies have suggested that women are often motivated toward intimacy and affiliation, which suggests that, to support intimate relationships, humanlike robots could benefit from a female form.

Most of the papers in this special issue grew out of the first workshop on the topic, 'Toward social mechanisms of android science', at the *27th Annual Meeting of the Cognitive Science Society*, on 25 and 26 July in Stresa, Italy (MacDorman and Ishiguro 2006a). A second workshop was collocated with the *28th Annual Meeting*, on 26 July, in Vancouver, Canada. The work of the authors of this special issue should serve as a basis for further developments in android science.

References

M. Asada, K.F. MacDorman, H. Ishiguro and Y. Kuniyoshi, "Cognitive developmental robotics as a new paradigm for the design of humanoid robots", *Robot. Automation*, 37, pp. 185–193, 2001.

B. Barry, "Better than people: why the Japanese want their robots to act more like humans", *The Economist*, 24 December 2005.

D.J. Calverley, "Android science and animal rights, does an analogy exist?", *Connection Sci.*, 2006 (this issue).

S.J. Cowley and T. Kanda, "Friendly machines: interaction-oriented robots today and tomorrow", *Alternation*, 12, pp. 79–106, 2005.

S.J. Cowley and K.F. MacDorman, "Simulating conversations: the communion game", *AI & Soc.*, 9, pp. 116–137, 1995.

S.J. Cowley and K.F. MacDorman, "What baboons, babies, and Tetris players tell us about interaction: a biosocial view of norm-based social learning", *Connection Sci.*, 2006 (this issue).

S.J. Cowley, S. Moodley and A. Fiori-Cowley, "Grounding signs of culture: primary intersubjectivity in social semiosis", *Mind, Culture Activity*, 11, pp. 109–132, 2004.

D.C. Dennett, "The practical requirements for making a conscious robot", *Phil. Trans. R. Soc. Lond.*, 349, pp. 133–146, 1994.

D. Hanson, "Exploring the aesthetic range for humanoid robots", in *Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*, pp. 39–42, 2006.

D. Hanson, A. Olney, S. Prilliman, E. Mathews, M. Zielke, D. Hammons, R. Fernandez, and H. Stephanou, "Upending the uncanny valley", in *Proceedings of the Twentieth National Conference on Artificial Intelligence*, Menlo Park, CA: AAAI Press, pp. 1728–1729, 2005.

H. Ishiguro, "Android science: conscious and subconscious recognition", *Connection Sci.*, 2006 (this issue).

K. Ishii, "A study on strategies to make good use of android science", in *Proceedings of the ICCS/CogSci-2006 Poster Session*, pp. 123–124, 2006.

P.H. Kahn, Jr, H. Ishiguro, B. Friedman and T. Kanda, "What is a human?—Toward psychological benchmarks in the field of human–robot interaction", in *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2006.

B. Lee, "Empathy, androids and 'authentic experience"', *Connection Sci.*, 2006 (this issue).

J. Lindblom and T. Ziemke, "The social body in motion: cognitive development in infants and androids", *Connection Sci.*, 2006 (this issue).

K.F. MacDorman, "Extending the medium hypothesis: the Dennett–Mangan controversy and beyond", *J. Mind Behav.*, 25, pp. 237–257, 2004.

K.F. MacDorman and S.J. Cowley, "Long-term relationships as a benchmark for robot personhood", in *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2006.

K.F. MacDorman and H. Ishiguro, "Toward social mechanisms of android science: a CogSci 2005 workshop", *Interact. Stud.*, 7, pp. 289–296, 2006a.

K.F. MacDorman and H. Ishiguro, "The uncanny advantage of using androids in social and cognitive science research", *Interact. Stud.*, 7(3), pp. 297–337, 2006b.

K.F. MacDorman and H. Ishiguro, "Opening Pandora's uncanny box", *Interact. Stud.*, 7(3), pp. 361–368, 2006c.

K.F. MacDorman, T. Minato, M. Shimada, S. Itakura, S.J. Cowley and H. Ishiguro, "Assessing human likeness by eye contact in an android testbed", in *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society*, 2005.

M. Mori, "Bukimi No Tani (the uncanny valley)", *Energy*, 7, pp. 33–35, 1970.

C.H. Ramey, "For the sake of others': the personal ethics of human–android interaction", in *Proceedings of the CogSci 2005 Workshop: Toward Social Mechanisms of Android Science*, pp. 137–148, 2005a.

C.H. Ramey, "The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots", in *Proceedings of the Views of the Uncanny Valley Workshop, IEEE-RAS International Conference on Humanoid Robots*, 2005b.

O. Sugiyama, T. Kanda, M. Imai, H. Ishiguro and N. Hagita, "Humanlike conversation with gestures and verbal cues based on a three-layer attention-drawing model", *Connection Sci.*, 2006 (this issue).

S. Turkle, W. Taggart, C.D. Kidd and O. Dasté, "Relational artifacts with children and elders: the complexities of cybercompanionship", *Connection Sci.*, 2006 (this issue).

M.L. Walters, K. Dautenhahn, S.N. Woods, K.L. Koay, R. te Boekhorst and D. Lee, "Exploratory studies on social spaces between humans and a mechanical-looking robot", *Connection Sci.*, 2006 (this issue).

T. Ziemke and J. Lindblom, "Some methodological issues in android science", *Interact. Stud.*, 7(3), 2006.

Guest Editors:
Karl F. MacDorman
Hiroshi Ishiguro