AI & Soc
DOI 10.1007/s00146-007-0145-y

ORIGINAL ARTICLE

# Socio-ethics of interaction with intelligent interactive technologies

Satinder P. Gill

Received: 5 March 2007 / Accepted: 20 June 2007


**Abstract** Socio-ethics covers the relation of the individual with the group and with society, as the individual acquires the skills for social life with others and the conduct of ‘normal responsible behaviour’ (Leal in AI Soc 9:29–32, 1995) that guides moral action. For a consideration of what it means to be socially skilled in everyday human interaction and the ethical issues arising from the new conditions of interaction that come with the integration of intelligent interactive artefacts, we will provide an analysis at multiple levels of these phenomena and draw on a variety of application domains, for example, healthcare and interactive media.

## 1 Introduction

Socio-ethics will be considered as the ethics of sociality and understanding what that means. This covers the relation of the individual with the group and with society, as the individual acquires the skills for social life with others and the conduct of ‘normal responsible behaviour’ (Leal 1995) that guides moral action. For a consideration of what it means to be socially skilled in everyday human interaction, and the ethical issues arising from the new conditions of interaction that come with the integration of intelligent interactive artefacts, we will provide an analysis at multiple levels of these phenomena and draw on a variety of application domains, for example, healthcare and interactive media.

When aesthetics is abstracted from the ethics of social norms (for example when artists explore genetics for creating living art forms, Gigliotti 2006) then this

S. P. Gill (✉)
School of Computing Science, Middlesex University,
Ravensfield House, The Burroughs, Hendon, NW4 4BT London, UK
e-mail: sattisan@yahoo.com

destabilises normal responsible behaviour. The consequences of this in the case of interaction technologies are a separation of how we judge action in the virtual representations of human life (for example, in virtual computer mediated spaces), in the artificial simulations of human life (for example robotics) and how we judge human life in the social everyday co-existence. The symbiosis between the artificial and the real needs a symbiosis between aesthetics and ethics in order to design socially sustainable systems. Both aesthetics and ethics are beyond data, beyond information, and beyond explicit representation as rules and evidence (Scha 2007). They lie fundamentally in human values and are made visible in human practices.

In our everyday communications and engagement with each other we could say that we feel resonance or dissonance in our connection with someone or no sensation at all. Such resonance may be described as having an aesthetic quality. The aesthetic resonance in human engagement, it is proposed, lies at one fundamental level in our timed movements of body and voice and breath, i.e., the rhythm, pulse, tempo, pitch. This resonance in the prosody of body, voice and breath shapes community and gives a sense of the nature of the community, that is, being social. A common experience we have of the aesthetic resonance of being a community is when we sing and dance together. The aesthetics and ethics of human engagement come together in the feeling that the communication is working well, that the communication makes sense and has a natural flow. We take this fine and precise coordination of the movements of our bodies and voice for granted and only notice it when there is a problem in the communication such that we feel awkward, feel dissonance, or just do not feel we connect. However, we are able to repair or resolve the problems if we share the ethics of interaction. Ethics is about knowing how and when to act and involves social judgment about appropriate behaviour. When we share communicative ethics, we have a mutual intention to be committed to engage our selves. The relationship of aesthetics and ethics in our everyday interactions carries in its timed movements and modularity, the past, present, and future possibilities of meaning.

## 2 Social intelligence, human synchrony and coordination

The advent of various kinds of multi-modal and multi-sensory interactive technologies is having an impact on the social dimension of being human. Two core aspects of 'being human' in social life that we focus on in this paper are 'coordinated autonomy' and 'collective action'. Autonomy is a concept that can be misunderstood as meaning the 'individual' and systems built on this principle may impact on human cognition in a manner that does not afford co-evolution in the symbiotic interactive structure where symbiosis is necessarily rooted in the social interaction. This constrained concept of autonomy, as meaning the individual, needs to be rethought as existing in the social dynamics of everyday life, for example, in the way we talk, move, engage with each other, misunderstand and repair misunderstanding, and in the way we search for information in our environment and make sense of it for our purposes. Much of this activity is taken for granted by us as we do it all the time and without awareness that we are doing so. Some call it 'social intelligence' (Gill and Borchers 2004). And we are very precise in our performance of social intelligence. Studies on the use of various kinds of large-scale interfaces, show how co-ordinated autonomy has various patterns of coordinated and synchronized body and speech movements, (emergent paralinguistic pulse periodicities), which become visible when they are interrupted by an interface that does not support them. These pulsing periodicities of coordination and synchrony perform an important function in sustaining our interactions (Gill 2004b).

These patterns of coordinated autonomy enable us to ground our understanding of any communicative situation, and the emergence of these pulse periodicities are a form of social entrainment, fundamental to social cohesion. Entrainment occurs when there is the coupling of oscillations of 'autonomous' entities/beings (when the rhythms of two or more people become coupled) (Large and Reiss-Jones 1999; Cross 2007). Take the example of walking in step with someone. It is not something we do consciously, but we notice it when we find we are out of step with the person we are talking with and try immediately to get back into step with them. Furthermore, such autonomy has a musicality developed within carer–infant interactions often described as motherese (Miall and Dissanayake 2003; Trevarthen 1994). We never talk to babies as we would speak with an adult. Rather we exaggerate our prosody of voice and gesture, and modulate the tones almost in a song like mode. Considerable research suggests that this is no accident. In this collaborative interaction, the baby learns about the patterns of social interaction involving those of coordination and cooperation which are embedded in the culture, and this learning facilitates the acquisition of knowledge essential for survival within the culture. Coordinated autonomy may be considered therefore as consisting in salient phenomenal 'beats' that already have the phenomenological experience embodied, in other words they have the iconicity of culture (Gill 2005, 2007).

Interactive technologies (e.g., the internet, multi-sensory ambient technologies, video-conferencing, multi-modal interactive systems, etc.) impact on coordinated autonomy and social intelligence when we engage with them and via them. These technologies are operating with us at the dimensions of human performance and their design needs to be sensitive to the very basic levels of 'being human' in that performance structure, otherwise it alters it and affects our self-expression and identity. For example, it may be harmful to subject a very young child or baby to an interactive technology that impinges on their development of social interaction with others.

The socio-ethical concerns lie in the impacts on human social intelligence, i.e., human cognition, communication, learning, and social cohesion, and particularly the impact on tacit knowing, cognition, and collective action. Such impacts may have an effect on social/cultural inclusivity and access. *Tacit knowing* is an essential concept of human knowing and action formed by Polanyi in the 1960s and adopted in a range of fields of study from hermeneutics, social studies of scientific knowledge, historians of science, engineers, information and communication technologists. It is the unspoken dimension of human knowledge, formed in practice (through action, adaptation, negotiation, repair of breakdown), and is considered in this paper as being essential for skilled communication and skilled judgments/decision-making.

*Collective action* is achieved through our understanding of the performance of representations of the tacit dimension in our communication—gestures, non-verbal cues, speech, silence, touch, and other structures of information in our environment, evident in how we perform with them when we are engaging with others (Gill 2007).

Social interaction in everyday life is about making sense and this is shaped by social intelligence. It is the dynamics of the interplay of our personal (self) and experiential (our experiencing), objective (articulation and abstraction) and subsidiary (knowledge that resides in those around us, family, friends, colleagues) dimensions that makes social interaction meaningful and learning possible. The design of interactive intelligent technologies is bounded by three orders of rationality (Gill 2006, 2007), the first order being to observe the present and describe and define the observations (conceptual gap). This raises a first order gap between what will be termed 'actuality' (the experiencing that draws on past, present, and expectations of future) and 'reality' (the observed present) (Gill 2006; 2007; Uchiyama 2003). The second order of rationality is design (technical) limitations that come in designing the technology—this creates the second order gap that is the design gap between human and the machine. The third order rationality of design is technical competence, the application gap—that can lead to the breakdown and disruption of our everyday practices as we interact with the technology, with the consequences that we become deskilled in our social intelligence without our even realizing it. The reason we cannot realize it is because this interaction is with our processes of embedded knowledge which are shaped, shared, and learnt within communities of practice. This is a socio-ethical concern to be addressed.

### 2.1 Embeddedness and embodiment

The tacit dimension of knowledge is embedded and embodied in our social interactions. To embody is to give concrete form to an abstract concept in performance. Embedded knowledge is about fixed and innate structures and we do not have conscious access to these but they are essential in our working, learning, and sharing knowledge within 'communities of practice' (Wenger and Lave 1991). Embedded knowing is 'work-related' practical knowledge, not expressed but implied and simply understood, often associated with intuition (Murphy [Ongoing Ph.D thesis]). It is non-codifiable knowledge acquired through informal 'take-up' of learned behaviours and procedures—embedded in the individual ingrained in their practice and expertise, only expressible through proficient execution and forms of learning involving demonstration and imitation (Fleck 1997). All knowledge has embedded dimensions in a spectrum—at one extreme it's almost completely embedded—semi-conscious, unconscious held in heads and bodies; other extreme almost completely explicit, codified, structures, accessible to people outside its origination. Certain aspects of embedded knowledge cannot be transferred—i.e., aspects involving insight, intuitions, decisions based on 'gut' feel (Leonard and Senisper 1998). Knowledge in organisations evolves and becomes embedded through the evolving work culture and within individuals through their social interactions within communities of practice.

A reflection on socio-ethics and embedded knowledge is important because intelligent interactive technologies, be they virtual/artificial agents, etc., necessarily seek to engage us at the embodied and embedded knowledge levels, and this raises socio-ethical concerns for the effect upon our interaction practices.

## 3 Normal responsible behaviour

Socio-ethics is concerned with 'normal responsible behaviour' (Leal 1995) in everyday interaction. It is concerned with bringing together the utilitarian perspective (i.e., the greatest good of the greatest number should be the guiding principle of conduct) and the moral perspective (i.e., socially responsible behaviour) within an ethical framework of socio-ethics as code of conduct. The realisation of the tacit and the objective dimensions of our knowledge resides in our co-action with others hence a code of conduct is necessarily a code of practice. It is only through practice that we can conduct socially responsible behaviour. We attain the tacit dimension of this behaviour in practice, and in this paper it is suggested we attain it as babies through out interactions with out mothers, and are not consciously aware of it, i.e., noticing it. However if something goes wrong we are made aware of the quality of the tacit-objective relation as something that is missing as we experience a fragmentation of the fluidity of the interaction. What was not in our focal awareness is now made objective, rendering it elusive. We could describe this visibility as the gap between the tacit and the objective. In the same vein we do not notice that someone is being ethical, as this is part of normal responsible behaviour, however, when someone behaves in a socially inappropriate manner, we become aware that this person is unethical. We need an ethical code of conduct to provide the checks and balances in our social system of interaction with interactive technologies.

## 4 Actuality and reality gap

The design of intelligent interactive systems makes salient the interplay between technology, application domain (context), organisational domain (embedded knowledge of organisation as process), and cultural domain (moral and social values). Interactive technology needs to be considered as more than simply the design of an artefact. This interplay provides a holistic framework for the design and application of interactive intelligent systems where the cultural domain drives the application process within an organisation so that interaction between human and technology could function in a manner that allows for normal responsible behaviour. Within the organisational context we look at gaps of responsibility and of knowledge, between actuality and reality, and consider the consequences of disengagement from ethical actions, arising out of the interaction with intelligent interaction artefacts/agents.

## 5 Ethical issues arising from interaction

### 5.1 Privacy and ethics

Interactive intelligent technologies range from digital devices that we wear in our clothing to virtual agents and moving artefacts in our social spaces. Privacy concerns our person in relation to others, involving our social values (from cultural practices in society), organisational values (through communities of practice), for example, the trust between a patient and a doctor (Carew and Stapleton 2005). Integrating interactive technologies into the organisational domain has implications for disturbing and altering the value system and how we trust and interact with others. This effect on social values and trust is of socio-ethical concern.

Consider the case of wearable technology. At a very basic level, when we wear anything be it clothes, jewellery, watches, mobile phone, etc., we are communicating something about ourselves to others and with others and expressing identities. Lets add to this the complexity of digital and networked devices such as wearable technologies (Smith 2006) and within a context such as health care architecture. Imagine a patient who is wearing a hybrid digital–biological material that monitors her heart rate and temperature, etc., and at the appropriate time provides her with medicine. The monitoring of the patient's health is undertaken via a network that links the patients information with the organisation's data base that is linked to monitoring and interacting networks that may lie outside medical information, for example, about her health and life insurance, diet, etc. Access to this information could lead to breaches of patient information and the patients may have no way of knowing whether such breaches have occurred. Third parties do use patient data and some demand patient information beyond that necessary for their purposes. For example, insurance companies ask for medical histories and results of tests, etc. (Slack 2001). In some cases the motive of an insurer seeking access to patient data is to eliminate high-risk cases, i.e., to refuse them. This is unethical and is a privacy concern. Third party access to data should be limited to a strictly need to know basis, as an ethical principle.

The patient and clinician relationship is traditionally shaped within communities of practice where both perceive each other as persons with social values and organisational values that enables them to trust each other. One critical practice on the part of patients is to confide in their clinician, and the opportunity to confide is one of the major functions of privacy (Pedersen 1997; Carew and Stapleton 2005). The use of information systems in health care practices could lead to reduced contact between patient and clinician and this could have an effect their trust relations. An individual's acceptance of an information system will depend on the trust between the individual and their doctor. Confiding by patients through electronic means necessarily depends on trust relations because of the private nature of the information involved.

Once interactive technology enters this relationship, it alters the patient–clinician relationship which is now being monitored. Furthermore, the clinician is also being monitoring by a monitoring network that in turn affects their relation to their patient. The information flows from the monitoring network to the patient and the clinician and updates their information.

The technology is conceived as a tool for standardising care and making healthcare personnel more efficient by controlling aspects of their lives. This has meant that doctors/clinicians no longer have control over how they work and treat their patients. In fact, there is now two-way dynamics of information flow and monitoring where the monitoring from the patient and clinician have different purposes. This again raise issues about privacy and information data flows and ethics of interaction. An example given further below from healthcare architecture provides an example of the effects of information technology and monitoring systems within an organisational context.

In summary, the discussion above identifies that it is the values (morals of normal responsive behaviour) that drive trust and that, in turn, both values and trust drive ethics (rules/codes of responsible conduct).

## 5.2 Managing privacy: the case of healthcare architecture

Healthcare informatics has underestimated the value of privacy treating it largely as equivalent to data integrity, security and availability, describing it in regulation terms, i.e., controllable in the information systems world. The management of privacy is about managing the calculation, i.e., the data, and leaving the judgement out (value driven). It is this separation of calculation and judgement that leads to the breakdown of communications and relationships. This is a socio-ethical concern for interaction.

Privacy is seen as a boundary control process (management of privacy) whereby individuals control how much or little contact they have with others at any given time in a given situation (Carew and Stapleton 2005). Dimensions of privacy include the social, informational, psychological and physical. The physical refers to the environment, e.g., office, home, hospital, etc. ‘Social’ refers to the freedom of individuals to withdraw from or enter into interaction with others. ‘Psychological’ is related to the social dimension but refers only to the individual’s psyche and ‘informational’ refers to the individual’s ability to control their personal information. However, we need to extend this picture of privacy from the individual to the relationship of the individual within communities that the individual engages within. From this perspective we consider the organisational values and wider social values as driving forces for ethics and interaction.

Carew and Stapleton (2005) consider the factors of privacy within the healthcare domain. They argue that these include safety, i.e., ensuring that correct patient data is recorded and appropriate treatment provided and that there is appropriate decision support, auditing system, and controlled access. Other privacy factors include the consideration of harm to individuals and their families as a side effect of healthcare long after the medical care/treatment. These are the issues that the designers of interactive systems need to take into account. In designing information systems we need to consider a broader definition of privacy where an individual suffers no harm (physical or otherwise) as a side effect of undergoing health care at any time during or after it has been completed. This would be an ethical principle.

The danger to patient safety comes from electronic storing and processing of data where there could be unauthorised access to data, e.g., changing of records and potentially dangerous treatment being provided. There are issues around having large quantity of personally sensitive medical information being stored as this could lead to social harm if a third party obtained this information. Physical harm or psychological occurs if the data infers that the individual deviates from expected norms (personal pro ling) and the illegitimate use of this information will affect their physical, social, and psychological lives. Designers of information systems need to recognise this and consider it as a design issue.

From a different perspective on safety, that of civil liability (Collste et al. 2006), it may be dif cult to clearly establish malpractice where an on-line prescription is issued. Whereas in the traditional clinician–patient relation a clear duty of care exists, it is debateable whether a doctor who prescribes medication on-line attracts the same duty of care as there is no direct physical or verbal contact with the patient (Kahn et al. 2000). Within current networked environments can on-line patients legally address the issue of medical malpractice especially where the medical practitioner is located in a distanced nation state? If not, what ethical mechanism needs to be put in place to address those issues.

Patient clinician contact in telecare services to deliver healthcare remotely, or the use of electronic health records information to make diagnoses instead of physically visiting patients, leads to the disembodiment of this relationship. Dreyfus (2001) notes that ‘telepresence can never give us a sense of the reality of far away things, nor can it convey a sense of trust of distanced human beings’ (p. 98). We can never truly get a grip on the reality as the true context cannot be felt arti cially from a distance. Healthcare workers cannot fully understand the reality of the patient where there is the lack of context that can only be established by the physical embodied presence. In treating the unreal (or hyper-real) patient the health care professional may unknowingly take risks. This is an ethical concern.

Rosenberg (2006) notes that implanting chips into humans to provide temporary identi cation is a serious step on the road to massive violation of privacy rights. In Canada there are privacy guidelines to address this focus by the of ce of the Privacy Commissioner RFID and focus 2006, and these are: (a) to focus on RFID information systems and not technologies; (b) that privacy and security must be built in from outset (at the design stage), and (c) that design and application should involve maximal individual participation and consent. Rosenberg’s concern is that it seems to be taken for granted that the bene ts to information society of intelligent artefacts is so obvious that any critical review is unnecessary, and this has to be challenged. Traditional ethics that de nes appropriate human behaviour in a variety of situations must be extended to the situation of human interaction with advanced technology. Ethical behaviour by humans must regularly be reinforced, given differences in ethnicity and religion as well as changing values in society. This is an ethical challenge for researchers to address.

## 5.3 Decision making, risk, and error

Technology is often perceived as a solution to a particular problem and the design focuses on problem solving with speci ed goals to be achieved and the best/

efficient/effective ways to achieve those (Miller 2006). Problem solving assumes there is certainty. If nothing is certain then following an action which is predetermined is potentially dangerous or may lead to error. We may want to eliminate the possibility of error before it can do too much damage. We may want to expose any unnoticed flaws in what we know and in what we do. Our aim is to find the flaws. Decision making, according to Miller, is ‘ad hoc’. Risk is in the future so we can only hypothesise or conjecture and this is the dilemma of evaluating risk. In this sense, we can only realise risk when we use systems in special application domains. Designing predetermined systems means that when error occurs it is too late to do anything about it. To overcome this error we need to constantly re-evaluate risk and redesign the system. You need to modify the system all the time.

Collste et al. (2006) argue that if intelligent systems are making decisions regarding the health status of a patient, e.g., using intelligent clothing, and deciding how much medicine to be give by incorporating decision making methods, e.g., a multi agent system, it is important to know the basis on which decisions are being made. For example, knowing the accuracy of data, what constitutes a decision, and how does this affect the right of the patient to refuse treatment (the decision). There is also the issue of informed consent: if the patient does not understand how the technology works or its possible consequences, will the technology have to be explained to patients together with all implications of data transfer and medical impact? What are the criteria for informed consent? Should a list be devised? how will it be judged that such consent has been given?

Our use of learning automata, intelligent agents and networks of agents, raise issues of responsibility (Marino and Tamburrini 2006). When the technology learns in an unpredictable manner the person using the technology and the person who has designed it cannot be held responsible, as we cannot be responsible for what we do not have control over. This raises a responsibility gap for traditional responsibility ascription and consequently, an ethical gap. This requires us to draw upon alternative frames for making ethical judgements, e.g., case based legal reasoning about relations, for example, parents, tutors and children, people at the workplace (disputes), and producers of goods, e.g., getting profits when their workers' health is suffering—even where there is no causal relationship. We do not need to look at the causal reasons for an action but need to look at assigning the social responsibility for the action.

Decker (2006) illustrates the responsibility gap through the exemplar of the automated pilot where human–machine control is identified as the main source of accidents. One of the ethical dilemmas according to him is that humans are used as a means to an end that they themselves have no interest in. This is seen within the framework of utilitarianism. This framework has been used to restrict dignity and autonomy (e.g., profiling and tagging) of persons to achieve higher level ends (e.g., national security) for a more comprehensive utility. Seeing this issue from a legal perspective it is argued that the human is responsible for say human–robot collaboration and that this must be delimited to the common actions, as humans do not know all of the robot's skill. However, this raises the responsibility gap that is discussed by Marino and Tamburrini (in this collection). When discussing the integration of interactive technologies in society we need to take account of the following: how do you meet/define the ends/needs in an ethical way? What's the purpose? What level of control should be given over to a machine (interactive technology) and how much should one keep, in levels and kind?

Torrance (2006) makes the point about conscious experience being at the heart of ethics where what concerns us is existence and the quality of experience. The difference between real and simulated consciousness is more clearly marked as being an ethical signi cant boundary than that between real and simulated intelligence. Property of being human—we do not like being the instrument of other's desires. But machines have to be treated like this. There is almost no ethical behaviour in relation to machines. Ethical issues only arise when interaction arises, i.e., we have to be responsible when we design machines to make sure they do not behave in ways that give rise to unethical action. Ethical rules and responsibilities include to be able to desire or want ethically appropriate situations, to have a morally virtuous character or disposition and appropriate moral emotions or sensibilities. Question of autonomy: can machines act following ethical codes of conduct. Even when machines/arti cial agents may display certain kinds of functionalities that may display richer kind of autonomy in decision making, e.g., the case of banking systems, we cannot assign the arti cial agent systems and judge them as having inherent responsibility.

### 5.4 Augmentation or merger

Warwick and Cerqui (2006) discusses ethical issues about what happens when technologies are built with the purpose of augmentation or merger of the human with the machine. She illustrates these concerns by the example of bionic implants, and brain computer integration and asks a general question of whether we are developing technology for therapy or for human enhancement: where therapy is treatment for de ciency or disorder, and enhancement is an extension or improvement of some capacity, characteristic or ability of the human. Brain–computer interfacing has implications for whether it is a merger between brain and machine or the augmentation of the brain by the machine. These raise ethical issues: what are ethical standards against which we can measure brain enhancement? This depends on de nition. The de nition of health is a dynamic one. Consider the World Health Organisation (WHO) where health is not just absence of disease but is also about well-being. In order to consider ethical issues we need to know how the de nition of health is evolving and at the moment, because the de nition is evolving, ethical evaluation is subjective. The concept of disease is now very elastic. This is in uenced by the complexity of the environment, i.e., in designing machines or tools for augmentation of our human capacities, how do we evaluate this augmentation? And before we can evaluate, we need to de ne this adaptable environment. The problem is not just about the design of technologies but about the evolving environment in which they are placed that creates a dif culty in identifying what is normal. Even the boundary of the normal and the pathological is shifting and transferable and the body needs to be plastic in this picture.

This has implications for socio-ethics. Ethical concerns are linked to the environment in which technology is applied. The interaction between technology, human, and environment needs to be understood and this is an ethical requirement. In this evolving picture of a plastic environment how do we define norms of ethics and responsibility?

### 5.5 The digital body—body as object, as disembodied

In the above discussions, we have covered a number of ethical issues, e.g., privacy, responsibility, control, and monitoring. Another ethical issue concerns the conception of the human body as an electronic body, caused by the integration of new interaction intelligent technologies. Rodotà (2006) argues that the human body is undergoing transformation because the very autonomy of the individual is being affected as these technologies are used to monitor (tag and track) and direct us remotely using, for example, electronic implants. The body is becoming a new object, a password (e.g., iris scan, finger print, etc.). Networks of interactive technologies are distributing the body and creating a networked person. Bionetworks, including biometric sensors, wireless connections, will enhance opportunities for round the clock medical support and multiply the opportunities for keeping bodies under control and violating their integrity. We are confronted with changes that have to do with anthropological features of individuals, with culture and social life. Persons who are permanently on the net are configured little by little in order to transmit and receive signals that allow tracking and profiling of movements, habits, and contacts, and thereby modify the meaning and contents of an individual's autonomy. Removal of the body and identity into virtual networks transforms persons into 'interactive beings that transmit and receive data'. This shaping of human interaction by the electronic conception of the body raises issues of physical and mental integrity (Article 3, Charter of Fundamental Rights of the European union). It reinforces the information model of signal and receiver that assumes individuals are autonomous entities rather than the fluid model of human interaction and cognition based on coordinated autonomy. The fragmentation in human engagement and social intelligence is a socio-ethical issue.

### 5.6 Moral disengagement and externalisation

Reinterpreting the work on bio-robotics and disengagement by Salvini et al. (2007; 2007) we can see that the integration of interactive technologies in complex social systems may lead to disengagement in human relations. For example, interaction with toys like the Tamogochi focuses on eliciting empathy, emotional attachment, and other emotional responses. The consequences of this are that it might endanger sociability and cause abstraction and moral disengagement. They identify the danger of applying the design and application of such technologies from the functional domain to the fundamentally human, as being the abstraction of what it means to be human. In relation this, in a broader organisational context, Moshowitz (2007) discusses how externalisation relieves us of responsibility and impedes the search for solutions to social problems caused by applications of technology.

## 5.7 Ethics is fragile, goodness is not

In essence, what the discussion so far has covered suggests that our normal responsible behaviour is proven robust and reliable most of the time and is only in danger if we put too much pressure to bear on it by integrating inappropriate technologies. Since this normal responsible behaviour embodies essential human goodness, the danger is that we will erode this dimension of being social by focusing on abstracted ethics which is fragile. This is a fundamental socio-ethical concern. Leal (1995) states that information, communication and automation technology is ethically neutral and this raises urgency for ethics as in 'normal responsible behaviour', as traditional ethics is then fragile. This technology raises issues concerning the loss of skill, communication, identity (sense of self), control, and privacy. Leal points out that one person's loss could be another person's gain so affects are not clear cut and depend on the contextual levels.

The comparison between people and computers creates the 'computer metaphor'. The machine metaphor was relatively harmless, but because we can assume that these machines 'think', in the computer metaphor, people become conceived as machines, and "this conception allows us to design the kinds of jobs we do, to pry into people's lives as we do, and suffer the kinds of TV programmes we suffer" (Leal 1995). Whether any of this is true or not, the point is that there are ethical issues and the people who build technologies disregard their existence and the computer metaphor is considered to be neutral and the technology 'value free'. Some would argue that technology should be value free otherwise politicians would prevent it from working. The question then is how we use it. Leal calls this 'neutralism' and opts for the weak version, i.e., technology could be used for good or bad purposes. However, the point is that there are good or bad consequences and for whom. Given this, a moral philosophy based on 'discourse' (e.g., Habermas 1981) becomes important—to bring differing opinions to full expression, and there is no unique method to do this because of increasingly complex environments of personal and corporate/organisational interests. Ethical reasoning and discussion in itself is unable to prevent bad consequences if it remains abstract. If an ethical theory is the masterplan underlying a technology, such a technology can also be liable to misuse, because of the factor of 'normal responsible behaviour' which is highly undervalued and underrated, almost ignored.

Leal cites the zoologist, Young (1951), said that it's not rationality or consciousness that that makes us human but the intricate network of cooperation 'on a steadily increasing space-time scale' that is social life. Cooperation has two aspects, the collective and the individual. The collective is often termed 'cooperation' by scientists and Leal would prefer to call it coordination (Argyle 1991)—this ranges from the intricacies of face-to-face conversation to the complexities of telecommunications, from simple localised division of labour to international transfer of capital and human resources. It works and it makes us human. There is an individual aspect of cooperation without which coordination would not be possible, namely 'normal responsible behaviour', and this would not be possible without coordination. Further above, this was discussed in terms of social intelligence.

However, the mutual dependency is not Leal's point. He believes that 'normal responsible behaviour' is central for any ethical consideration. We become blind to considering the normality of cooperative social life because it happens most of the time and instead we focus on human error, etc. In literature, goodness is only salient as heroism. But we only need heroes when systems break down. We should be focusing on designing systems that reduce the need for heroes and can only do that if we pay attention to normal responsible behaviour. Such behaviour embodies human goodness 'at its best'—it's reliable, robust, so reliable that badly designed systems are kept going because responsible human operators try to do the right thing at the right time. It does not work all the time and that is because of the way we design the system.

The embodiment of human goodness in normal responsible behaviour makes it invisible to usual ethical theorising and psychological and sociological analysis. For Leal, human goodness is integral to all human skills, embedded in any of them. Skills can be misused and ordinary goodness embedded in a skill is not infallible. However, skills are not supposed to hit the mark all the time, only most of the time, and hitting the mark most of the time does not make human goodness fragile unless the system puts inordinate strain on them. This fact makes it important to bear in mind the reality of 'normal responsible behaviour' when applying new technologies.

Leal describes the history of philosophical thinking as having been about the attainment of certainty and perfection (Dewey 1929) where practice and experience are accidental and too variable for this purpose. The twentieth century threw this picture out as we became used to theories being fallible and replaceable. But we still do not think of ordinary human experience and practice, ordinary judgement and skill, human goodness as the core. If we do, then we will be better able to build technologies which enable people to do what they 'try to do anyway'. If technology is neutral, then ethics, under the pressures and powers of new technologies is increasingly fragile—where ethics denotes both the theory and the moral structures and the inbuilt values of technology.

## 6 Human-centredness—socio-ethical framework

The above discussions present a picture of how our present everyday interactions, in various spheres and working and social life, with intelligent interactive and information technologies are increasing the complexity of our interaction environments in a manner we do not fully understand. This complexity is such that it is causing us to be aware of growing problems in human social practice that necessitate ethical addressing such as that of moral disengagement and loss of privacy. The interaction complexities are disrupting our communities of practice that embed social and cultural moral codes of conduct and within which we acquire normal responsible behaviour by sharing and engaging in the social practices. Our uses of interactive technologies are already overstraining that very 'goodness' that Leal has warned us to pay attention to. We need to re-appraise what we mean by human–machine interaction and understand its consequences for human social understanding and responsible behaviour, and use this to re-appraise design and application requirements for interaction environments, for example in bringing together the aesthetic and the ethic.

At present, human–machine interaction is conceived as a dualistic relationship and this is extended to networked spaces. The dualistic picture is a very limited conception of social interaction that actually consists in multiple structures of horizontal and vertical communications and operates at complex dimensions of human knowing in any communication setting. The dualism limitation creates gaps in the reality and actuality of our experiences in the world (Gill 2007). While the actuality is rooted in the past experiences and is shaped by the present reality and future possibilities, the reality is defined by the observable facts and data of "now", the present. In an attempt to model the user interactive interface, the designer is limited to objectifying the situation as 'observed', thereby excluding many of the possibilities of the situation as 'being observed'. This leads to a widening of the 'actuality gap' arising from the gaps between the actuality and reality of both the user and the designer. For example, in addition to the actuality gap our interaction experiences within our social life, especially the tacit dimension, are 'transparent' to us most of the time while they may remain hidden to users from other contexts. This limits the inclusive conception of user interaction from a cross-user perspective (Gill 2007). Because of the actuality–reality gap (tacit dimension) the rich interaction experiences of the social and cultural contexts of users remain excluded from the design of the techno-centric systems and technologies. This exclusion of the 'tacit' dimension impoverishes the design and applications of interactive intelligent technology within a wider societal context.

The consideration of interactive technologies within an ethical framework requires us to examine the processes of the design and application of these technologies. At present the design of interactive technologies and the concept of technology as a tool, is limited by a technocentric framework of reality. This limitation is recognised by proponents of the human centred systems framework (Gill 1996, 2004a, 2006, 2007; Cooley 1987) that aims to fill the gaps between the reality (observed) and actuality (practice, experiencing) of human interaction within the broader societal context.

At the heart of the human centred framework are the ideas of symbiosis, tacit knowledge, and machines with purpose. Symbiosis enables the continuous interrelationship between the personal, the experiential and the objective dimensions of human knowing and interaction. This extends the scope of seeing information and interactive technologies within an enriching and the holistic framework of technology, organisation, society, and culture, to support ethical action embedded in the conduct of normal responsible behaviour. The 'tacit' is seen here as the inter-relationship between the 'personal' (person) (feeling/experiencing) and 'experiential' (group) (collective experience/practice) and the objective (society) (Gill 1995). This articulation of the tacit provides a conceptual handle to articulate interdependent (symbiotic) relationships between the 'personal', the 'experiential' and the objective. It can be argued that part of the 'personal' knowledge can become part of the 'experiential' dimension over time during the process of participation in a group, and that part of the 'experiential' knowledge can become absorbed into the 'objective' dimension over time through the process of collaboration. Following the similar argument, it is proposed that part of the 'objective' knowledge can also be transferred to the 'experiential' domain, and part of the 'experiential' knowledge to the 'personal' domain. It is further proposed that this symbiotic idea of transference between 'personal' 'experiential' and the 'objective' forms the core catalyst for designing technological architectures for interaction design. Actuality–reality gap: 'in comprehending reality, we construct a model which represents facts we observe, and use this model to design technology/ techno-economic solutions. However in comprehending actuality, we experience actuality as it is practiced and experienced from within. The model of actuality represents practice (tacit dimension). Designs built on reality can only weakly be applied to actuality and we need to cultivate a design culture that overcomes this gap between what is experienced and the experiencing in order for there to be the merger between morality as practiced, and utilitarianism (rules/principles).

One way to handle this gap is to consider machines with purpose (Rosenbrock), an ethical concern. 'Purpose' widens the design and application scope of systems and technologies whilst enhancing and enriching human potential. This is a developmental view that shifts the predominant technical focus of technology, that is concerned with human functions or characteristics rooted in observation (e.g., based on observations of behaviour in social practices), to include ethical concerns and normal responsible behaviour. It can be argued that human centred systems may provide an ethical framework of 'governance' of technological architectures and of their operations as they impact societies and cultures. Human-centredness raises three questions:

1. Could we design a technology even if it technically possible? Feasibility question.
2. Should we design a particular technology? Social responsibility question.
3. Is the technology socially sustainable? Ethical question.

In order to understand the gaps between actuality and reality, we can take the health care scenario. In the human-centred perspective there are conceptual gaps between (a) the medical model of health, (b) the practice based model of health, and (c) the community model of health. These conceptual gaps are mirrored by the knowledge gaps between explicit knowing (e.g., medical), practice based knowing (e.g., health care professionals), personal (e.g., patient and community workers) and subsidiarily knowing (e.g., family and friends). Awareness and understanding of these conceptual gaps and knowledge gaps would help us to seek collaboration and interfacing between these models of health care. There are some fundamental research challenges that need to be met to develop a conceptual framework of health care which:

a. nds a symbiotic relationship between the medical, practice based, personal and subsidiary conceptions of health care provision.
b. lls the gap between the technical vision of health system and the social vision of health care systems.

c. develops a holistic approach that can bridge the gaps between the information and data handling requirements of the management and health care needs of both the professionals and users of health.
d. enables the development of interfaces that bridge the gap between the scientific knowledge of medical practice, the communities of practice knowledge of the health carers, and the social knowledge of the personal relations and social networks, within the health care and welfare chain (Gill 2006, 2007).

The health system example illustrates that there are interactive and overlapping roles of stakeholders who include medical and health care professionals, voluntary and community organisations, the patients and their families and friends. The complexity of these interactions and roles needs to be considered as part of design requirements when building interactive technologies in any organisational context and their consideration makes for a complex picture of ethical concerns and overlapping social responsibilities. The challenge is to come up with an understanding of socio-ethical conduct for any organisational complexity.

In summary, the design of intelligent interactive systems needs to attend to the interplay between technology, the application domain (context), the organisational domain (embedded knowledge of organisation as process), and the cultural domain (moral and social values). This need is made visible by the discrepancies and fragmentations arising with the integrations of such systems in our communities of practice, i.e., our interactions environments. Such fragmentations include responsibility gaps, gaps in knowledge, and gaps in actuality and reality. These have consequences of our moral disengagement and the need for ethical responsibility. Understanding and integrating the interplay between the dimensions of our interaction environments would provide a holistic framework for the design and application of interactive intelligent systems where the cultural domain drives the application process within an organisation, so that interaction between human and technology could function in a manner that allows for normal responsible behaviour.

## References


Argyle M (1991) Cooperation: the basis of sociability. Routledge, London

Carew PJ, Stapleton L (2005) Privacy, patients and healthcare workers: a critical analysis of large scale, integrated manufacturing information systems reapplied in health. Paper presented at the 16th IFAC World congress, Prague

Collste G, Duquenoy P, Carlisle G, Hedstrom K, Kimppa K, Mordini E (2006) ICT in medicine and health care: assessing social, ethical and legal issues. In: Berleur J, Numinen MI, Impagliasso J (eds) Social informatics: an information society for all? In remembrance of Rob Kling, IFIP vol 223. Springer, Boston, pp 297–308

Cooley M (1987) Architect or bee? The human price of technology. Hogarth Press, London

Cross I (2007) The evolutionary nature of musical meaning. In: Musica scientiae, (in press)

Decker M (2006) Replaceability of humans by autonomous robots? Ethical reflection within the framework of an interdisciplinary technology assessment. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Istituto Italiano per gli Studi Filosofici, Naples, 17–18 October, 2006

Dewey J (1929) Experience and nature. Dover, New York

Dreyfus H (2001) On the internet. Thinking in action. Routledge, London

Fleck J (1997) Contingent knowledge and technology development, Technol Anal Strateg Manage 9(4)

Gigliotti C (2006) Leonardo's choice: the ethics of artists working with genetic technologies. AI Soc 20:22–34
Gill SP (1995) Dialogue and tacit knowledge for knowledge transfer. Ph.D. thesis. University of Cambridge, UK
Gill KS (ed) (1996) Human-centred systems. Springer, London
Gill KS (2004a) The dance of the user in arti cial cultural space. In: International conference on the 'Culture of the Arti cial', Ascona, 23–25 April, 2004
Gill SP (2004b) Body moves and tacit knowing. In: Gorayska B, Mey JL (eds) Cognition and technology. Benjamin, Reading
Gill SP (2005) Pulse periodicity in paralinguistic coordination. In: International conference on 'Interacting Bodies', Lyons, France
Gill KS (2006) Human centredness agenda: actuality and reality. In: Invited talk at Waterford Institute of Technology, 25 October 2006
Gill SP (2007) Knowledge as embodied performance. In: Gill SP (ed) Cognition, communication and interaction: Transdisciplinary perspectives of interactive technology. Springer, London (in Press)
Gill KS (2007) Rethinking the interaction architecture. In: Gill SP (eds) Cognition, communication and interaction. Springer, London
Gill SP, Borchers JO (2004) Knowledge in co-action: social intelligence in collaborative design activity. AI Soc 17(3):322–339
Habermas J (1981) The theory of communicative action (trans. by Thomas McCarthy) 2 vols. Cambridge, Polity, 1984–1987. (Trans. from Theorie des kommunikativen Handelns, 2 vols. Frankfurt am Main, Suhrkamp, 1981)
Large EW, Reiss-Jones M (1999) The dynamics of attending: how people track time-varying events. Psychol Rev 106:119–159
Leal F (1995) Ethics is fragile, goodness is not. AI Soc 9:29–32
Leonard D, Sensipar S (1998) The role of tacit knowledge in group innovation. Calif Manage Rev 40(3):112–132
Marino D, Tamburrini G (2006) Learning automata and human responsibility. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Istituto Italiano per gli Studi Filoso ci, Naples, 17–18 October, 2006, pp 27–29
Miall DS, Dissanayake E (2003) The poetics of babytalk. Hum Nat 14:337–364
Miller D (2006) Piecemeal decision-making. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Nella Sede Dell' Instituto, Naples, 17–18 October, 2006
Moshowitz A (2007) Technology as excuse for questionable ethics. AI Soc, in this special issue (in press)
Murphy FM. A framework for eliciting user embedded knowledge requirements for information systems development. Ongoing Ph.D. thesis
Pedersen DM (1997) Psychological functions of privacy. J Environ Psychol 17(2):147–156
Rodota (2006) Adventures of the human body. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Istituto Italiano per gli Studi Filoso ci, Naples, 17–18 October, 2006, pp 46–47
Rosenberg RS (2006) The social impact of intelligent artifacts. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Istituto Italiano per gli Studi Filoso ci, Naples, 17–18 October, 2006, pp 47–50
Salvini P, Laschi C, Dario P (2007) Exploring technoethical issues in bio-robotics technology. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Nella Sede Dell'Instituto, Naples, 17–18 October, 2006, pp 54–56
Salvini P, Datteri E, Laschi C, Dario P (2007) Scienti c models and ethical issues in hybrid bionic systems research. AI Soc, in this special issue (in press)
Sha XW (2007) The poetics of performative space. In: A&Society. Springer, London
Slack WV (2001) Cybermedicine. How computing empowers doctors and patients for better health care. Jossey-Bass, CA
Smith D (2006) Are you wearing the fridge? Wearable technology and the construction of identity. In: The body interface seminar series, Middlesex University, 6 April, 2006
Torrance S (2006) The ethical status of arti cial agents – with and without consciousness. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Nella Sede Dell'Instituto, Naples, 17–18 October, 2006, pp 60–67

Trevarthen C, Aitken KJ (1994) Brain development, infant communication, and empathy disorders: intrinsic factors in child mental health. Dev Psychopathol 6:579–633
Uchiyama K (2003) The theory and practice of actuality. Institute of Business research, Daito Bunka University, Japan
Warwick K, Cerqui D (2006) Therapy versus enhancement in brain computer. In: International workshop on ethics of human interaction with robotic, bionic, and AI systems: concepts and policies, Nella Sede Dell'Instituto, Naples, 17–18 October, 2006
Wenger E, Lave J (1991) Situated learning. Legitimate peripheral participation. CUP, Cambridge
Young JZ (1951) Doubt and certainly in science. The Clarendon Press, Oxford