Ethics Inf Technol
DOI 10.1007/s10676-016-9387-z

ORIGINAL PAPER

Should we welcome robot teachers?

Amanda J. C. Sharkey[1]

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Current uses of robots in classrooms are reviewed and used to characterise four scenarios: (s1) Robot as Classroom Teacher; (s2) Robot as Companion and Peer; (s3) Robot as Care-eliciting Companion; and (s4) Telepresence Robot Teacher. The main ethical concerns associated with robot teachers are identified as: privacy; attachment, deception, and loss of human contact; and control and accountability. These are discussed in terms of the four identified scenarios. It is argued that classroom robots are likely to impact children's' privacy, especially when they masquerade as their friends and companions, when sensors are used to measure children's responses, and when records are kept. Social robots designed to appear as if they understand and care for humans necessarily involve some deception (itself a complex notion), and could increase the risk of reduced human contact. Children could form attachments to robot companions (s2 and s3), or robot teachers (s1) and this could have a deleterious effect on their social development. There are also concerns about the ability, and use of robots to control or make decisions about children's behaviour in the classroom. It is concluded that there are good reasons not to welcome fully fledged robot teachers (s1), and that robot companions (s2 and 3) should be given a cautious welcome at best. The limited circumstances in which robots could be used in the classroom to improve the human condition by offering otherwise unavailable educational experiences are discussed.

✉ Amanda J. C. Sharkey
a.sharkey@dcs.shef.ac.uk; a.sharkey@shef.ac.uk

[1] Department of Computer Science, University of Sheffield, Regent Court, Portobello Rd, Sheffield S1 4DP, UK

Keywords Robot teacher · Robot companion · Robot ethics · Attachment · Deception · Privacy · Classroom

> One looks back with appreciation to the brilliant teachers, but with gratitude to those who touched our human feelings. The curriculum is so much necessary raw material, but warmth is the vital element for the growing plant and for the soul of the child. (Carl Jung 1953)

Introduction

Many children find the idea of robots exciting. Imagine a school visit to a museum, where a small friendly humanoid robot explains to a group of children why they should eat enough vegetables. The children are likely to pay attention and to enjoy the encounter. They might even remember the lesson more than they would if it had been delivered at school by their regular teacher. There seems little reason to object to such a presentation. But if the children were to arrive at school the next morning to find a robot in the place of their familiar teacher, they (and their parents) might not be so happy.

People are worried about the use of robots in schools. In 2012, a European survey of public attitudes to robots of over 27,000 people found that 34 % thought robots should be banned from the field of education (Eurobarometer 382 2012). 60 % thought that robots should be banned from the care of children, the elderly or the disabled. Only 3 % of those surveyed thought that robots should be used in education. Are these negative views justified? In this article, we will look at current and near future uses of robots in the

Published online: 10 February 2016

classroom and discuss the extent to which there are good reasons to be concerned about their use.

Robotics has progressed to a point where there is a real possibility of robots taking on social roles in our lives, and it has become crucial to look at the ethical issues raised by such developments. We need to think about where robots can and should be used, and where they would be best avoided, before we travel too far along a path towards complete automation. The field of robot ethics is currently undergoing quite a rapid development (Lin et al. 2012), and there have been a number of ethical assessments of the use of robots in society. These range from considerations of the advantages and risks posed by robot nannies (Sharkey and Sharkey 2010), to using robots to care for older people (Sparrow and Sparrow 2006; Sharkey and Sharkey 2012), or even for the provision of sexual services (Levy 2007). In this paper, we focus upon the ethical issues raised by the idea of robots teaching in the classroom. In order to provide a realistic grounding for this discussion, we begin with a review of the social robots that are currently being used in classrooms. On the basis of this review, four representative scenarios will be identified. These will be used as the basis for an ensuing discussion of the ethical concerns that they raise.

Current robots in the classroom

Robots as objects to be manipulated and operated by students have become common place in schools. For quite some time they have been used as intermediary tools to explain concepts in mathematics and science, and as a means of involving students in technology by building and programming robots and working in teams (for reviews see Benitti 2012; Mubin et al. 2013). Our focus here is instead on the idea of using 'social' robots to act as teachers, or as classroom companions. When a robot is acting as a teacher, or as a companion, the children are encountering an apparently social being, and are not involved in programming, or building it.

There are already some examples of social robots being used in classrooms. One example is Saya, a humanoid robot deployed in classrooms in Japan to deliver material about the principles of leverage, and an introduction to robotics. Saya is a remote controlled humanoid robot with a female appearance. She (or it) consists of a movable head with the ability to make emotional facial expressions, attached to a manikin body (Hashimoto et al. 2011). An operator in a control room can hear and observe the students in the classroom by means of a video camera and the robot's CCD camera. When the robot is operated in 'interactive' mode, it can articulate brief sentences and accompany them with an appropriate facial expression, for instance telling the class to "Be quiet!" whilst displaying an angry facial expression. Despite this, questionnaire responses from elementary school pupils indicated that the class was enjoyable.

The Saya robot was presented in the role of a teacher. More often, classroom robots are presented in the role of a companion or peer. For instance, 'Rubi', a low cost 'sociable' robot, was used to explore whether a robot could improve toddlers' vocabulary skills (Movellan et al. 2009). The robot was immersed in an Early Childhood Education Centre for 2 weeks. It operated autonomously during this period, and could sing and dance; play a physical game of taking and giving back objects using its physical actuators; and play Flash-based educational games targeting vocabulary development. It switched between games depending on an 'interest estimator' which took into account the number of faces detected and the number of touches received in the past minute. The researchers reported evidence of a 27 % improvement in 18–24 month toddlers' knowledge of the target words taught by the robot as compared to a matched set of control words that were not taught. They concluded that 'sociable robots may be an effective and low cost technology to enrich Early Childhood Education environments'.

Kanda et al. (2004) describe an 18 day field trial at a Japanese elementary school in which two English-speaking 'Robovie' robots operated autonomously to interact with first and sixth grade pupils. The robots could identify the children by means of the wireless tags they wore. The robots spoke English with children that approached it, and had a vocabulary of around 300 sentences for speaking, and 50 words for recognition. A picture-word matching test was administered to the children before the study, after 1 week and after 2 weeks, and the frequency of interactions between the children and the robots was recorded. Improvements in English skills as measured by the picture-word matching test were found in those children who interacted with the robot more often. Kanda et al. stressed the need to investigate the development of longer term relationships with robots, as opposed to the brief initial encounters that are often studied. The robots' ability to use the children's names was found to encourage interaction. Evidence of an improvement in English scores was found for those students who continued to interact with the robot over the 2 week period, and who could be said to have formed some kind of a relationship with it.

In a subsequent study, Kanda et al. (2007) developed a classroom robot installation designed to encourage children to continue to interact with a robot for a longer period. The field trial was performed over an 8 week period in an elementary school in Japan, placing a Robovie robot in a class with 37 students aged 10–11 years. The children were given the opportunity to interact with the robot during the

lunch time break. As well as being able to identify the children by means of RFID tags, the robot could keep track of how often individual children interacted with it. It was programmed to exhibit 'pseudo development': adding more behaviours to its repertoire over time. In addition, the robot informed the children that it would tell them a secret if they spent time with it: the personal information it divulged varied depending on how long the child had spent with it. For instance, it would tell children that it liked the teacher, or what its favourite baseball team was.

As well as studies in which classroom robots are presented as companions, some researchers have explored young children's interactions with robots designed to elicit care-giving behaviour from them. Tanaka et al. (2007) placed a robot in a classroom of 18–24 month old toddlers for 45 sessions each lasting approximately 50 min over a period of 5 months. The aim was not to get the robot to teach the children, but to look at the social interactions between the children and the robot. The QRIO robot received some input from a human operator specifying a walking direction, head direction, and six different behavioural categories (dance, sit down, stand up, lie down, hand gesture and giggle). An automatic giggle reaction when its head was patted was set up for the robot, to ensure a contingent response. The researchers claim to have found evidence of 'long term bonding' between the robot and the children in their study. The children continued to interact with the robot over time, and exhibited a variety of social and care taking behaviours towards the robot. They touched the robot more often than a static toy robot or a teddy bear, and the researchers claimed that they came to treat it as a peer rather than as a toy.

Tanaka and Matsuzoe (2012) introduced a 'care-receiving' robot into an English language school for Japanese children. The robot was smaller than the children, made mistakes, and seemed to need their help. 17 children aged between 3 and 6 years were involved in the study with the aim of seeing whether the children would learn English verbs if they 'taught' them to the robot. They identified a set of 4 previously unknown English verbs for each child. Two of the verbs were taught to the children by the experimenter (who asked them to match up the word and the appropriate gesture). For the other two verbs, the experimenter showed the child how to teach the robot to match the word and the gesture, and then encouraged the children to teach the robot in the same way. The verbs the children taught to the robot were remembered better than the words the experimenter taught them directly. The authors conclude that these preliminary results suggest that getting children to teach 'care-receiving' robots can have some educational benefits. Hood et al. (2015) also report research in which children taught a robot. Children aged 6–8 years taught a Nao robot to form handwritten letters. The robot was programmed to make the same errors as typically made by children, and to gradually improve its performance based on the example letters that the children formed on a tablet computer. The children were keen to teach the robot, although it is not clear whether or not teaching the robot led to improvements in their own handwriting.

Telepresence robots represent another form of robots to be found in classrooms. They have been used to enable telepresence communication between pupils and remote teachers and also between pupils in different classrooms. Tanaka et al. (2013) report a study in which a child-operated telepresence robot was used to link remote classrooms of children aged 6–8 years old in Australia and Japan. Their preliminary results suggested that when individual children controlled a remote robot to interact with a distant English teacher, they were more engaged than when they interacted with the teacher via a Skype screen. Similarly, when Australian children remotely controlled a robot in a Japanese classroom, the Japanese students were keen to interact, and to try using English phrases to address the operator.

There has been considerable interest in South Korea in using robots for English language teaching (Han 2012). Han et al. (2005) report studies of the educational effectiveness of the IROBI robot, a so called home educational robot. They found better learning of English from the robot compared to other media (books with an audio tape, and a computer program). The EngKey robot has been deployed in South Korean classrooms to teach students English via telepresence. The EngKey has a dumpy egg shaped appearance and was designed to seem friendly and accessible. It has been used to enable remote teachers in the Phillipines to teach English to South Korean students, and found to improve students' performance when deployed in field tests in 29 elementary schools in South Korea (Yun et al. 2011). The Robosem robot developed by Yujin Robotics has also been used as a telepresence robot for remote language teaching in Korea (Park et al. 2011).

Telepresence robots differ from the autonomous robots used in some of the studies described here, in being overtly controlled by a human operator. They usually have an anthropomorphic appearance, and some, like the EngKey and Robosem robots, can operate in either telepresence or autonomous mode (a useful classroom feature when the remote connection breaks down). By contrast, the Robovie robots investigated by Kanda and colleagues are designed to operate autonomously, without human input. Then there are robots such as Saya that are presented as if they were autonomous, but are remotely controlled in a Wizard of Oz set up. Other robots are operated semi-autonomously, with some human supervision. For example, the QRIO robot researched by Tanaka and colleagues exhibited some

autonomous behaviour but could also be directed by a human operator so that it responded more appropriately to what was going on in the classroom.

It is evident from this review that the idea of robot teachers in the classroom is not just the stuff of science fiction. At the same time, it is apparent that the current abilities of robot teachers to operate autonomously are still quite limited, and often aided by covert or even overt human intervention or remote control. The underlying motivation of several of these studies is often more one of exploring whether the robot would be accepted in the classroom than of demonstrating its effectiveness at teaching. Some of the studies, such as those by Kanda et al. (2004, 2007) and Tanaka et al. (2007) are designed to investigate children's relationships to robots over a longer time period than many human-robot interaction studies cover. Others explore some of the factors that affect children's interest in the robots, such as the ability of the robot to call the children by name (Kanda et al. 2004), or to give them privileged (secret) information (Kanda et al. 2007).

The studies do show that children *can* learn from robots, particularly in the application area of robot language teaching. Kanda et al. (2004, 2007) found improvements in English scores. Tanaka and Matsuzoe (2012) found better learning of the words that children taught to robots as compared to the words the experimenter taught them. Movellan et al. (2009) report an improvement in vocabulary scores, and Yun et al. (2011) report an improvement in student performance as a consequence of the robot's telepresence operation. Nonetheless there is a need for more careful experimental design here. It is important to compare the robot's teaching efficacy to other teaching methods; especially so given the greater cost usually associated with robotics. Comparisons between the effectiveness of a human teacher and a robot teacher are rarely undertaken (for an exception see Rostanti 2015). Comparisons between the effects of language teaching by means of a telepresence robot and a Skype interface (Tanaka et al. 2013), and an educational robot and other media (Han et al. 2005) represent steps towards more convincing assessments.

There is also scope for more detailed investigations of the extent that children learn and retain the information delivered by a robot, and of the factors that determine the robot's teaching effectiveness. The appearance of the robot and its ability to interact with and respond to its audience are prime candidates for such factors. Komatsubara et al. (2014) carried out a field trial with a social robot that quizzed children about science lessons they had received from a human teacher, and told them the correct answer together with a simple explanation. However they found no evidence that the robot increased their knowledge and it was suggested that the children may have been bored by its delivery, especially when it continued to give an explanation when the children had already understood.

An interesting question that needs to be explored is the extent to which children trust and believe in robots that are presented in a teaching role, and the factors that affect that trust. Some recent work on selective trust has begun to explore the factors that influence a child's beliefs in what they are told (Sabbagh and Shafman 2009). As Koenig and Sabbagh (2013) point out, 'children do not blindly trust the words of others', but exhibit selective learning; making decisions about who to believe about what. A robot that is unable to answer children's questions when they stray beyond the featured topic would probably be viewed quite sceptically by the children it is 'teaching'. It is also likely that the appearance and behaviour of a robot will affect the extent to which the information it provides will be believed, with different results from robots with different appearances and behaviours. It is also possible that trust and belief in a robot will depend on the topic being considered. When Gaudiello et al. (submitted) considered people's trust in an iCub robot and the robot's influence on their decision making, they found it had more influence when its answers related to functional and technical questions (e.g. the weight of objects) and less when they related to social questions (which items were more suitable for different social contexts). This implies that people might be more willing to believe and trust information provided by a robot when it concerns factual and functional topics, than when it deals with emotional and social issues. The phenomena of automation bias (Carr 2015) and algorithm aversion (Dietvorst et al. 2015), are also relevant here, although their relationship to robots in the classroom have not yet been explored. There is a need for further research here: if robots are to be placed in classrooms, it is important that they are given an appropriate level of trust and acceptance.

Four scenarios for robots in the classroom

The ensuing discussion of the ethical issues raised by robot teachers will be made more specific by basing it on a set of four representative scenarios. These scenarios are identified on the basis of the classroom contexts exemplified in the studies described above, in a review that presents a picture of the current state of the art in 2015. First, the Saya robot was presented in the role of an authoritative classroom teacher (even though it was actually remotely controlled). This leads to the identification of Scenario 1, Robot as Classroom Teacher. In the investigations reported by Movellan et al. (2009), Tanaka et al. (2007), Kanda et al. (2004), the robots were presented to the children as

companions and peers rather than as a teacher. On the basis of these studies, we identify Scenario 2, Robot as Companion and Peer. Some researchers (Tanaka et al. 2007; Tanaka and Matsuzoe 2012; Hood et al. 2015) used companion robots designed to elicit care-giving from children: these examples form the basis for Scenario 3, Robot as Care-eliciting Companion. And finally there are the Telepresence robots, used to enable a remote teacher to teach the class, which lead to the identification of Scenario 4, Telepresence Robot Teacher.

Scenario 1: Robot as Classroom Teacher.
Scenario 2: Robot as Companion and Peer.
Scenario 3: Robot as Care-eliciting Companion.
Scenario 4: Telepresence Robot Teacher.

As well as being based on the reviewed studies of classroom robots, it is claimed that these scenarios represent an interesting range of roles for robots in the classroom. They vary in the extent to which the robot replaces or supplements the human teacher. The most extreme version of a robot teacher is represented by the Classroom teacher in Scenario 1, since it involves the robot replacing the human teacher for at least a limited period. A Classroom teacher robot would need to act as a figure of authority and as an explicit source of knowledge. By contrast, the Companion robot and the Care-eliciting companion robot scenarios do not involve replacing a human teacher, and could depend on a human teacher to be present and in charge of the classroom. Neither would require the presentation of the robot as an authoritative figure, and both have a goal of implicit rather than explicit teaching. The Telepresence robot by contrast could be used to replace (or to supplement) physically present human teachers with a remote, albeit human, educator. These four scenarios are not the only ones possible, and different situations could arise in future studies. However it is claimed that identifying and discussing these provides a necessary and useful first step towards an ethical consideration of robot teachers, and enables a consideration of whether some of these scenarios represent better goals than others.

Ethical concerns about robot teachers

In order to determine the ethical issues that are most relevant to the idea of robot teachers, as exemplified by the four scenarios, we begin by examining the ethical concerns previously raised elsewhere in discussions about social robots in related situations and contexts.

A number of questions about the impacts of social robots on the privacy of individuals have been previously raised (see Sharkey and Sharkey 2010, 2012). Social robots can affect the privacy of individuals by collecting personal identifying information about them that can be accessed by other people. The privacy of individuals would be intruded upon if a social robot was used to enable direct surveillance. For instance, information picked up by the robot's sensors that enabled the identification of the person being monitored could be directly transmitted to human monitors, even though that person might consider themselves to be alone and unobserved. Alternatively (or additionally) such personal information could be stored on the robot, and subsequently accessed by others. An insightful discussion about the impact of robots on the privacy of individuals can be found in Calo (2012). As he points out, robots in social spaces highlight questions about increased direct surveillance, since they are 'equipped with the ability to sense, process and record the world around them'. As mobile devices, robots may be allowed increased access to historically protected spaces such as the home. Also, by dint of what Calo terms their 'social meaning', and their apparent social nature, robots may extract confidences from people that computers or other machines would not. There are particular reasons to be worried about the privacy implications of robots in the classroom, and this forms the starting point for the ethical consideration that follows.

As well as privacy, there is another set of interrelated concerns that arise as consequence of the presentation of robots as social entities. If a robot is built to resemble a human being, or at least a being with emotions, those who encounter it may expect it to be able to care for and look after people. However, this appearance is, in some respects, deceptive (although the issue of deception is a complex one, as discussed in "Attachment, deception and loss of human contact" section). Questions have been asked about the ability of robots to provide meaningful care for older people (Coeckelbergh 2010; van Wynsberghe 2013), and about the impact of robot care on the dignity of older people (Sharkey 2014). Such questions are relevant to robot teachers, since one aspect of what is required of a good teacher is that they should provide care for the children in their charge.

In related work, Sharkey and Sharkey (2010) identified a number of ethical concerns associated with the idea of Robot Nannies. As well as misgivings about their effects on children's privacy, several of these concerns were related to questions about attachment, or lack of attachment, between children and robots, and about the deception this could involve. The idea of robot nannies differs in several respects from that of robot teachers. For a start, a 'nanny' robot is more likely to be used at home than in the classroom. A robot nanny is also more likely to be used with very young children and babies, and to come with a strong risk of psychological harm if used for any extended periods of time. Nonetheless, many of these concerns are still relevant to the idea of robot teachers in the classroom.

Others have also highlighted concerns about the deception that may be engendered by robots, particularly in discussions of robot companions and robot pets. Sparrow (2002) takes exception to the deception and self-deception that he claims robot companions and robot pets rely on. Wallach and Allen (2009) also suggest that the techniques used to enable robots to detect and respond to human social cues are 'arguably forms of deception'. It therefore seems important to consider the extent to which robot teachers or classroom companions involve some form of deception, and whether this could lead to negative consequences.

Another common concern is the loss of human contact that could result from the deployment of social robots in some circumstances. Sparrow and Sparrow (2006) were suspicious about the reduction in human contact that would result from the introduction of any robots, social or not, into the care of the elderly. As well as loss of human contact, Sharkey and Sharkey (2012) were also concerned about the reduction in human contact that could result from the use of robots to care for the elderly, or from their use as robot nannies (Sharkey and Sharkey 2010).

Attachment, deception and loss of human contact are all pertinent to the idea of robot teachers. The concepts cannot be easily disentangled from each other. For instance, the deceptive appearance of robots as real social entities could lead people to form attachments to them, or to imagine that they were capable of or worthy of attachment. This could in turn increase the loss of human contact that could result from the introduction of robots in the classroom. Because they are so interrelated, the ethical issues relating to attachment, deception and loss of human contact are considered together under one heading.

Other pressing ethical concerns that have been raised in papers on robot ethics, and that seem particularly relevant to the use of robots in the classroom, are those that pertain to control and accountability. Placing robots in charge of human beings, whether they are vulnerable elderly people (Sharkey and Sharkey 2012), or young children (Sharkey and Sharkey 2010), gives rise to questions about control and accountability. To what extent should robots be trusted to make the right decisions about what humans should do? To what extent can they be held accountable for such decisions? It seems important to consider these questions with respect to robots in the classroom.

Following this analysis, in this ethical assessment of the idea of robot teachers, we will concentrate on discussions of (1) Privacy; (2) Attachment, deception and loss of human contact and (3) Control and accountability. These topics do not exhaust the list of possible issues for consideration—there are others such as safety and liability, which are also relevant. However safety and liability issues are common to all robotic applications that involve contact with humans, and we suggest that they are best discussed in the context of robotics as a whole. The three headings used here have been chosen because they seem the most relevant and the most in need of reinterpretation and articulation in terms of robot teachers.

Privacy

The more technology is used in the classroom, the more issues about privacy of information come to the fore. A robot's ability to interact with children is enabled by sensors. If those sensors are used to enable a reactive response, without storing information, there seems little reason to worry. For instance, a robot might use its sensors to detect whether or not a child or a group of children was standing in front of it in order to trigger its presentation. But if the information detected by the robot is recorded, or categorised and recorded, this gives rise to concern about what information should be stored, and who is permitted access to it. In addition, even if the information is not stored, the use of sensors and associated algorithms that make it possible to detect children's emotional state could be viewed by some as a step too far.

As is apparent from the studies described earlier, robots in the classroom can be enabled to recognise individuals. This can be accomplished by means of RFID tags worn by the children enabling the robot to call them by their names. Alternatively, face recognition algorithms could be used to recognise individual children. Recognising and naming a child does not necessarily mean that further information about that child will be stored, but it raises questions about record keeping. Indeed, Kanda et al. (2007) describe how the robot they used kept a record of which children had interacted with it, and even of friendship groups amongst the children. This strikes a disturbing note. Is it too far-fetched to imagine that, in the future, robots might be used to categorise and monitor children's behaviours; keeping a record of disruptive behaviour, or alerting the teacher?

In the present post-Snowden climate, there is uneasiness about technologically based invasions of privacy. Of course, when experimental research studies are conducted in a classroom there are established protocols to follow about the storage of personal information. However, if robots are to be really used in the classroom, the personal information they store will not be deleted at the end of the study in the same way. There are many questions to be considered here, including the extent to which such information should be used as the basis for educational decisions made about the child. The storage of personal information is covered in the UK by legislation such as the Data Protection Act, but the mobility and connectedness of robots provide new challenges. In 2015, concerns about the collection of 'big data' in schools were raised by President Obama in USA where there are plans to introduce the

Student Digital Privacy Act to curtail the use of information about students collected by schools in order to provide personalised educational services and to limit targeted advertising and selling of the data. In UK, although the Data Protection Act provides some protection of personal data, the full implications of an increasing ability to sense and store enormous amounts of personal data have not yet been thoroughly addressed.

Sensors in the classroom also give rise to the possibility of a more intimate form of privacy invasion. Physiological measures, and emotional facial expression recognition, offer the potential to detect and possibly record information about the emotional state of children interacting with a robot. For instance, a biometric bracelet named the Q-sensor was developed by Affectiva (an MIT media lab spin off company) to measure Galvanic Skin Response (GSR) and the emotional arousal of the wearer (http://affect.media.mit.edu/projectpages/iCalm/iCalm-2-Q.html). It was suggested that it could be used as an 'engagement pedometer', indicating when students are engaged and when they are bored and disinterested. Affectiva has subsequently diverted its attentions to the development of other software. Nonetheless, both physiological measures, and emotional expression recognition have the potential of being used in the classroom to track students' engagement.

Of course, a robot that is able to detect the level of engagement of its audience may deliver a better performance. Mutlu and Szafir (2012) programmed a Wakamaru humanoid robot to monitor the engagement of its users and to adjust its behaviour to increase that engagement. They monitored real time student attention using neural signals captured using a wireless EEG headset as the robot told a story to individual participants. The robot was able to nod its head, and engage in eye contact during the story. In addition, it could display 'immediacy cues' by increasing its volume and making arm gestures. In three different conditions it (1) displayed these immediacy cues at random intervals, or (2) displayed them adaptively when the EEG indicated a drop in the participant's level of engagement or (3) did not change its volume or use gestures. Performance on a memory test for a story told by the robot indicated that participants' memory for the story was significantly better when the robot responded adaptively to detected decreases in engagement.

Even though a robot might well increase the engagement of its audience through the use of sensors, there are still reasons to be concerned about their use. One problem is that high levels of arousal might have nothing to do with the material or delivery but could be caused by other events in the classroom. Higher levels of arousal could also be created by exciting behaviours on the part of the robot that do not result in better learning or understanding of the material being communicated, and would push the educational system towards a form of 'edutainment' in which any difficult and potentially boring topics were avoided.

In addition, the use of emotional detectors and sensors can be viewed as an invasion of privacy. Although a human teacher may be able to recognise the emotions and feelings of their pupils to some extent, this is not the same as the kind of recognition that might become possible if the pupils had to wear sensors on their body that could transmit information about their present emotional state. Teachers and other adults sometimes complain about the eye rolling behaviour of children, but what if they could further legitimise this complaint by referring to data on the children's emotional response tracked by a digital device?

Does the relevance of these privacy issues differ for the four scenarios for robots in the classroom? Most apply equally to all four scenarios, because personal data might be stored or used in any of them. There are additional concerns about the privacy of information stored and conveyed by means of a telepresence robot (Scenario 4) because of the potential to cross national boundaries (e.g. South Korea to the Philippines), complicating the application of national legislation such as the Data Protection Act. There are also concerns that apply particularly to the scenario in which the robot is presented as a companion or peer (Scenarios 2 and 3). Presenting a robot to children as their 'friend' could encourage them to share information, and even confide secrets, in ways that could result in a violation of their privacy. This issue is tied up with the questions about deception and attachment that will be explicated further in the following section.

Attachment, deception and loss of human contact

There is a growing knowledge of the factors that contribute to the illusion that a robot is able to relate to humans. A robot's sensors, as discussed earlier, can allow it to respond to and interact with humans in ways that foster the semblance of understanding. For example, a robot that is able to detect a person's emotional facial expression and respond with a matching one of its own, or an appropriate verbal comment, will seem responsive. A robot that can look into the eyes of the person talking to it is more likely to seem sentient. Likewise a robot that can detect when a person is paying attention to it, or what they are paying attention to may be seen as one that understands what is going on. A robot's ability to respond contingently to humans can be enabled by its sensors, and makes for a more convincing robot (Yamaoka et al. 2007). The appearance and behaviour of a robot also plays an important role. The illusion of understanding can be more convincing if the robot's appearance avoids the uncanny valley, and it behaves like a human whilst not looking too

much like one. This is probably because we are so skilled at rapidly evaluating human behaviour and monitoring it for any signs of abnormality. A good match between a robot's voice and its appearance helps (Meah and Moore 2014), as does its ability to respond with emotional expressions that are appropriate to the surrounding context (Zhang and Sharkey 2012).

The creation and development of robots that are able to respond appropriately to humans can certainly have the effect of making them easier to interact with and more fun to have around. At the same time, the argument has been made by some that such development is inherently deceptive. Sparrow and Sparrow (2006) writing about the use of robots for elder care, argue that 'to intend to deceive others, even for their own subjective benefit is unethical, especially when the result of the deception will actually constitute a harm to the person being deceived'. Wallach and Allen (2009) also consider the techniques being used to give robots the ability to detect social cues and respond with social gestures, and conclude that 'from a puritanical perspective, all such techniques are arguably forms of deception'. Sharkey and Sharkey (2006) pointed out that much research in Artificial Intelligence, from robotics to natural language interfaces, depends on creating illusions and deceiving people.

Of course, the terms 'deception' and 'deceptive' in the context of robotics do not necessarily imply any evil intent on the part of those keen to create the illusion of animacy. The harm that could be created by a robot that gives the illusion of sentience and understanding is not going to be immediately obvious, and researchers attempting to create robots able to respond to humans in a social manner may not have even considered that their endeavours could lead to any kind of damage. Since the time of automata makers, or even earlier (Sharkey and Sharkey 2006), inventors have enjoyed creating apparently life-like machines. In addition, those who view and interact with such machines can be seen as contributing to their own deception, since the human tendency to enjoy being anthropomorphic is well known (Epley et al. 2007, 2008; Sharkey and Sharkey 2011).

Despite these points, there is a strong risk that robots which create the illusion of a greater competency than they actually possess could engender some harm to the person or persons being deceived. Friedman and Kahn (1992) point out some of the risks of imagining that machines are capable of more than they actually are. A robot that is too good at emulating the behaviour of a human could lead people to expect too much of it, and to use it for educational purposes for which it is not well enough suited. It could for instance encourage the view that it could be placed in a position of authority such as that of a classroom teacher. The problems associated with such a view will be discussed in "Control and accountability" section.

The idea of deception and the creation of a convincing illusion give rise to several important issues relating to the emotional attachments that might, or might not develop between children and classroom robots. There are particular risks associated with the convincing presentation of a classroom robot as a companion or peer. If a classroom robot is presented as a friendly companion (Scenarios 2 and 3), the children might imagine that the robot cares about them. They might feel anxious or sad when the robot is absent, or choose to spend time with the robot in preference to their less predictable human peers. Instead of learning how to cope with the natural give and take involved in playing with fellow students they might get used to being able to tell their robot companion what to do. In other words, some of their learning about social skills could be impeded.

Children do sometimes try to abuse robots (Bartneck and Hu 2008; Brščić et al. 2015). A child could be unpleasant and cruel to a robot and it would not notice. The child might as a result learn that bad behaviour in friendships does not have any consequences. Tanaka and Kimura (2009) mention the expectation that 'people who treat non-living objects with respect naturally act in the same way towards living things too'. However, the impact that human-robot relationship have on subsequent relationships with other human beings is unknown. Supposing that a child were to treat the robot badly, what impact would this have on their behaviour towards other children?

There is also the possibility that a child's trust in relationships could be weakened if they thought the robot was their friend, but came to realise that the robot was just a programmed entity, and as likely to form a friendship with the next child as with them. Similarly, the pseudo relationship formed with the robot could affect the child's views and understanding of how relationships work. Thinking you have a relationship with a robot could be like imagining you have a relationship with a psychopath: any kind and empathetic feelings that you have for the robot are definitely not reciprocated.

It could be argued that an attachment formed to a robot is no different to the attachments that children feel for their favourite cuddly toy. But there are important differences. A cuddly toy does not move, and any attachment that the child feels for it is based on their imagination. A social robot is also not a living entity. However unlike the toy, it can be programmed to move and behave as if it were alive. As a result it is can be more compelling to interact with, and children may be less clear about whether or not it is a living being, and about whether or not it can reciprocate in a relationship. There is good reason to believe that a robot

(like the Care-eliciting robot in Scenario 3) that seems to be vulnerable and in need of care is particularly hard to resist. There have been various computer games that have exploited children's caring natures: think about the Tamagotchi craze at the turn of the century, when children spent hours looking after a digital pet on a screen.

It might also be claimed that attachments to a robot are no more based on deception than a child's attachment to the family pet. But again, there are important differences. First, the family pet is a living creature, and something with which the child can genuinely form a relationship. Even though it is often argued that we should not be anthropomorphic and imagine that animals have human-like feelings for us, the family pet will know the child, and will be directly affected by its actions. The robot, on the other hand, will only be able to simulate any affective response to the child. Some might suggest that robots will eventually be able to feel real emotions, but there is little evidence that this will happen any time soon.

As well as concerns about robots presented as children's companions (Scenarios 2 and 3) there are also questions to be asked about the attachments children would form, or fail to form with Telepresence Robot Teachers (Scenario 4). Any relationship with the distant teacher could be complicated by the children's views of and relationship with the Telepresence robot itself. The extent to which a human teacher's relationship to the children in the classroom would be affected by not being physically present is unknown, and in need of further investigation.

Problems seem likely to result from placing a robot in the role of a Classroom Teacher (Scenario 1). Children do form attachments to their human teachers, and can be attentive to their direction: learning more from them than just the explicit educational material they deliver. The quotation from Carl Jung at the beginning of this paper is apposite here. Teachers are most effective when they function as an attachment figure. Bergin and Bergin (2009) summarise research on attachment style relationships with teachers where attachment is defined as a deep and enduring affectionate bond that connects one person to another across time and space (Ainsworth 1973; Bowlby 1969). It is a bond that is first formed with a baby and child's primary caregiver, and affects their relationship to the world. A securely attached child feels at liberty to explore the world, secure in the knowledge that they have a caregiver they can rely on. Although the main attachment bonds will be to the child's primary caregiver, teachers can also function as attachment figures. Bergin and Bergin (2009) claim that attachment is relevant to the classroom in two respects. First, an attachment bond between child and teacher can help in the classroom by encouraging the child to feel secure, and able to explore their environment. Second, an attachment to a teacher can help to socialise children, as they adopt the adult's behaviour and values and are encouraged to interact harmoniously with other children. An attachment bond is more likely to be formed with a teacher who is sensitive to the child's emotions and needs.

If robots were to be increasingly deployed as Classroom Teachers in the future, there is a risk that children would *not* view them as attachment figures, and so would lose that emotional security. By contrast, if they were to perceive the robot as an attachment figure, this would open the possibility of the children adopting the robot's apparent values, and as in the case of the robot companion, basing their social skills and world outlook on the behaviour and apparent attitudes of a machine rather than on a living, breathing, empathising human.

Control and accountability

The notion of robot teachers highlights concerns about robots being in charge of human beings. The idea of robots being in a position to exert control over humans, even (or especially) when those humans are children, should be controversial. However it is hard to imagine how a robot could function as a teacher (Scenario 1) without being able to exert its authority over the children in the classroom. Surely it would need to be able to recognise, and prevent, disruptive behaviour? It would also need to be able to recognise and reward positive behaviour and successful learning, and find ways of reducing or eliminating negative behaviour and poor learning outcomes.

Many people might be concerned by the idea of giving robots the power to restrict the activities of humans. At the same time, others might like to think that robots would be fairer than humans. Those who had uncomfortable relationships with teachers in their childhood could argue that a robot would do better: it would not be prejudiced, vindictive or angry. A similar argument has been made in other contexts, from care-giving to the battlefield. Borenstein and Pearson (2013), in a discussion of robot caregivers, suggest that robots could be preferable to humans in some respects, because 'robots are unlikely to suffer from certain kinds of human failings' (Borenstein and Pearson 2013) since they lack empathy and are therefore not susceptible to the 'dark side of empathy': namely indifference and even sadism (Darwall 1998). In a military context, Arkin (2009) has argued that robot soldiers could be more ethical than human soldiers because they would not get angry or want to take revenge. The suggestion that a robot would be fairer and less prejudiced than humans in the classroom is related to Arkin's claim that robots can be more ethical than humans.

Arkin (2009) proposed the idea of an *ethical governor* for robot soldiers, which would evaluate possible actions

against a set of constraints such that unacceptable levels of collateral damage would be avoided, and only morally permissible actions selected. Winfield et al. (2014) also discuss the possibility of an ethical robot that evaluates possible actions against a set of constraints before selecting one. They describe an example in which a robot risks its own safety in order to preserve the safety of another robot (representing the idea of a robot protecting the safety of a human). Could a robot teacher be programmed in a similar way to make ethical decisions in a classroom; decisions for instance about when to praise or castigate children for their behaviour?

One problem with this idea is that making a good decision about what to do in the classroom depends on having the ability to discriminate between different kinds of behaviour, and to understand the intentions that underlie them. Recognising which children are misbehaving and disturbing the classroom requires a detailed understanding of the intentions behind a child's actions. A quiet child could be studying, or sullenly refusing to participate. A vociferous child might be actively contributing to the class discussion, or interfering with it. The problem is further compounded by the rapidity with which pupils can change states; the previously studying child can switch to being a disruptive ring leader. For a robot to exert effective (and fair) control over children's behaviour in the classroom, it would also need a reasonable idea of their probable next actions, and to have strategies for encouraging good behaviour and discouraging bad behaviour. These are abilities that humans have, and that the best teachers can exploit effectively.

Could a robot have these abilities? It seems unlikely in the near future. Christof Heyns, the UN Special Rapporteur on extrajudicial, summary or arbitrary executions argued against the use of autonomous robots to make lethal decisions on the battlefield on the basis that robots lack 'human judgement, common sense, appreciation of the larger picture, understanding of the intentions behind people's actions, and understanding of values and anticipation of the direction in which events are unfolding' (2013, A/HRC/23/47). Clearly robot teachers would not be required to make lethal decisions, but their actions would still impact the lives of children, and they also lack the abilities listed by Heyns.

Could robots develop these abilities in the distant future? There are good reasons to think they will not. It has been argued that understanding good and bad behaviour depends on a sense of morality, which itself has a biological basis. Churchland (2011) argued that morality depends on the biology of interlocking brain processes: caring (rooted in attachment to kin and kith and care for their well being); recognition of other's psychological states (rooted in the benefits of predicting the behaviour of others); problem solving in a social context; and learning social practices. She argues that the basis for caring about others lies in the neurochemistry of attachment and bonding in mammals. Neuropeptics, oxytocin and arginine vasopressin underlie the extension of self-maintenance and avoidance of pain in mammals to their immediate kin. Humans and other mammals feel anxious and awful both when their own well-being is threatened, and also when the well-being of their loved ones is threatened. They feel pleasure when their infants are safe, and when they are in the company of others. Churchland (2011) extends her argument about morality originating in the biology of the brain to explain the development of more complex social relationships.

This argument implies that robots do not have the necessary biological nature required for a sense of morality. Without this, how could they make fair decisions about good or bad behaviour in the classroom? The robot teacher could 'decide' by means of pre-programmed rules, but their effectiveness would depend on the programmer having anticipated the situations likely to arise and the appropriate response to them. The variety of situations and social encounters that could arise in a classroom makes this unlikely.

Although it may be possible to create the illusion of understanding and empathetic robots, it remains the case as Wallach and Allen (2009) acknowledge, that 'present-day technology is far from having the kinds of intelligence and intentions people demand from human moral agents' (p. 45). Roboticists have begun to consider the relevance of artificial empathy to robotics (e.g. Damiano et al. 2014), but this research is at an early stage. In the meantime, robots' lack of understanding of children's behaviour provides a major stumbling block for suggestions that robots will be able to replace human teachers any time soon.

As well as deficits in moral understanding, robots are also not necessarily fair and unbiased. Because robots are developed and programmed by humans, they can exhibit *technological bias*. Forms of technological bias were already being discussed nearly two decades ago (Friedman and Nissenbaum 1996). The idea was illustrated in 2009 by reports showing that Hewlett-Packard webcams' face tracking algorithms worked only with white faces, and not with black faces (the problem was subsequently fixed). Ensuring that a robot treats all children equally requires the developers and programmers of the robot to be aware of possible inequalities that could result from the robot's behaviour or sensors. Hewlett-Packard is unlikely to have intended their face tracking algorithm to be racist; the developers had just failed to notice that the algorithms they were using did not perform well with black and darker skin. It is possible to imagine other forms of bias that a robot might show, if they were not anticipated by its

programmers and developers. For instance, any speech recognition systems they use are likely perform better for children without strong regional accents, or dialects.

As well as questions about a robot's ability to make appropriate decisions, robot teachers would also give rise to legal issues about accountability. Teachers need to be able to reward and punish the behaviour of children in the classroom. Under the Children Act 1989, teachers have a duty of care towards their pupils, a concept referred to as 'in loco parentis' that has evolved through legal precedent. Legally, while not bound by parental responsibility, teachers must behave as any reasonable parent would do in promoting the welfare and safety of children in their care. The principle of 'in loco parentis' can be used to justify a teacher's reasonable use of punishment, although corporal punishment in schools has been outlawed in most of Europe for some time. Questions about legal responsibility and robots are complex and increasingly discussed (Asaro 2012). It is unlikely that the 'in loco parentis' principle would be applied to a robot, but a robot engaged in teaching activity would need recourse to some forms of sanction. Apart from rewarding or punishing behaviour, a robot teacher might need to prevent a child from performing dangerous actions, or from hurting their classmates, or injuring the robot. It is not clear what kinds of sanctions a robot could acceptably use. It might be that such questions mean that a robot could not feasibly be left in charge of a classroom of children, and would always need to be able to rely on a human supervisor to maintain classroom control.

Different roles and scenarios for classroom robots do create differing perspectives on these questions about control and accountability, and about decisions about what to teach. They are particularly salient when considering the possibility of an autonomous robot teacher (Scenario 1). A robot teacher could be programmed to teach on a particular topic, or to follow a given curriculum. However a human teacher will continuously make decisions about when and how to teach something, adjusting their delivery in response to their understanding of the situation and the audience. A robot, for reasons discussed above, is unlikely to be able to do this. In addition, in order to function as a classroom teacher, a robot would have to be able to control and make decisions about children's behaviour in the classroom. The argument is made here, and elsewhere, that robots do not have the necessary moral and situational understanding to be able to adequately, or acceptably, fulfil this role.

Control and autonomy are less of a concern in the case of the Telepresence robot in Scenario 4, since a human operator, or operators will presumably be involved; although the extent to which the remote teacher is distanced from the classroom situation is likely to limit their awareness of what is going on in the classroom. When the robot is presented as a companion or peer (Scenario 2 and 3), it is not seen as being in a position of authority, and there is less reason to be concerned about questions of control and autonomy. Nonetheless, if companion robots are to be used for teaching purposes, there is still a need to think carefully about any delegation of decision making capabilities. Even a robot presented as a companion could be required to make some decisions about a child's learning, or performance. Care needs to be taken to ensure that any such decisions are ones that it is appropriate for a robot to make. In other words, it should be clear that the decisions are made by programmed algorithms, and not the result of human-like judgement.

Reasons in favour of robot teachers

Although we have identified and discussed the main ethical concerns associated with the introduction of robots in the classroom in terms of 4 different scenarios, we have not yet considered the arguments that could be made in favour of classroom robots. Perhaps the ethical concerns raised here could be outweighed by compelling reasons in favour of deploying robots. In order to address this possibility, we consider the main arguments and reasons for replacing humans with robots in social roles, and the extent to which they apply equally, or differently, to the four classroom scenarios.

There are at least five general arguments that have been made in favour of the use of robots in society. First, it is often suggested that robots are particularly appropriate for situations that involve tasks that are dangerous, dirty or dull for humans to undertake, and that by taking on such tasks robots could free up humans for more interesting and rewarding activities (Takayama et al. 2008). A second reason for placing robots in social roles would be if they were found to outperform humans. For instance, if it were shown that children generally learned better from robot teachers than they did from human teachers, this could be a reason in favour of their adoption. A third reason for turning to robots is when they can offer something that would not otherwise be available. A fourth reason for deploying robots is as a signal that the organisation deploying them is technologically advanced and 'cutting edge'. A fifth reason in favour of replacing humans with robots is an economic one, based on claims that they will be more cost effective than the human alternative.

Do any of these reasons provide compelling justifications for the introduction of teaching robots into the classroom? Several do not stand up to much scrutiny. The first does seem particularly relevant to teaching. Few would see the teacher's role as being so dangerous, dirty or dull

for humans that we need to replace them with robots, as in Scenario 1. There is generally no shortage of people wanting to become teachers. If, in the future, teaching came to be perceived as a boring activity best left to robots, this would not augur well for the future of humanity.

The second reason depends on finding robot teachers to be better than human ones, and is one that is particularly relevant to Scenario 1, and the idea of a robot replacing the classroom teacher. Given the limited ability of robots to have a good understanding of what is going on in the classroom and in children's minds, this is unlikely to be the case in the near future. So far research on robots in the classroom does not usually involve a comparison between the effectiveness of robots and humans in conveying information. In much of the research reviewed in the section on "Current robots in the classroom" (e.g. Movellan et al. 2009; Kanda et al. 2004, 2007) the concern was to show that children can learn from a robot and accept it in the classroom, and not to compare the robots' effectiveness to that of human teachers. When a small scale comparison to human teachers was undertaken (Rostanti 2015), the robot did not fare well. Claims that robot teachers will be more motivating and effective for students than a humans need to be backed up by convincing evidence, and that evidence is not yet available. The possibility of robots making fairer decisions in the classroom than humans was discussed in "Control and accountability" section and argued to be an unlikely one.

The third reason is more viable, as there are situations and scenarios in which classroom robots could conceivably offer something otherwise unavailable. Telepresence robots (Scenario 4) for instance can be used to enable such a learning experience. The EngKey robot reviewed earlier was being used to give South Korean students access to English tutors in the Philippines. Likewise, a robot companion (Scenario 2) could augment a human teacher's lessons by providing some individual coaching. Children may even be more willing to admit their lack of understanding to a robot than to a human. Similarly, they might prefer to practice speaking a foreign language with a robot companion than with a person. As discussed in "Current robots in the classroom" section, an effective use of robots in the classroom that is beginning to emerge is when they are presented as a peer in need of help (Scenario 3, the Care-eliciting companion), so that the child has to teach the robot something. This was the case in the study described earlier by Tanaka and Matsuzoe (2012). Their preliminary results, and those of Hood et al. (2015), suggest that this approach can work well since the robot can be programmed to seem to need help from even a struggling student, thereby giving that student a rewarding feeling of competence.

The fourth reason for using robots as a means of indicating to the technological sophistication of the school or educational establishment may reflect the underlying motivations for such developments: but is of questionable value. Ensuring that children have some knowledge and experience of robots may well be a good thing, but it is important to critically evaluate the evidence about the extent to which robots can be used to enhance and facilitate learning before diverting too much of the limited educational funding budget towards them. This argument applies to all four scenarios discussed here. This leads us to the fifth reason. Replacing human teachers or assistants with robots because they are more cost effective is surely not something to be encouraged. Governments and local authorities might see some advantages to employing robot teachers; they would not demand pay, or strike, or complain about being asked to follow a prescribed curriculum. However, to be justifiable there would need to be good evidence of the robots' adequacy and competence for the role as compared to human teachers. This argument applies to all four of the scenarios we have considered. The cost effectiveness argument is one that may increasingly be made about robotics in various domains. It is to be hoped that discussions such as this that highlight the associated ethical concerns will help to reinforce and strengthen the arguments against such developments, and to ensure that robots are only introduced in situations where they can be shown to lead to an improvement in the human condition.

Most of the reasons we have considered here are not found to be good ones. Teaching is not a dangerous, dirty or dull task for which robots could appropriately replace humans. There is no compelling evidence that robots are better than humans at teaching children. The economic reason is not a powerful one unless the robots were shown to outperform humans, and the same is true for their use as a signal of the technological sophistication of the school or organisation.

The most convincing reason then in favour of robots in the classroom is that they can sometimes offer a beneficial educational experience that might otherwise not be available. This might be the case for the companion robots in Scenarios 2 and 3, and the Telepresence robots in Scenario 4. Generally, it makes sense to use robots in circumstances in which they can offer people access to resources and abilities that would not otherwise be realisable, rather than in situations where they are being used to replace competent humans. A related argument was made in the context of robots for older people in favour of deploying robots and robotic technology that expanded the set of capabilities accessible to them (Sharkey 2014).

Conclusions

Now that we have considered the main ethical issues raised by, and the reasons in favour of, classroom robots, some implications about the relative acceptability of the four

classroom robot scenarios can be drawn. These conclusions are based on the current and likely near future abilities of social robots, and it is acknowledged that they might need to be revisited if robots with significantly greater abilities are developed.

There are reasons to support the use of Telepresence robots (Scenario 4) when they are used to provide educational opportunities that would otherwise be inaccessible. For instance, they could be used to facilitate children's access to remote skilled teachers unavailable in their school. Their use as a cost-cutting measure should still be viewed with suspicion, and they do give rise to concerns about privacy and sharing of information, but nonetheless they could usefully supplement regular classroom teaching in some circumstances. Their use to facilitate contact with teachers and speakers of a foreign language seems appropriate, and if they are deployed in a classroom in which a human teacher is also available, there would be less need to be concerned about the issues of control and autonomy, and attachment and deception.

Companion and peer robots designed to foster implicit learning (Scenario 2 and 3) seem quite likely to appear in schools because they can function under the auspices of the human teacher without the need to control the classroom, or to appear fully competent. If such robots are to be welcomed, their welcome should be a cautious one because of the need to establish the educational effectiveness of such measures, particularly when compared to cheaper alternatives such as educational software and virtual coaches. In addition, since such robots masquerade as children's friends, there are concerns about the extent to which they would violate their privacy, and a risk that they would have a deleterious impact on their learning about social relationships. Nonetheless, if concerns about privacy and social relationships were addressed, it is possible that such robots could be used to offer new educational opportunities. For example, the idea of developing a care-eliciting robot that encourages children to teach it new concepts or skills (and thereby reinforce their own learning) seems a promising one. Similarly companion robots could be developed to provide individualised practice for children on tasks that require repetition (and that might be too dull or time consuming for human teachers). It also seems plausible that children might be more willing to admit a lack of understanding, or a need for repeated presentation of material to a robot than to a human adult.

The use of fully fledged robot teachers (the extreme of Scenario 1) is surely something that should not be encouraged, or seen as a goal worth striving for. There seems no good reason to expect that robot teachers would offer extra educational benefits over a human teacher. It is also apparent that robot teachers will not be able form an adequate replacements for humans in the near future. Robots are unlikely to have the ability to keep control of a room full of children in the absence of a human teacher (except in a nightmare situation where they could administer physical restraint and punishment to make up for their own shortcomings). A robot could be programmed to deliver educational material, but it is not at all clear that children would learn that material once the initial novelty of the robot teacher had worn off. In addition, even if it were possible to program robots to deliver a curriculum, that would not make them *good* teachers. A good teacher should be able to identify the zone of proximal development for a child, and be able to teach them just what they need to know, just when they need to know it (Pelissier 1991). As discussed by Sharkey (2015), a robot is unlikely to be able to determine the relevant information to teach to a student in any meaningful way. As non-humans, how could robots determine what human children need to know, or have the intention to pass on the information that is needed to accomplish the tasks required in human culture (Kline 2015)? First and foremost, children need to be taught by fellow human beings who understand them, care for them, and who form appropriate role models and attachment figures.

Acknowledgments This work was partially supported by the European Union Seventh Framework Programme (FP7-ICT-2013-10) under grant agreement No. 611971.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Ainsworth, M. D. S. (1973). The development of infant–mother attachment. In B. Caldwell & H. Ricciuti (Eds.), *Review of child development research* (Vol. 3, pp. 1–94). Chicago: University of Chicago Press.

Arkin, R. (2009). Governing lethal behavior in autonomous robots. Chapman-Hall review. *Computers and Education, 58*(3), 978–988.

Asaro, P. (2012). A body to kick, but still no soul to damn: Legal perspectives on robotics. In P. Lin, K. Abney, & G. A. Bekey (Eds.), *Robot ethics: The ethical and social implications of robotics* (pp. 169–186). London: MIT Press.

Bartneck, C., & Hu, J. (2008). Exploring the abuse of robots. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 9*, 415–433.

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers and Education*, *58*(3), 978–988.

Bergin, C., & Bergin, D. (2009). Attachment in the classroom. *Education Psychology Review, 21*, 141–170.

Borenstein, J., & Pearson, Y. (2013). Companion robots and the emotional development of children. *Law, Innovation and Technology, 5*(2), 172–189.

Bowlby, J. (1969). *Attachment and loss: Volume 1: Attachment*. London: Hogarth Press.

Brščić, D., Kidokoro, H., Suehiro, Y., & Kanda, T. (2015) Escaping from children's abuse of social robots. In *Proceedings of ACM/IEEE international conference on human-robot interaction* (pp. 59–66).

Calo, M. R. (2012). Robots and privacy. In P. Lin, K. Abney, & G. A. Bekey (Eds.), *Robot ethics: The ethical and social implications of robotics* (pp. 187–202). London: The MIT Press.

Carr, N. (2015). *The glass cage: Where automation is taking us*. London: Bodley Head.

Churchland, P. S. (2011). *Braintrust: What neuroscience tells us about morality*. Oxford: Princeton University Press.

Coeckelbergh, M. (2010). Health care, capabilities, and AI assistive technologies. *Ethical Theory and Moral Practice, 13*(2), 181–190.

Damiano, L., Dumoouchel, P., & Lehmann, H. (2014). Artificial empathy: An interdisciplinary investigation. *International Journal of Social Robotics, 7*(1), 3–5.

Dietvorst, B., Simmons, J., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126.

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating social connection through inferential reproduction: Loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychological Science, 19*, 114–120.

Epley, N., Waytz, A., & Caciopo, J. T. (2007). On seeing human: A three factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886.

Eurobarometer 382. (2012). *Public attitudes towards robots*. Bussels: European Commission.

Friedman, B., & Kahn, P. H. (1992). Human agency and responsible computing: Implications for computer system design. *Journal of Systems and Software, 17*(1), 7–14.

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transaction on Information Systems (TOIS), 14*(3), 330–347.

Han, J., Jo, M., Park, S., & Kim, S. (2005). The educational use of home robots for children. In *Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005)* (pp. 378–383). Piscataway, NJ: IEEE.

Han, J. (2012). Emerging technologies: Robot assisted language learning. *Language Learning and Technology*, *16*(3), 1–9.

Hashimoto, T., Kato, N., & Kobayashi, H. (2011). Development of educational system with the android robot SAYA and evaluation. *International Journal Advanced Robotic Systems, 8*(3), 51–61. **(Special issues assistive robotics)**.

Gaudiello, I., Zibetti, E., Lefort, S., Chetouani, M., & Ivaldi, S. (submitted). Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to the iCub's answers. arXiv:1510.03678[cs.RO]

Heyns, C. (2013). Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, A/HRC/23/47

Hood, D., Lemaignan, S., & Dillenbourg, P. (2015). When children teach a robot to write: an autonomous teachable humanoid which uses simulated handwriting. In *HRI '15*, March 02–05 2015, Portland, OR, USA

Jung, C. (1953). The development of personality, *Collected Works of C.G. Jung*, volume 17, Princeton, N.J. Princeton University Press.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human Computer Interaction, 9*, 61–84.

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human–robot interaction. *IEEE Transactions on Robotics, 23*(5), 962–971.

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioural and Brain Sciences, 38*, 1–17.

Koenig, M., & Sabbagh, M. A. (2013). Selective social learning: New perspectives on learning from others. *Developmental Psychology, 49*, 399–403.

Komatsubara, T., Shiomi, M., Kanda, T., Ishiguro, H., and Hagita, N.,(2014) Can a social robot help children's understanding of science in classrooms? In *Proceedings of the second international conference on human–agent interaction* (pp. 83–90).

Levy, D. (2007). *Love + sex with robots: The evolution of human–robot relationships*. Ltd: Gerald Duckworth and Co.

Lin, P., Abney, K., & Bekey, G. A. (2012). *Robot ethics: The ethical and social implications of robotics*. London: MIT Press.

Meah, L.F.S, and Moore, R.K. (2014) The uncanny valley: A focus on misaligned cues. In M. Beetz, B. Johnston, M. Williams (Eds.), *Social Robotics* (vol. 8755, pp. 256–265). LNAI

Movellan, J., Eckhart, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. In *Proceedings of 2009 international conference on human robot interaction (HRI2009)*.

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J. J. (2013). A review of the applicability of robots in education. *Technology for Education and Learning, 1*, 1–7.

Mutlu, B., and Szafir, D. (2012) Pay Attention! Designing adaptive agents that monitor and improve user engagement. In *Proceedings of Human Factors in Computing (CHI 2012)*.

Park, S., Han, J., Kang, B., & Shin, K. (2011). Teaching assistant robot, ROBOSEM, in English class and practical issues for its diffusion. In *Proceedings of workshop on advanced robotics and its social impacts*. http://www.arso2011.org/papers

Pelissier, C. (1991). The anthropology of teaching and learning. *Annual Review of Anthropology, 20*, 75–95.

Rostanti, S. (2015). *Can a robot teach?* University of Sheffield, Department of Computer Science undergraduate dissertation.

Sabbagh, M. A., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition, 112*, 415–422.

Sharkey, A. (2014). Robots and human dignity: The effects of robot care on the dignity of older people. *Ethics and Information Technology, 16*(1), 53–75.

Sharkey, A. (2015). Robot teachers: The very idea! *Behavioural and Brain Sciences, 38*, 46–47.

Sharkey, N., & Sharkey, A. (2006). Artificial intelligence and natural magic. *Artificial Intelligence Review, 25*, 9–19.

Sharkey, N. E., & Sharkey, A. J. C. (2010). The crying shame of robot nannies: An ethical appraisal. *Interaction Studies, 11*(2), 161–190.

Sharkey, A., & Sharkey, N. (2011). Children, the elderly, and interactive robots. *IEEE Robotics and Automation Magazine, 18*(1), 32–38.

Sharkey, A. J. C., & Sharkey, N. E. (2012). Granny and the robots: Ethical issues in robot care for the elderly. *Ethics and Information Technology, 14*(1), 27–40.

Sparrow, R. (2002). The march of the robot dogs. *Ethics and Information Technology, 4*, 305–318.

Sparrow, R., & Sparrow, L. (2006). In the hands of machines? The future of aged care. *Mind and Machine, 16*, 141–161.

Takayama, L, Ju, W., and Nass, C. (2008) Beyond dirty, dangerous and dull: What every day people think robots should do. In *Proceedings of Human Robot Interaction 2008*, March 12–15, pp 25–32

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Science, 194*(46), 17954–17958.

Tanaka, F., & Kimura, T. (2009). The use of robots in early education: a scenario based on ethical consideration. In

Proceedings of the 18th IEEE international symposium on robot and human interactive communication (RO-MAN 2009) (pp. 558–560).

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human–Robot Interaction*, *1*(1), 78–95.

Tanaka, F., Takahashi, T., Matsuzoe, S., Tazawa, & Morita, M. (2013). Child-operated telepresence robot: A field trial connecting classrooms between Australia and Japan. In *Proceedings of IEEE/RSJ international conference on intelligent robots and systems (IROS 2013), Tokyo, Japan, November 2013,* (pp. 5896–5901).

Van Wynsberghe, A. (2013). Designing care robots for care: Care centered value-sensitive design. *Journal of Science and Engineering Ethics, 19*(2), 407–433. doi:10.1007/s11948-011-9343-6.

Wallach, W., & Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. New York: Oxford University Press.

Winfield, A.F., Blum, C., and Liu, W.(2014) Towards an ethical robot: Internal models, consequences and ethical action selection. In M. Mistry, A. Leonardis, M. Witkowski, & C. Melhuish (Eds) *Advances in autonomous robotics systems: Proceedings of the 15th annual conference, TAROS 2014* (pp 85–96). Birmingham, UK, 1–3 September

Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2007). Interacting with a human or a humanoid robot? In *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*

Yun, S., Shin, J., Kim, D., Kim, C.G., Kim, M., & Choi, M. T. (2011). Engkey: Tele-education robot. In B. Mutlu, et al. (Eds.) *Social robotics: Proceedings of the third international conference on social robotics,* LNAI 7072, pp. 142–152.

Zhang, J., & Sharkey, A. (2012). It's not all written on the robot's face. *Robotics and Autonomous Systems, 60*(11), 1449–1456.