Robot nannies get a wheel in the door

A response to the commentaries

Noel Sharkey & Amanda Sharkey
University of Sheffield, UK

"We are right and all of the commentators are wrong", would be one approach to the commentaries. But our aim was to open a conversation among researchers from different cognizant disciplines related to robotics, childcare and society about the ethical dangers of raising children with robots. We wished to test our ideas and learn from a range of expertise different from our own. In this sense, the range of commentaries represents success. We are caught in the middle ground with a cluster of authors sharing similar positions surrounded by extremes on both sides. We are seen as both alarmist and too conservative in our concerns. We do not agree with all of the opinions but we find value in all of the commentaries. Some misunderstood or missed the points we made, some we think make errors of judgment and some caricature our arguments with disregard for the fine print. These receive a robust defence and perhaps a little too much of the space. All of the commentators are right some of the time but none is right all of the time.

We freely admit that much of what we said about the future development of robot technology is based on our own (conservative) extrapolation from what is currently available. There is little disagreement about our projections in the commentaries that robots have insufficient capability to be adequate nanny substitutes into the forseeable future – even from those who disagree about other issues. We are looking at a 5 to 15 year window into the future but make it clear that we can't be precise with the time estimates. That is a trap that many in AI and robotics have fallen into in the past. Many predictions made in the past have been out by a few decades and we are still waiting for more to come in. One can never be sure what breakthroughs in technology will accelerate or inhibit some of the expected developments.

The commentaries have not changed our opinions about the coming of robot nannies except that, if anything, they have strengthened our view that they will come – although we sincerely hope that we are wrong. We have amended our views about how the robot nanny will get its wheel in the door and our views about potential psychological damage have been broadened. We also take on board new ideas about regulation, over-regulation and developing robot literacy.

Interaction Studies 11:2 (2010), **302–313**. DOI 10.1075/is.11.2.20sha
ISSN 1572–0373 / E-ISSN 1572–0381

A number of commentators make a case for the positive application of research on educational and socially assistive robotics as an augmentation of normal care. We welcome their inclusion here in showing the beneficial aspects of robotics (which are many) and we have discussed some of the benefits in areas such as therapeutic applications for children in paragraph 2 of our target article. Some are defensive about their research domain and perhaps entertain the false belief that we are attacking them in our target article. Maybe we have not been clear enough that it is not the research that worries us. It is its potential commercial applications.

We value socially assistive and educational robotics research and research on children and robots in general. We rely on such research to find out the good, the bad and the ugly of robot use and we harvest the results to gain ideas about the ethical problems of potential commercial exploitation. However, research, for example on a social ecology of robots (Sabanovic), does not preclude the use of robots in the home for exclusive childcare (although it may help the public to understand the limitations). It would be naïve to believe that research results are not often used in ways that the scientists did not intend or sometimes in ways that they have not have imagined. Think of how much of all of our research on robotics has contributed to the development of advanced military (killer) robots which many of us may not have wanted. We cannot prevent our research being used for any purpose once it is in the public domain.

The target article is focused on the ethical dangers that might arise as a result of robots being used as a substitute for human nannies. We have not considered the types of educational use of robots as discussed in the excellent commentary by Tanaka and Kimura or the meso level of analysis discussed by Sabanovic. We do have some ethical questions about the eventual use of social and educational robots and the idea of increasing the allure of a robot by making it an object that needs to be cared for. But these are concerns about the potential exploitation of the research and not the research itself. Melson cogently expresses her worries about this area in her commentary but our response to the commentaries is not the place to open a new debate.

The main concern in the target article is with the potential for near-exclusive or exclusive care of children by robots and their protection from psychological harm. Some have hinted that we might be a bit alarmist. We are aware that we are presenting the extreme case but that was a deliberate strategy. An engineering testing strategy that we have often employed in our other work is to take systems to breaking point and then work backwards to find out at what point the problem began to emerge. So we decided to examine what kinds of psychological damage could result in the extreme case of robot care where experiments cannot be conducted and work backwards from there with the help of the commentaries.

We chose this forum to directly target researchers because we believe that they are best placed to bridge the gap between public expectations of robotics, often informed only by science fiction, and the limitations of the technology. We encourage researchers to engage with the public and policy makers. Not just to educate (or patronise) but to learn from the public and take their concerns into account in design and in discussion with funders and policy makers. We hope that others from different disciplines will take up the challenge of working backwards from the extreme. The realisation of this hope has had a great start in the commentaries.

Unfortunately it was not possible to have a detailed discussion of each of the commentaries due to limitation of space and time. We have organised our responses into three major themes that are most relevant to the target article: (i) what are the main drivers to get childcare robots into the home? (ii) what type of harm could robot nannies cause? (iii) what sorts of regulation and guidelines should we have? There are many more ideas in the commentaries than can be covered by these themes. Some, such as Carme Torras' excellent piece on science fiction and robot nannies, hardly get a mention at all. We apologise to the authors for not being able to discuss them here and hope that their commentaries will still have an impact on the readership.

1. The creep of the robot nannies

Alarm bells started ringing for us after the issue of childcare robots came up in discussions that Noel regularly holds with large numbers of the general public in different countries. After some research, we discovered that several companies were developing childcare robots and that set us thinking about how robots could come to be acceptable in a childcare role. We examined the properties of these robots carefully to get an idea of how the robots might get a wheel in the door.

At the start of Section 2 of the target article, we asserted that "An essential ingredient for consumer trust in childcare robots is that they keep children safe from physical harm." Disappointingly, few of the commentators picked up on this issue except tangentially. We suggested that most of the childcare robots we had reviewed provided some form of mobile monitoring. Typically the child can be viewed through the robot's camera eyes on a mobile phone or on a window in the corner of a computer monitor. The child can also wear a tag or beacon so that the robot knows where she is located and many of the robots could go in search of the child. We pointed out that to be commercially successful in the future, much of the monitoring would have to become autonomous. If this works well, and we are pretty sure that it is possible, it could make the childcare robot attractive to many more busy professional consumers.

We suggested that through clever marketing, "busy working parents might be tempted to think that a robot nanny could provide constant supervision, entertainment and companionship for their children." We used the example of how the very simple Hello Kitty robot appeared to be marketed for the purpose of childcare. Movellan reports that it was not the company who were making these claims but an independent marketing company. While this distinction is useful when it comes to the allocation of responsibility, it makes little difference to our argument about the exploitation of people's misunderstanding about the abilities of robots.

Feil-Seifer and Mataric argue that childcare robots will not appear on the market because Socially Assistive Robots would not fool people for long into believing that they are *socially competent*. They say that the children in their studies aged 5 to 15 can very quickly determine that a robot is not as socially intelligent as a human being (our target age is younger). They state that this is a refutation of our arguments. But this is a difference of opinion and not a refutation. A refutation would require, at the very minimum, some analysis of the evidence we present from other studies (Section 3 of the commentary). Not only that, but they are attempting to refute a position that we do not hold.

We agree with their expert analysis of the state of robotics. We make no claim that robots will be seen as socially competent. We don't need to. Our argument rests on the exploitation of natural anthropomorphism. The human tendency to view inanimate objects having certain "life giving" trigger characteristics as if they were alive has been understood and employed by toymakers, puppeteers and automata makers for millennia and more recently by cartoonists and animators. Our point about the illusion of animacy (and mental states) in robots is simply that by adding dynamic properties such as movement of the body, facial expressions, features and other interactive devices, natural anthropomorphism will be amplified. We present considerable evidence from other studies about the relationship and bonding between children and robots in Section 3 of the target article. Other commentators agree. Mercer tells us, "it would be possible for basic attachment behavior to develop; toddlers show this in rudimentary form to preferred blankets, teddies, and even familiar places." Melson also agrees but extends our discussion considerably.

Movellan states that children above the age of 9 months have no problem in assessing the limits of robot skills after just a few minutes of interaction. Apart from the fact that even professional robot testers cannot assess the limits of a robot's skills "after a few minutes", he does not address our evidence in Section 3. In fact some of the work he cites to support this claim, his classic study with Tanaka on 18 to 24 month olds, is actually part of our evidence: "Results indicate that current robot technology is surprisingly close to achieving autonomous bonding and socialization with human toddlers for sustained periods of time" and, in the conclusions,

"Children exhibited a variety of social and care-taking behaviors toward the robot and progressively treated it more as a peer than as a toy" (Tanaka et al., 2007).

Melson's commentary resonates with our article in pointing out how endowing robots with ever more sophisticated capacities combined with marketeers touting the myriad benefits of robot to harried parents will lead to proliferation of robots in many aspects of children's lives. But she goes further than us by opening the issue into the wider context of the decreasing social investment of face-to-face human–human interaction and relationships. This is also echoed in the commentaries of Marti and Kubinyi et al.

Marti discusses how children in the 21st century differ from those of previous generations. She presents compelling evidence for the change in their preferred media to the computer and the web. She suggests that the children of today are "native speakers" of the digital language of computers, video games, and the internet. In other words, they are digital natives. Where Marti takes a positive outlook on the digitalisation of culture, Kubinyi et al., from an ethological perspective, worry about the damage this could cause to the nature of humanity. They warn that if we continue along the road of substituting artifacts in human functions that a new form of human, *Homo technicus*, could emerge (a great term).

Melson rightly terms these increasing interactions with and through technology, e.g. voice-activated systems, as *pseudo-interactions*. The digital native will not need to believe that the robot is socially competent or human-like to engage with it. As Melson puts it, a robot's approximations to the 'real thing' may seem *good enough*.

Others believe that the robot minders may get a wheel in the door by less direct means. Bryson provides an insightful suggestion that manufacturers will not mislead the public directly through advertising as it would leave them open to legal liability. Instead they will make robots that are very attractive and appealing to parents and children and leave parents to discover the potential for childcare themselves. The manufacturer may even issue disclaimers on the box advising that the robot should only be used for entertainment.

Wallach's commentary supports Bryson's when he talks about how regulation might just result in manufacturers changing what is on the tin to the likes of educational toy or "mommy's second set of watchful eyes". This has a strong ring of plausibility and reminds us of a sign we once saw in the window of a store in San Francisco in the 1980s selling marijuana pipes: "all pipes in this store are intended only for use with tobacco".

In a similar vein, Whitby discusses how the robot nanny could arrive in the home by stealth. He provides a possible trajectory in which parents start out with good intentions and may even limit the time the child spends alone with the robot. But the "long hours culture" can create extreme pressure that will progressively

reduce the good intentions. Thus one of the dangers of the augmented childcare suggested by Petters et al., is mission creep. It starts off helping mommy but through successive upgrades and trust, it takes on more and more of the minding role. Yujin's Jupiter robot video shows how it can help mom in the kitchen with recipes on the internet. It can help to look after granny and the kids.

We accept that this stealth scenario may be a more likely way for the nanny robots to progressively get both wheels through the door and into the home rather than through manufacturers making false claims for which they may be held accountable. And of course there are other stealthy means to imply the childcare uses of robots in the home. For example, if you watch NEC's PaPeRo videos (received from NEC on request) you will see a young child sitting alone and then you see the mother in another location calling the robot on her cell phone. The robot goes and finds the little girl and the mother interacts with her via the phone and the robot. The power of the "Hollywoodisation" of robot marketing should not be overlooked.

2. What's the harm of a robot nanny?

We drew on evidence from the psychological literature on attachment and neglect to look at the possible harm that could result from the extreme case of leaving babies and young children in the exclusive, or near-exclusive care of robots. Taking the lead from research suggesting that children could form close bonds and relationships with robots, our focus was "on the ways in which a child's relationship with a robot carer could affect the child's emotional and social development and potentially lead to pathological states". Looking at current and near future robot abilities, our conclusion was that a robot nanny could not constitute an adequate replacement for human care. Robots cannot interact with the level of sophistication needed for the type of attachment required for healthy development. Children raised exclusively, or mostly by robots would be likely to show signs of neglect, and could even develop attachment disorders.

All of the commentators, with the possible exception of Belpaeme and Morse, agree that the idea of exclusive childcare by robots would be a bad idea. Most recognise that it could be harmful to children. There is some disagreement from Petters et al. who suggest that children are unlikely to form attachments to a robot nanny if there are any other humans around – particularly those in control of the robot.

Of course, attachment to a caregiver, as we point out, is not an end in itself. The caregiver needs to maintain that attachment, and guide the baby, and subsequently the toddler, as it develops. Various commentators have looked at the behaviours required of a good childminder or nanny, and the extent to which robots might be

able to produce them. Most commentators emphasise the gap between what has been achieved in robotics and what would be required.

Belpaeme and Morse are the exception again here in suggesting that what they describe as the "doom scenario" won't happen because progress in robotics means that robots will be much better at social interaction etc than Harlow's wire covered monkeys. They claim that such progress means that it is possible to make robots "with which children engage over a prolonged period of time". But this is part of the problem and not the solution. It touches on a 'raw nerve' at the core of our concerns that robots are not an adequate replacement for human care and prolonged interaction could lead to serious psychological damage.

Zoll and Spielhagen take a more extreme position than us about the dangers of even short term interaction with a robot nanny. They criticise our suggestion that leaving a child alone with a robot for short periods would be no worse than leaving them in front of television. First they cite research showing the harmful effects of television. Then, more relevant to the target article, they argue that robot care would be worse because an interacting robot inevitably encourages the formation of a relationship that could impact on attachment styles. For them, even part-time robot care of children might affect their attachment styles. We feel that this is worth further investigation. They also make a crucial point that it is not just a matter of showing that robots do not harm children before we use them but we should be sure that they are actually beneficial.

Castellano and Peters minimally require the robot nanny to be socially perceptive: "The first thing that a parent would require from a robot that is advertised as capable of taking care of children is the ability to understand where the children are and what they are doing, what their emotional states and future intentions may be, and relate these to the events unfolding (and about to unfold) in the environment." This goes beyond recognising simple emotion through facial features to identifying the causes of an emotional state. These commentators are clear that we seem "far from a system that is capable of interpreting the events unfolding in the environment, or processing them in an integrative or reflective manner based on mentalising and empathising capabilities".

Van den Broek similarly emphasises the difficulties involved in recognising the emotion of the child but suggests that robots could be equipped with biosignal sensors (electromyography, electrocardiography, and electrodermal activity). He looks at the problems faced by long term robot care: the need for a robot to change its behaviour as the child develops, and the need to take personality, and cultural differences into account. This leads him to the conclusion that the challenges are so great that robot nannies may remain the subject of fiction. But we maintain that not having the required abilities will not prevent the use of the robot as a human substitute as we have been arguing throughout.

Mercer, a developmental psychologist specialising in attachment, encapsulates this need for subtle human–baby interactions in her commentary. She examines some of the delicate interactions between infant and carer that form the basis for attachment. For example, mothers of young babies quickly imitate and mirror emotional expressions, providing feedback about their baby's emotions. Since even depressed human mothers often provide poor feedback, resulting in babies who are slowed in language development and who often look inappropriately sad, she suggests that the chances of a robot nanny being able to provide a good foundation "from which personality can develop and on which appropriate social behavior can be based" are small.

Apart from issues about attachment, some other interesting suggestions were made about further harmful effects that could result from interacting too much with robots. Kubinyi et al. cite work on cross-species adoption in animals to make the point that interacting with robots might affect human behaviour to such an extent as to influence mating preferences. Bryson suggests that some children might come to prefer interactions with robots, as more predictable, and that "Those children who prefer predictable interactions may be setting themselves up for a life-long preference for machines to humans".

Melson suggests that even if robots are not used for the exclusive care of children, they may nonetheless cause a reduction in the amount of time that children spend interacting with other human beings. She cites some interesting evidence of the effects that interacting with technology might have on our view of and relationship to living things. Torres also discusses the effect that interacting with robots is likely to have on us, quoting from the philosopher Robert Soloman: "It is the relationships that we have constructed which in turn shape us". Thus if we interact with a slave robot, that might make us despotic, or with an efficient robot we might stop thinking for ourselves.

Finally, there was some discussion from the commentators about age related effects of robot care. The extent and type of ill effects or problems that children may suffer as a result of robot care are likely to be a function of how old they are. We were deliberately a little vague on this in our target article: "Given the paucity of research on childcare robots we have not been age specific, but our concerns are predominantly with the lower age groups – babies to preschoolers up to five years old – that appear to be the target group of the manufacturers." (p. 174).

Clearly, the requirements for a robot nanny differ with the age of the child. There is some disagreement amongst the commentators as they discuss the possible harmful effects of robot carers in terms of age. While Movellan suggests caution for babies under the age of 12 months, he argues anecdotally that there is a "consensus" that children older than this will not be adversely affected since their primary attachments will already be formed. Ruiz-del-Solar agrees that robot care

of the young should be avoided, but considers that a robot nanny could be used to mind older children who would otherwise be home alone. It could keep them company and away from dangerous influences such as street gangs. Ruiz-del-Solar gives 5 years old as the dividing line.

In contrast, Wallach suggests that "the development of a child sleeping in the warm soft arms of a rocking robot" is unlikely to be worse than that of a child sleeping alone in a crib, whilst thinking that questions about developmental damage are more complex with older children. We agree with Wallach in this specific instance of rocking a baby but there are many other circumstances where the robot interaction could adversely affect a baby.

Petters et al. argue that the behaviours required of a robot nanny become more challenging as the child gets older: "As the child's cognitive skills increase, the attachment figure's contributions become increasingly abstract and difficult to implement in software".

Mercer considers the need for an infant's carer to adjust their responses and behaviours as the child develops and grows older. Even nappy changing is done in a different way for a wriggly 8 month old, and a tiny baby. Mothers constantly change the way in which they speak to babies, and the content of their speech; for a young baby, a high pitched cooing voice is more appropriate, whereas as the baby become a little older, some simple vocabulary that reflects what the child is looking at becomes more appropriate.

Clearly, further detailed analysis of the psychological literature beyond the scope of our article is required to look at the likely effects of robot care during critical periods in development of cognition, affect, language, social awareness and so on. We hope that some of the developmentalists reading this will take up the task.

3. Regulations, guidelines and discussion

A major concern in our target article was about the protection of children. To this end, after discussing potential psychological damage, we considered the adequacy of current legislation and international ethical guidelines to handle the overuse of robot care. The main legal protection would be under the general laws of neglect. We pointed out that as robots become safer and can protect children from physical harm, it will become harder to make a case for neglect. One of our points was that psychological harm is more difficult to demonstrate as it could be attributable to other factors such as the child's predispositions or even to insensitive parenting. Thus leaving a child in "the safe hands" of a robot that protects it from physical harm may not always be considered neglectful under current legislation. This is a very good reason to call for discussions with the public, the policy makers, scientists

and professional bodies. Iterative engagements would help to set up guidelines and regulate the potential for harm.

In his commentary, Whitby explicitly addresses the problem of codes and legislation and concurs with us on the inadequacy of current legislation in the UK: "the current complete absence of any guidelines leaves a situation in which the worst side of commercial pressure will work to not only make the technology freely available, but also to encourage unethical lines of development." He also offers useful advice about why the creation of professional codes can bridge the gap and put a brake on the rapid manufacture of nanny robots until proper regulation is in place. While not wanting to shift responsibility for bad parenting onto the manufacturers, Whitby discusses how society has already been remiss in allocating responsibility effectively to designers and programmers for the bad consequences of their products. He points to the role to be played by professional bodies in ensuring that the manufacturers don't shift all of the blame onto the parents.

In stark contrast, Movellan sees our call for protection as overly zealous. He appears to find the mere suggestion of guidelines and regulation to be offensive to teachers and parents. But surely pointing out potential dangers of using childcare robots is no more offensive than trying to establish any laws or treaties for child protection. There is a long history of child cruelty, abuse, undue hardship and harsh labour practices with the knowledge of parents that have dramatically been reduced in richer countries through legal measures. Of course the majority of parents want to do the best for their children and can use their common sense, but common sense can be misled. We certainly see horror stories in the news about the treatment of children on a daily basis.

Movellan argues oddly against regulation by providing an example of an early regulation requiring a man to walk in front of every moving vehicle with a red flag. However, this should not be taken as an argument for non-regulation of all new technology including robots any more than it should be taken seriously as an argument against traffic regulation. As Bryson points out, there is "horrific cost in loss of life and well being, to say nothing of environmental damage" caused by the automobile. And yet it is still with us because it is critical to the economy and individual freedom. Worryingly, she points out that if robots become as essential to our economy as automobiles, they will also become immune to arbitrary legislative bans.

Movellan's anti-regulation stance comes out strongest against our comments about South Korea having problems in the past because they "strongly resist interference in family lives by outsiders". As the basis for these problems we cited the findings of Hahm and Guterman (2001), "South Korea has had a remarkably high incidence and prevalence rates of physical violence against children, yet the problem has received only limited public and professional attention until very recently."

Movellan's response was, "I was surprised that anybody would consider this [resisting interference in family lives] to be a problem. I would hope for as many cultures as possible to adhere to the wisdom of this aspect of South Korean culture".

We have a certain amount of sympathy with Movellan's position on regulation. We do not want interference from the government in our private lives either. But we would like to see protections for children everywhere from potential parental abuses and even from businesses. Ruiz-del-Sol makes very useful points about the need for regulation, as indicated in how it has been used effectively for toys, video games, medicines – even saunas and Jacuzzis etc. He reminds us of advice on labels for toys to make them age appropriate and on devices in sports facilities that are not appropriate for young children. We can't see why guidance and regulation on the use of robots for childcare should be any different.

Wallach provides a counterbalance with his sensible reminder of the dangers of over-regulation when dealing with a new technology such as care robots. Such new technologies are moving targets and so short-term issues may be ameliorated over time. His worry is that regulation may be, "perceived as stop signs, which might lead to regulations that are perceived as being unnecessary and as interfering with progress." We accept this as a reasonable point but the jury is still out as to whether we actually need a "stop sign" for robot nannies or not.

Castellano and Peters discuss the idea of educating parents and providing guidelines about what is acceptable practice with new sensitive technologies. One example in their discussion of socially perceptive robots is about the use of knowledge about a child lying and how this could be exploited in persuasion. They suggest drawing up clear guidelines as to what response to a child's emotional state can be considered ethical and safe and that people need to be made aware of the risks of using a socially perceptive robot.

We take the point about legislating too early (although this seems very unlikely) and about the problems of over regulating. However, at the very minimum we should begin to draw up guidelines with practitioners and parents. The nanny codes of ethics are a good model of a professional code that, while not legislative, is used in training on good practice.

4. Conclusion

We responded to the commentaries on our target article in three sections. In the first, despite some differences of opinion, we felt that the commentaries reinforced our view that robot nannies will find their way into homes in the not too distant future. We accepted the idea that this is more likely to happen by stealth than through explicit advertising. In the second section, we saw a greater range of possible

psychological problems than we had first proposed. We acknowledged that more research is needed on the amount of time children can safely spend with robots and on the impact of robot care on different age-critical points. In the third section, we defended our view that the current protections for children are insufficient. There was considerable support for this position amongst the commentaries and we took on board a number of suggestions about how best to proceed with guidelines and how to be cautious about over regulation.

Now having looked into the "pit of despair", we would like to offer one ray of hope that between us all, we can find a way to limit the potential dangers of robot care. The method we have suggested, as a first step, is to have an iterative public engagement process between public, scientists, professional bodies and policy makers – inform, listen, discuss, listen and act. This combines well with Melson's idea of creating *robot literacy*. Parents and children can learn about how robots are produced by humans, about their limitations and about how they differ from the living.

This idea of robot literacy has echoes through several of the commentaries with suggestions about involving children and parents in a collaborative design process. If robots are inevitably going to be part of our lives, the work of the social/educational roboticists and theorists may promote greater experience and understanding of robotics that will help to save children from the excesses of extreme robot care. We hope that the commentators and readers will feel sufficiently moved to continue this conversation beyond our closed circle.