

**HUMAN-COMPUTER
INTERACTION**
A Journal of Theoretical, Empirical, and Methodological
Issues of The Interaction of Humans and Computers

Volume 35, Number 4
October-December 2013



The Intentions of Washing Machines

Journal:	<i>Human-Computer Interaction</i>
Manuscript ID	Draft
Manuscript Type:	Commentary
HCI Keyword Taxonomy:	Artificial Intelligence

SCHOLARONE™
Manuscripts

1
2
3
4 **Commentary: *The Intentions of Washing Machines***
5
6
7
8
9

10
11 **Richard H. R. Harper¹**

12 *¹Lancaster University, England*
13
14

15 .
16
17 **Richard Harper** (r.harper@lancaster.ac.uk, <https://rhrharper.com>) is Director of the Institute for
18 Social Futures of Lancaster University.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Hancock makes a range of claims but the most important is this: if a machine ‘learns’, then,
4 eventually, it will become ‘self-aware’. It is self-awareness, he argues, that will distinguish
5 machines that are merely autonomous (i.e., which work without human intervention, of which
6 there are many) and those which do something else, which become, in the things they do, like us
7
8

9 I cannot understand why one would think this move from learning to awareness would
10 happen but Hancock is convinced. One might add that it is not his discipline that leads to this
11 view - there is no human factors research that asserts or demonstrates that self-awareness
12 emerges through learning, for example; or at least as far as I am aware of. Certainly, Hancock
13 does not cite any. On the contrary, it seems that Hancock takes this idea from the AI community,
14 though as it happens it is an argument that coat-tails on similar notions put forward by cognitive
15 scientists. Some philosophers argue the same, too, such as Dennett (For the view from AI and
16 computer science, see Russell 2019; for the view of cognitive science, see Tallis, 2011; for a
17 review of the philosophy see Harper, et al, 2016). Be that as it may, let me focus on this claim
18 and ask what ‘self-awareness’ might mean or how it might be measured.
19
20

21 It seems to me that this is a question to do with anthropology. Hence, one way of
22 approaching this is through imagining how people would act when self-awareness is at issue
23 (Pihlström, 2003: 259-286). Or, put another way, what does someone mean when they say they
24 are ‘self-aware’ and why would they say it? I think they do so if they are ‘conscious’ of such
25 things as their intentions. ‘I am about to do this’ they say when they are wanting some advice on
26 that course of action. Intentions are a measure of self-awareness.
27
28

29 So, is Hancock saying that autonomous machines would be conscious of their intentions and
30 would that mean, too, that they would treat these intentions as accountable matters? Would that
31 mean, say, that a washing machine could have intentions of various kinds? And more, would it
32 mean that these emerge from the learning that the washine does?
33
34

35 There are a number of thoughts that arise given this anthropological ‘vignette’ of washing
36 machines and their intentions. How would these intentions be shown? Would the machine need
37 to speak? Besides, when would the machine have these intentions? At what point during learning
38 would they arise? After it has been working a while? One might presuppose some answers here –
39 a machine might only ‘speak’ (if that is its mode of accountability) only once it is switched on.
40 Moreover, one imagines a washing machine would not have any intentions when it was being
41 assembled nor would it have any when it was being disassembled either (as it happens, Hancock
42 refers to similar matters when he reminds the reader of one of his many phrases in earlier human
43 factor articles: this time, the phrase ‘isles of autonomy’. This is an allusion to how current
44 machines are only autonomous at particular moments in their life, as elsewhere in that life they
45 are subject to the control of and management by people. So, here: a washing machine might only
46 have intentions once it has been made and when it is switched on).
47
48

49 OK. Let us presuppose a washing machine can have intentions only when it is on, that it has
50 done enough washing to learn and is now ready to be self-aware. And let us presuppose too that
51 this awareness is manifest in how intentions come to be accountable. Now, in being accountable,
52 it is not only others that might comment on intentions; a washing machine might too. In having
53 intentions, washing machines can have a relationship to their own intentions. A washing machine
54 could say, sometimes, ‘I do not intend to do the washing’. Self-awareness is not just a question
55
56
57
58
59
60

1
2
3 of knowing what one's intentions are as being able to sometimes disavow intentions. Awareness
4 therefore means choice (Harper, et al, 2016). Once seen this way, one can see that this might also
5 mean that a washing machine choose not to do something. That would be an expression of its
6 choice. 'I will not do the washing!' the machine might say. The manner of a teenager comes to
7 mind – sometimes words, sometimes movements, but all too often, 'No!' Given this, one wants
8 to ask why would a washing machine say no? Is it because it wants (has the intention) of doing
9 something else? Is its 'intelligence' somehow derived from its learning and self-awareness such
10 that it could, say, choose to be a search engine? Or a game player (teenagers again)? Could it opt
11 to do the hoovering? (Teenagers take note).
12
13

14 The reader will be thinking that this is getting fatuous. But this is where claims about
15 awareness that Hancock happily builds on end up – not with subtle notions of how people are
16 richly complex, at times unfathomable creatures whose intentions can be opaque even to
17 themselves (Das, 2020). It doesn't lead to an anthropology where the phrase 'I will not do this!'
18 is a proof of the humanity of those being studied.
19
20

21 Is this what Hancock is thinking of when he suggests that machines might come to be like
22 us? My remarks show this metaphorical assertion is curiously floppy and hence potentially
23 unhelpful (for a review of the problems of metaphor see Donoghue, 2014). It limits what is
24 implied rather than offers a comprehensiveness that adequately characterizes what people do
25 (and of the arrangements constitutive of society that allow those doings to be done) such one
26 could judge whether some machines are 'like us'. For machines to be like us they need to be
27 wilful and able to 'walk away' as we sometimes do; as teenagers too often do. But this does not
28 seem to be what Hancock is imagining. His autonomous machines are not like 'us' in these
29 respects. He eschews these concerns and simply asserts that AI machines learn and therefore
30 must become self-aware 'like us' without wanting to explore what this might mean.
31
32
33

34 And this brings to mind one of the more interesting debates in common rooms at the current
35 time (if not yet in the journals). It is about what AI machines 'are'. The answer to this is often a
36 function of the disciplinary view that the protagonists in the debates take. Computer scientists
37 and statisticians take quite different starting points, for example. Statisticians think AI is an
38 unnecessary name for what are 'statistical engines' which function through consuming their own
39 output: statistics in, and then statistics out; endlessly, until a resolution is made: a call about
40 adequacy, about when the numbers have said enough (Bayes would turn in his grave in delight!).
41 The philosophy of statisticians is loosely encapsulated in George Box's credo that 'all models are
42 wrong, some models are useful' and in any case, models 'are only useful when their purpose is
43 well-defined' (Box: 2013). Consequently, for statisticians, data are always particular, always
44 devised to answer particular questions; hence the machines in question are not generalised or
45 generalisable. They are constrained to the particulars.
46
47
48

49 In contrast, computer scientists (or rather those doing AI) think their 'engines' are greater
50 than, and independent of, the 'data' they use. They reach beyond the particular to the general. It is
51 not particular purposes that guide them, but general ones. This leads to the somewhat odd notion
52 that the 'New AI' (as it is called, see Taylor, 2021) will lead to 'generalised intelligence'. It is
53 odd because these same advocates readily admit that 'Good Old-Fashioned AI' (or 'GOF AI')
54 failed precisely because it sought generalised intelligence. In contrast, the very success of their
55 own (New AI) is derived from its sticking with the particulars. Search engines are good because
56
57
58
59
60

1
2
3 they only do search queries; not because they do some abstract form of intelligence.
4

5
6 So why do advocates of the New AI say there will be a return to the general? I think I know
7 the answer, and it has nothing to do with empirical proofs about Turing Tests or the applicability
8 of metaphors that hold that machines are like us. It has more to do with disciplinary premises –
9 ones in computer science. It seems to me that they are taking a classic computer science view
10 which says that programmes are essentially distinct from the phenomena being processed, the
11 data. This is the Turing theoretic approach, typically combined with Von Neuman architectures
12 (that store both input data and data to do with the programme) that has delivered the
13 marvellously adaptable generalised machines we see around us today. AI researchers are
14 mistaking the programme for the machine, and in their mistake, encourage others, outside the
15 discipline of computer science, to misunderstand what AI can do (and how it does it). It leads
16 people to think that machine learning will lead to self-awareness, for example, and not just to
17 machines that do their particular tasks better. In my mind, AI researchers need to be reminded of
18 some basic home truths. To restate: the applications that one sees around us today called ‘AI’ are
19 designed to do particular things. They have no self-awareness that leads them to ask about their
20 intentions. Search engines, as a case in point, don’t muse on whether they might do the washing
21 up; intelligent hoovers don’t offer the user guidance on the design of their PPT slides; Siri
22 doesn’t drive the car (or do the laundry). What we know of as AI machines today, indeed nearly
23 all of which are instantiations of the ‘New AI’, are good at what they do since they limit their
24 functionality to specifiable tasks. They don’t become self-aware. If they tried, they would end up
25 like the Old AI, a failed enterprise. Hancock take note.
26
27
28
29
30

31 REFERENCES

32
33 Box, E.P. (2013) *An Accidental Statistician: The Life and Memories of George E.P.Box*, Wiley
34 & Sons: New York.
35

36 Das, V. (2020) *Textures of the Ordinary*, Fordham Press: New York.
37

38 Harper, R., Randall, D. & Sharrock, W. (2016) *Choice: The Sciences of Reason in the Twenty*
39 *First Century: a Critical Assessment*, Polity Press: Cambridge.
40

41 Donoghue, D. (2014) *Metaphor*, Harvard University Press: Mass.
42

43 Pihlström, S. (2003) On the concept of a philosophical anthropology, *Journal of Philosophical*
44 *Research*, Vol 28, 259-286.
45
46

47 Russell, S. (2019). *Human Compatible: AI and the problem of control*, Penguin: London.
48

49 Taylor, P. (2021, January 21). *Insanely Complicated, Hopelessly Inadequate*, *London Review of*
50 *Books*: London.
51

52 Tallis, R. (2011). *Aping Mankind*, Routledge: Abingdon..
53
54
55
56
57
58
59
60