AI & SOCIETY
https://doi.org/10.1007/s00146-021-01230-z

ORIGINAL ARTICLE

In principle obstacles for empathic AI: why we can't replace human empathy in healthcare

Carlos Montemayor[1] · Jodi Halpern[2] · Abrol Fairweather[1]

Received: 16 January 2021 / Accepted: 11 May 2021

Abstract
What are the limits of the use of artificial intelligence (AI) in the relational aspects of medical and nursing care? There has been a lot of recent work and applications showing the promise and efficiency of AI in clinical medicine, both at the research and treatment levels. Many of the obstacles discussed in the literature are technical in character, regarding how to improve and optimize current practices in clinical medicine and also how to develop better data bases for optimal parameter adjustments and predictive algorithms. This paper argues that there are also in principle obstacles to the application of AI in clinical medicine and care where empathy is important, and that these problems cannot be solved with any of the technical and theoretical approaches that shape the current application of AI in specific areas of clinical medicine in which care for patients is fundamental. This is important, because it generates specific risks that may be overlooked otherwise, and it justifies the necessity of human monitoring and emotional intervention in clinical medicine. Consequently, difficult issues concerning moral and legal responsibility may ensue if these in principle problems are ignored.

Keywords Empathy · Consciousness · Attention · Care · Clinical medicine

1 Introduction

It is generally assumed in AI research that there are no "in principle" or a priori limitations about the applications and range of AI. In contrast, we argue that empathic AI is impossible, immoral, or both. Empathy is an in principle limit for AI. While the current argument is likely to generalize to other professions that rely on empathy, our attention here is confined to outcomes improved by empathy in clinical medicine and why AI cannot achieve these. Thus, our argument is not dependent on practical considerations and is limited in scope. Since it is an in principle problem, considerations about architecture, design, computer power or other pragmatic issues are not helpful in addressing it. But given that we consider primarily the area of patient care, rather than other aspects of medical applications of AI such as diagnostics, resource optimization or data gathering, in which AI has enormous potential for improving medical services, the difficulty we present does not amount to a categorically general objection to the use of AI in medicine.

The problem we present centers on the notion of empathy. As has been clarified recently in the literature,[1] various definitions of empathy in the literature emphasize three key components of empathy: (i) *emotional* empathy; (ii) *cognitive* empathy; and (iii) *motivational* empathy. Emotional and motivational empathy can be viscerally or biologically associated by *experiencing* emotions that lead to empathic concern for others, motivating us to offer help. Cognitive empathy, by contrast, is very different, because it allows us to *detect or recognize* the emotional mental states of conspecifics by representing their situation, thereby allowing us to identify salient interpretations of their emotions based on features of their expressions. Doing so can also lead to motivational empathy (offering help), but for quite different reasons, including manipulative ones. In fact, psychopathic patients are very good at cognitive empathy while lacking completely the remorse associated with emotional empathy. If we create AI that is very good at cognitive empathy but which is incapable of emotional empathy, are we creating "psychopathic" and potentially inhuman machines? This paper argues that this is a risk that must be examined and

✉ Carlos Montemayor
cmontema@sfsu.edu

[1] San Francisco State University, San Francisco, CA, USA

[2] University of California, Berkeley, Berkeley, CA, USA

[1] Zaki (2017).

Published online: 26 May 2021

considered very seriously. We provide a more rigorous definition of these kinds of empathy by distinguishing two possible types of simulated empathy in AI applications.

Finally, while a comprehensive analysis of empathy and its relation to care may require a more thorough discussion of the nature of emotions and their relation to intentional action and rationality, the present discussion concerns the specific use of AI in medical care settings in which experienced empathy is essential. Therefore, a circumscribed and qualified approach suffices and in any case, delving into the philosophy of emotion and intentional action would lead us astray from our main goal. We focus on a classification of empathy in AI medical care situations to provide a clear basis for our argument, which is that in these settings, the lack of genuine or experienced empathy presents an in principle problem for the use of AI-simulated empathy, thereby creating risks in patient care that should be avoided for moral and legal reasons.

2 AI and experienced empathy

Having clarified the nature and scope of our argument, let us grant that AI will eventually simulate many forms of thought, skill, and cognition, including emotional reasoning. Like Moravec's paradox, most problems about AI development are conceived as hurdles that scientists will eventually solve. Some authors consider some challenges about moral machines as already addressed in systems that represent the satisfaction of goals without much autonomy, and as part of a larger AI project of developing more morally sensitive and autonomous AI.[2] Although some authors are more optimistic than others, and the range of issues about AI risk and value alignment are multiple,[3] there is consensus that there is no fundamental and insurmountable limitation for the development and application of AI in multiple fields, including morality. Here, we argue that there is at least one a priori limit in the application of AI: empathy, specifically in the field of medical care. Machine interactions in hospitals and similar care-based situations can achieve some neighboring phenomenon, but cannot, in principle, achieve empathy.

Clinical empathy, the use of empathy by nurses, doctors, therapists, etc. is emotion-guided imagining of what a particular moment, or slice of life (a time-slice or segment of one's life), feels like or means to the patient. The moment might be receiving a difficult cancer diagnosis. The slice of life may be going through a difficult post-surgical recovery. The clinician's ability to experience empathy in real time by resonating with a patient's emotional shifts while imagining what the situation is like from inside the patient's perspective (as-if in the subject position of the experience) enables not only more meaningful but more effective medical care for at least three reasons. Getting an accurate history is crucial for medical diagnosis. Did the patient first feel physically exhausted and then, when unable to work for a while, become depressed? Or did she feel exhausted every morning when thinking about depressing aspects of life and then lose her motivation to work? Replicated empirical studies show that patients disclose their histories selectively to physicians according to how emotionally attuned and empathic in real time they perceive their physician to be.[4] Studies show that they do not reveal information at first, but give emotional hints—saying that "my headache just kept coming back" with a lot of anxiety, till they sense that their physician is resonating with the importance of this moment in their story. When they sense this attunement, they reveal information, when they don't, they don't disclose.[5]

Second, effective medical care depends on patients adhering to treatment—the biggest cause of poor results in medicine (for people with access) is that approximately half of medical recommendations including prescriptions are not followed or taken as prescribed. The biggest predictor of adherence to treatment is trust in the physician, and it turns out that a major predictor (in some studies, the biggest predictor) of trust is the patient's perception that the physician is genuinely worried when they talk about something worrisome, that the physician is empathically accompanying them in real time.[6]

Third, a big part of medical care is helping patient's cope with bad news and regain agency to take necessary next steps to help themselves. Oncology patients who sense that their physician was empathizing with them when discussing their cancer diagnosis cope better, seeking treatment and support groups more actively than patients who did not feel so accompanied and had longer periods of confusion and anxiety after receiving difficult information.[7]

Note that all three of these benefits are based on the patient's experiencing the clinician's occurrent empathy in real time. They are not about whether the physician is highly knowledgeable about human behavior and good at predicting what certain people will do; while such knowledge and predictive abilities might have value in healthcare as well, and for that matter, might be exactly what AI systems can provide, this is not what was proven to improve outcomes in medical practice.

[2] Wallach and Allen (2009).

[3] Russell (2019).

[4] Eide et al. (2004); Finset and Ørnes (2017).

[5] Suchman et al. (1997); Halpern (2007).

[6] Roter et al. (1997); Kim et al. (2004); Halpern, "Empathy and Patient–Physician Conflicts, op.cit.

[7] Girgis and Sanson-Fisher (1995); Rosenzweig (2012).

These are virtues of attention that result in genuinely empathic interactions. Successful empathic communication manifests the skills and autonomy of agents—it is not accidentally the case that they empathize with each other; they empathize because they *consciously practice empathic attention*.[8] They know through their shared experiences what is salient in a situation and use their attention to virtuously select only what should be salient from an information set that contains a large amount of information that is not relevant to a specific situation. Moravec's paradox (1988) states that abstract and complex thought is easier to program than the easiest of motor skills. This has proven to be a fundamental challenge in robotics and AI in real, dynamic environments. This is not very surprising. Dexterous behavior has, after all, evolved through millennia of evolution, while the so-called "abstract thought" may be the result of the very recent ability humans have to express themselves linguistically. Here, we argue that the prospects for empathic AI face even greater challenges than Moravec saw for dexterous behavior. While relatives of empathy (compassion, sympathy) might be reproducible in AI systems and have some benefit in clinical settings, when empathy is properly understood it is clear that the capacities that manifest empathy are not capacities that any AI system can manifest.

However, our criticism is consistent with the distinctions mentioned above, and therefore, we are not arguing that AI cannot possibly achieve any cognitive capacities concerning empathy. In particular, AI may be quite good at cognitive empathy. It could accurately represent emotions (although this would be no trivial task) and properly relate situations with desired outcomes. However, because of the incapacity of AI to have emotional or experienced empathy, considerable risks regarding manipulation and unethical behavior need to be avoided, similar to the risks associated with psychopathological patients. Below, we introduce the notion of "empathy*" which could be considered a kind of cognitive empathy that is also insufficient for experienced empathy in medical care settings.

Turkle[9] has presented this problem as a slogan: simulated intelligence may be intelligence, but simulated emotion cannot be emotion. This is especially true of empathy. Humans, like other animals,[10] empathize with each other through the visceral and biologically based emotions our social brains evolved. Human empathy needs these emotional guides but builds upon them a cognitive apparatus that enables imagining not merely how the world "looks" from another person's perspective, but how the specific situation feels to another person inside it. Empathy is *imagining feeling* where the feeling is about something specific, that is the focus of another person's emotion.[11] What enables empathy includes emotional resonance with another's emotions, and the way that this resonance provides a kind of "mood" lighting/stage setting as well as a source of associative linkages that guide the activity of imagining another person's lived experience.[12]

3 Properly guided attention and empathy

It is because empathy has particularity—it is not a thought about how in general people feel, but an experience of imagining how this person feels in this particular situation, that the attention central to empathy is quite focused.[13] This emotion-guided cognition is a manifestation of our mental agency to direct our attention, which involves both conscious and unconscious components. We cannot just choose to resonate with another person, but we can choose to pay attention, try to imagine what they are going through, and in so doing will often find ourselves resonating, which improves our ability to imagine their situation, setting off a virtuous cycle. Importantly, we can have a second-order intention to be receptive that guides us to listen in this way to people over time, developing a habit of receptivity. Critically for this paper, the cognitive activities involved in not just paying attention but cultivating receptiveness to emotional resonance, which then enables imagining how their experience feels to them cannot be reduced to representations of rules and their application to cases.

Consciously empathic attention makes salient the needs of others, and most importantly, allows us to imagine having an experience that is filled in by the details of another person's life, not our own. That is, empathy is not imagining how we would feel if in their situation, but rather, how it feels to be in a situation delineated by their particular predicament.[14]

If genuine care needs to be provided to others in need, then consciously empathic attention must be involved. AI cannot provide consciously empathic attention, because empathy is based on our biological conscious and unconscious mental experiences and our attention capacities to select the most salient and important information for a patient in a situation of care. And this selection is rooted in biologic experiences like resonating with another's emotions. All AI will be able to do is represent the situation of

[8] Montemayor and Haladjian (2015); Haladjian and Montemayor (2016).

[9] Turkle (2005).

[10] de Waal (2019).

[11] Halpern (2001), pp. 85–92.

[12] ibid, pp. 92–95.

[13] Halpern (2014).

[14] Halpern, From Detached Concern to Empathy, op.cit., pp 82–85.

a hypothetical patient and apply it to a concrete data set concerning a specific patient according to some reliable algorithm or rule of inference. Therefore, AI cannot provide empathic attention, genuine care for humans—at best, they can provide emotionally unengaged care through representations and rules about cases.

The situation is much worse than the way Turkle put it. Simulated empathy is not only not really empathy; it is the opposite of empathy, because it is manipulative and misleading to the recipient. It generates responses in the receiver's social brain (i.e., the neural networks responsible for experienced and motivational empathy) that should not be triggered, because there is no biological agent in tune with their emotions at the other end. If so, we are doing a great disservice to those in need of genuine empathy and care by putting them in the "hands" of AI.

The unique elements of empathy that prove elusive to AI are grounded in Ludwig Wittgenstein's way of distinguishing the logical form of first person reports of occurrent psychological states from second or third person reports. For human beings, first person reports such as "My head aches" or "I expect him any moment now" are expressive and thus not in the business of depicting facts, or of being checked against the facts. We do not check our own winces or grimaces against the facts any more than we check "My head aches" against the facts. For the third person report that "Jones is sad" we cite behavioral evidence related to Jones to form the belief that Jones is sad. However, we do not do this for the first person case, we do not say "It seems to me that I am sad, for I am letting my head hang so." The upshot of this asymmetry is that we do not empathize with the psychological states of others through an analogical inference from our own case.[15]

According to Halpern, this is an important point in understanding the nature of empathy. Empathizing with another person is not a case of understanding what you would be experiencing were you in their circumstances, it is not analogically inferred.[16] If AI systems can only form analogically inferred presumptions about the psychological states of others, it cannot empathize with others. We empathize because we are implicitly guided by our own emotional experiences, including experiences of occurrent resonance, rather than as a consequence of an inference. Consider what such inference would have to be like: you observe the other's expressions from a third person perspective, match this to your internal state when expressing yourself in some similar way, then infer that this must be their internal state. Wittgenstein's point is that we do not match our external expressions of emotions to our internal states in this way. Empathy is much more automatic than this, because imputing subjective experiences to another is guided by occurrent resonance or similar emotional experiences rather than a third person process of inferential reasoning. Note for example that when we empathize with another, we have a sense that something is happening to us—we are having an experience—we experience empathy. This corresponds to being moved to imagine something rather than trying hard to imagine something. This is a problem in principle for AI, because empathy does not arise from understanding or observing one's own internal states to glean information that can then be matched to the other's external expressions, and this is the only way AI systems are capable of doing so.

4 Response to objections

Some may object, however, that this "in principle" problem is too unfairly anthropocentric. Why should AI be challenged with obscure notions concerning "what it is like" to experience empathy. Turing[17] called such objections "solipsistic", because there could be no way of verifying or sharing information. Thus, considerations about the qualitative character of subjective experience should not prevent us from designing AI that could enormously help in the field of medicine and other sectors where personal care is needed. Consider how advantageous it would be to have "caring" robots that attend to patients in the context of the current COVID-19 emergency. We already have some evidence that robotic pets that keep elders with mild-to-moderate dementia company, decrease loneliness and improve well-being[18] (real pets and even plants have provided the same benefits).[19] Could seemingly human companions or sites that convey empathy* do even more—help people feel understood as well as less lonely? This might bring the more powerful clinical effects that empathy, as opposed to company, can offer. Maybe, this is not genuinely felt empathy, but we can call it "empathy*." Empathy*, even if not biologically based, is empathic enough.

Again, there are logical, as well as ethical problems with empathy*. Empathy* is not a source of ethical and attentive care, because there is no affective resonance. However, there is a more fundamental problem with the argument in favor of empathy*, which requires the following assumptions: the conscious attention routines involved in empathy can be replaced, through deep-learning and predictive coding, with some kind of attention routine that is equivalent in results. Thus, it will produce behaviors that resemble attention-based

[15] Malcolm (1978).

[16] Halpern (2001), op, cit.

[17] Turing (1950).

[18] Portacolone et al. (2020).

[19] Stanley et al. (2014).

empathy by making salient the needs of patients through inferences or predictions about the patient's behavior.

The problem with this argument in favor of empathy* is that empathy* is not a form of attention either, and therefore, it cannot be based on an organizing helping attention that would enable such cognitive attunement through good inferences. This is also an a priori difficulty. If inference is guided towards helping others, the attention of an agent must be drawn from the premises of the inference to its conclusion. AI systems do not have the required autonomy, agency, and motivations to draw the conclusion from premises in a relevant and non-accidental way. Therefore, AI cannot be genuinely guided towards helping others through their agency, even if their results seem somehow equivalent to those of an empathic agent. AI lacks a helping intention towards another person as the basis of its attentional selection, because it does not have the appropriate motivational and inferential structure.

Mental action, motivations, and attention are hallmarks of solving problems concerning salience in a non-haphazard or accidental way.[20] We are building in too much luck and risk if we depend on AI, because it lacks empathy and empathy*. Since these are the only options, AI systems cannot provide empathic care. Autonomy and agency are the hallmarks of attentive agents—this is how their mental actions are not simply accidentally correct—and only attentive agents can care for others. The risk we create by ignoring this problem is quite serious: we put vulnerable people at the care of systems that, at best, can only accidentally "care" for them.

It might be objected that the argument above simply applies an old and well-known problem for AI in general to the particular case of empathy. It has been argued that conscious experience essentially possesses a phenomenal quality or "qualia." There is something that it is like for a subject to be in a mental state, and this is distinct from the information processing capacities of such states. Since AI cannot undergo qualitative mental states, AI cannot be in states of empathy. One might find this argument too quick and easy to merit much attention.[21]

Let's take a closer look at the role of qualia in empathy-based relations. While an empathizer might have various qualitative mental states, it is not essential that they have them, and attending to one's own mental states can even be counterproductive. The focus of the empathizer is the person being empathized with, it is an essentially *second person* perspective. While, the person receiving empathy may well need to be in qualitative states of mind to feel that they are being empathized with, the provider of empathy needs no qualitative states. Qualitative mental states are essential to the recipient of empathy, but not the giver of empathy. In fact, all that is required is for the interaction to involve emotional beings, which is compatible with judgment-based views of emotions.[22] Thus, our argument does not rely on the premise that AI would have to exhibit qualitative mental states to provide empathy.[23]

Halpern's account of 'affective resonance' may appear to require some qualitative states in both the giver and receiver of empathy. A subtle psychology of empathy will not preclude qualitative states in the empathizer, but will explain the nature of affective resonance without attributing any special qualitative states to the empathizer in the resonance achieved. And in fact, Halpern makes it clear that it is the empathizer's attentively following the flow of the receiver's emotional shifts that is crucial for the recipient of empathy to feel empathized with, and thus experience the benefits described above.[24] Such communicative attunement requires no special qualitative states in the empathizer. In practical terms, the improved clinical outcomes correlated with clinical empathy require that the recipient of empathy feel empathized with, so the role of specific qualitative affective states in the giver should be an open question. Rather than resonance with some unique qualitative feature of the care that empathy provides, the receiver might be said to resonate with the giver's *thoroughly second person, or receiver directed, attentional perspective.* The receiver is perceiving mental shifts in the giver of empathy, and this involves resonating with the thoroughly second person perspective of the giver, not a special and rather specific subjective state the giver is in. AI will have to replicate this to have clinically effective empathy.

The arguments above should suffice to establish the impossibility of empathic AI. However, there is a very different and equally problematic *ethical* obstacle. Even if the genuine article cannot be reproduced by AI, one can still defend an empathic AI program, because coming close to real empathy (e.g., as in empathy*) is good enough to put the attempt on solid footing. After all, human beings often have difficulty manifesting genuine empathy and we can only

[20] Wu (2011).

[21] For discussion, see Block (1980); Dennett (1990).

[22] Nussbaum (2001).

[23] Since our argument does not depend on phenomenal consciousness or qualia per se, but rather on the motivational aspects of empathic engagement through the second person/attentional perspective, it also constitutes an in principle objection to future artificially "conscious" systems that are conceived in terms of a non-organic substrate with a similar cognitive architecture. In particular, Susan Schneider's (2019) "substrate problem" or the difficulty that we don't know if chip-substrates with a complex architecture could be conscious or not, does not affect our argument, because these hypothetical systems would lack the motivational perspective that attentional engagement provides, which is an essentially social and emotional phenomenon.

[24] Halpern, op. cit., (2001). See also Murdoch (1971).

hope that, nonetheless, the effort is good enough to justify the attempt. On one hand, empathy* could be deployed towards people with mild-to-moderate dementia who may believe that it is real and thus experience clinical benefits. This would raise a classic ethical conflict between respect for persons and beneficence. One of us has argued elsewhere in the context of the robot pets that help reduce loneliness that when such results are based on deception, the violation of respect for persons is serious enough to shift to alternative means of decreasing loneliness such as supervised visits with real pets, or using AI to help people connect with real others.[25]

But what about when competent patients know that that their empathy* comes from AI and not a real person? Here, specific uses of something like AI empathy* may be ethically acceptable—for example, an AI responder akin to a smart journal to help people get feedback about how they are feeling (akin to biofeedback) that people can use to reflect. While we would not anticipate the same clinical benefits described above from this kind of intervention (though empirical research is needed to test that), there could be other benefits in self-awareness and perhaps deployment of cognitive behavioral therapy and other tools. Such interventions could enhance recipient's autonomy as well as well-being. The question that arises is why call the AI emotion "reflector" or "advisor" empathy*? Why not call it: AI feedback on emotions? Or something else that describes what it actually is? Most importantly, while getting feedback on one's own emotional states by a kind of smart journal may have real value, this is in fact not what has been shown to be deeply therapeutic about empathy when people are ill, suffering, grieving and afraid.[26] As described above, the therapeutic benefits of empathy, as opposed to journaling, meditating, and other salutary activities, have all been shown to depend specifically on the recipient's feeling that someone else is paying attention to them, curious to know more about them and worrying about them—in short caring about them. It misuses the term "empathy" to apply it to other important sources of understanding one's emotions. Given that health systems are under financial pressure to fund less human hours to provide the empathic care that patients need, we are concerned that labeling these AI programs "empathy" is a way to mislead the public (even if they know it's AI listening to them) and ultimately, to degrade expectations for human empathy in healthcare. Thus, we argue that it is ethically inappropriately to label even potentially very sophisticated AI emotional reflecting/advising as empathy.

In conclusion, empathic AI is thus either impossible or unethical. Impossible because of the lack of genuine empathy on the part of the AI. Unethical because of the risks empathy* produces not only for deception in specific instances, but for reducing and undermining the meaning and expectations for real empathy. The nature of the attention to another human being central to empathy is not exclusively a matter of rule following or good consequences represented under some algorithmic structure. Ignoring this could erode the normative expectation that when a person is suffering, they ought to elicit real human empathy, an expectation which is central to the morality of medical care as well as other relational practices.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Block N (1980) Troubles with functionalism. In: Block N (ed) Readings in the philosophy of psychology, vol 1. Harvard University Press, Cambridge, MA, pp 268–305

de Waal F (2019) Mama's last hug: animal emotions and what they tell us about ourselves. W. W. Norton & Company, New York

Dennett D (1990) Quining qualia. In: Lycan W (ed) Mind and cognition. Blackwell, Oxford, pp 519–548

Eide H, Frankel R, Haaversen AC, Vaupel KA, Graugaard PK, Finset A (2004) Listening for feelings: identifying and coding empathic and potential empathic opportunities in medical dialogues. Patient Educ Couns 54(3):291–297

Finset A, Ørnes K (2017) Empathy in the clinician-patient relationship: the role of reciprocal adjustments and processes of synchrony. J Patient Exp 4(2):64–68

Girgis A, Sanson-Fisher RW (1995) Breaking bad news: consensus guidelines for medical practitioners. J Clin Oncol 13:2449–2456

Haladjian HH, Montemayor C (2016) Artificial consciousness and the consciousness-attention dissociation. Conscious Cogn 45:210–225

Halpern J (2001) From detached concern to empathy: humanizing medical practice. Oxford University Press, Oxford

Halpern J (2007) Empathy and patient-physician conflicts. J Gen Intern Med 22(5):696–700

Halpern J (2014) From idealized clinical empathy to empathic communication in medical care. Med Health Care Philos 17(2):301–311

Halpern J (2017) The therapeutic effects of empathy in healthcare. In: Emotion researcher, International Society for Research on Emotion, pp 66–71

Kim SS, Kaplowitz S, Johnston MV (2004) The effects of physician empathy on patient satisfaction and compliance. Eval Health Prof 27(3):237–251

Malcolm N (1978) Wittgenstein's conception of first person psychological sentences as expressions. Philos Exch Brockport NY 2(4):59–72

Montemayor C, Haladjian HH (2015) Consciousness, attention, and conscious attention. MIT Press, Cambridge

Murdoch I (1971) The sovereignty of good. Routledge, London

Nussbaum MC (2001) Upheavals of thought: the intelligence of emotions. Cambridge University Press, Cambridge

[25] Portacolone et al. (2020), op. cit.

[26] Halpern (2017).

Portacolone E, Halpern J, Luxenberg J, Harrison K, Covinsky K (2020) Ethical issues raised by the introduction of artificial companions to older adults with cognitive impairment: a call for interdisciplinary collaborations. J Alzheimer's Dis 76:445–455

Rosenzweig MQ (2012) Breaking bad news: a guide for effective and empathetic communication. Nurse Pract 37(2):1–4

Roter D, Hall J, Merisca R, Nordstrom B, Cretin D, Svarstad B (1997) Effectiveness of interventions to improve patient compliance: a metaanalysis. Med Care 36(8):1138–1161

Russell S (2019) Human compatible: artificial intelligence and the problem of control. Random House, Viking

Schneider S (2019) Artificial you: AI and the future of your mind. Princeton University Press, Princeton

Stanley IH, Conwell Y, Bowen C, Van Orden K (2014) Pet ownership may attenuate loneliness among older adult primary care patients who live alone. Aging Ment Health 18(3):394–399

Suchman A, Markakis K, Beckman H, Frankel R (1997) A model of empathic communication in the medical interview. JAMA 277(8):678–682

Turing AM (1950) Computing machinery and intelligence. Mind 59:433–460

Turkle S (2005) The second self: computers and the human spirit, 20th anniversary. MIT Press, Cambridge

Wallach W, Allen C (2009) Moral machines: teaching robots right from wrong. Oxford University Press, Oxford

Wu W (2011) Attention as selection for action. In: Mole C, Smithies D, Wu W (eds) Attention: philosophical and psychological essays, pp 97–116

Zaki J (2017) Moving beyond stereotypes of empathy. Trends Cogn Sci 21(2):59–60