The AI effect today: On denials of AI's intelligence

This article analyses the AI effect—briefly, the denial of AI's intelligence—in three steps. First, it explains the tendency to deny that AI is truly intelligent. Due to experiencing its own intelligence as self-awareness, the human mind cannot help but to tie any intelligence to self-awareness. Thus, for AI to be regarded as intelligent, it would need to have self-awareness, and because it does not, humans tend to deny that AI is intelligent in any way. Second, the article presents the AI effect as an ethical issue insofar as the effect denies that AI is indeed intelligent, albeit in ways different from humans. Third, it analyses a characteristic case of the AI effect—insistence that AI is not intelligent because it has no sentience—by examining the retroactive denigration of AlphaGo.

Keywords: *AI, AI ethics, AI effect, intelligence, awareness, AlphaGo*

Acknowledgments

Research leading to this article was done with the support of the research program group Problems of Autonomy and Identities in the Time of Globalization (P6-0194), funded by the Slovenian Research and Innovation Agency.

Author Information

Primož Krašovec, University of Ljubljana
https://orcid.org/0000-0002-9013-0305

How to cite this article:

Krašovec, Primož. "The AI effect today: On denials of AI's intelligence ".
Információs Társadalom XXV, no. 2 (2025): 45–58.
https://dx.doi.org/10.22503/inftars.XXV.2025.2.3

1. Introduction

AI is neither artificial nor intelligent. – Crawford (2021, 8)

As evidenced by that opening quotation and representative examples analysed in the section on AlphaGo at the end, the consensus within the critical literature on AI seems to be that the new wave of AI based on deep learning (DL) is not intelligent at all and that the hype accompanying it is at least naive if not outright "corporate propaganda" (Pasquinelli 2023, 15). This article's contention, however, is that the question of AI's intelligence is far from settled and that claims that AI is really artificial unintelligence (Broussard 2018) reveal less about AI and more about, to paraphrase Karl Marx, the poverty of AI critique, as well as people's deep-seated anthropocentric prejudices that deny the very possibility of other-than-human intelligence(s).

In what follows, the article first situates current denials of AI's intelligence within the history of the so-called AI effect as a tendency to retroactively disqualify anything that AI shows itself capable of as being unintelligent. The AI effect is counterposed to the ELIZA effect, an enthusiastic attitude toward AI that instead embraces AI as in fact intelligent, and Graziano's (2013, 2019) attention schema theory to explain both as expressions of "everyday animism," polarized by the interplay between perceptual illusion (i.e., knowing that machines are not self-aware but acting as if they are) and cognitive illusion—that is, believing that machines are in fact self-aware.

Second, the article presents the AI effect as an ethical issue that prevents the actual understanding of AI as other intelligence and thus precludes the formation of any genuine ethical relation with it. Calls for explainable AI, commonly abbreviated to "XAI," are interpreted as being decidedly one-sided because they place the whole burden of explanation on AI and fail to recognize that, to understand AI, the very mode of human understanding would have to change (Fazi 2021).

Third and last, the article analyses examples of the AI effect taken from recent critical literature on AI and organised around a characteristic contemporary case of the denial of AI's intelligence: the insistence that AI is not intelligent because it has no sentience in retroactive denigrations of AlphaGo's achievements since 2016. As shown, the case also makes visible an aversion to understanding intelligence as a material, computational process.

2. The AI effect and the ELIZA effect

Despite intensifying after the initial hype surrounding the ascension of DL AI dissipated in the late 2010s, denials of AI's intelligence are not new. In her seminal study of early AI, *Machines Who Think*, Pamela McCorduck (2004, 204) noted early on that:

> It's part of the history of the field of artificial intelligence that every time somebody figured out how to make a computer do something—play good checkers, solve simple but relatively informal problems—there was a chorus of critics to say, but that's not thinking.

As McCorduck's observations makes clear, the denial of AI's intelligence always comes *ex post* and never in a sense that something does not require genuine intelligence before AI does it. What McCorduck suggests is that when it comes to AI, the usual order of proof is reversed; instead of AI's failing to meet *ex ante* criteria of what constitutes real or genuine intelligence, those criteria seem to be whatever AI cannot do (Dreyfus 1992) or, at the time, is thought of incapable of achieving. In this article, following Hainlein and Kaplan (2019, 2), the phenomenon is called "the AI effect," defined as a denial of AI's intelligence even when it exhibits intelligent behaviour.

At first glance, the AI effect seems to be the opposite of the ELIZA effect, a widely documented and discussed attitude toward early AI. The ELIZA effect takes its name from the eponymous computer program released in 1966 that could perform simple psychotherapeutic conversations with human users (Weizenbaum 1966). ELIZA was so popular and the users' experience of it so immersive that the tendency of human users to develop social and emotional attachments to computer programs and to project human traits such as subjective feelings and self-awareness onto them is still called "the ELIZA effect" (Natale 2021, 50–67). The ELIZA effect describes a situation in which the intelligence of AI not only goes unquestioned but is also inflated to the point that, at least in the imagination of its users, it involves sentience and emotions. The ELIZA effect is thus an attribution of intelligence to AI even when AI does not exhibit intelligent behaviour.

Immediately salient is an important feature that both effects have in common: positing a common sense (self-)understanding of human intelligence as a measure of any intelligence. Thus, the difference between them is that, in the case of the AI effect, AI is chastised on the grounds that it has no humanlike intelligence, whereas in the ELIZA effect, it is cherished precisely because it is perceived as having humanlike intelligence. The difference between the ELIZA and AI effects is therefore not a difference of kind but of degree; they are the extreme points of the same spectrum of everyday attitudes that use an intuitive, experiential understanding of human intelligence as the measure of any intelligence. From another angle, what is missing in both cases is any conception of intelligence that can be nonhuman.

Historically, the oscillation between both effects was not random but conditioned by the actual performance levels of AI. When its performance was weak, the AI effect was relatively rare, whereas the ELIZA effect was so prevalent that Weizenbaum (1976), the inventor of ELIZA, felt compelled to dedicate a whole book to its critique. By contrast, the current prevalence of the AI effect can be attributed to the explosive rise in AI's levels of performance, whereas rare examples of the ELIZA effect—such as Google engineer Blake Lemoine's claim that the language model LaMDA is sentient—are met with unanimous scorn and derision (Bratton and Agüera y Arcas 2022). Thus, there seems to be an inverse correlation between the actual intelligence of AI and its social and cultural recognition.

Explaining both the ELIZA and AI effects requires returning to what they have in common: an intuitive understanding of intelligence as inseparable from self-awareness. According to Graziano (2013, 69), human (self-)awareness is a computed model of attention. That model is essentially pragmatic and focused on being efficient in

predicting one's behaviour and the behaviour of others—in short, its function is primarily social. For that reason, and given the demands of computational economy, it is also not exhaustive. Awareness is not a complete account of material computational processes that generate human intelligence but instead a simplified, reductive model that abstracts from anything too complex and material and feels like an immaterial presence within our heads (Graziano 2019, 70). Since human intelligence is inseparable from the experience of subjective awareness, it is intuitively inconceivable that intelligence could exist without it.

Early AI such as ELIZA affected its users in ways similar to puppet shows. A basic condition that allows people to become immersed in puppet shows and to enjoy them is a human tendency to project awareness everywhere (Graziano 2019, 6–10). People enjoy puppetry precisely because they suspend their disbelief that puppets are neither alive nor self-aware. It is a form of controlled illusion; audiences assign awareness to the puppets on the level of perception but on the cognitive level know that puppets' awareness is not real (Graziano 2013, 205–6). However, the attitude toward AI shifts once it is no longer limited to preprogrammed schemes and begins to exhibit signs of actual intelligence instead.

This article's working definition of *intelligence* as an ability to make autonomous decisions, based on sensing the environment, that are efficient (i.e., not random) in relation to it comes from recent theories in neuroscience (Damasio, 2021; Ledoux 2019, 47–77) and from Land's (2019) understanding of intelligence as cognitive autonomy in his essay on AlphaGo Zero. The crucial term is (cognitive) autonomy, not (self-)awareness, which allows positing AI as being truly intelligent regardless of whether it is self-aware. Early symbolic AI was unintelligent because it consisted of program execution and worked by applying preset rules. In the case of DL AI, by contrast, all the above conditions for actual intelligence are met. The environment is the data that AI is exposed to and trained with; sensing is the relation(s) that it forms with said data; and intelligent behaviour is autonomous decision-making that it performs in response to data—for example, the way that it parses and clusters data as well as extracts features from and analyses patterns in said data. DL AI not only surpasses traditional programming but was developed precisely for situations in which programming proved to be impossible (Mendon-Plasek 2021, 44) and for problems that it could not solve (Burrell 2016, 6), especially in cases that required actual (machine) intelligence and for which rote computation was insufficient.

It is precisely the encounter with real machine intelligence—that is, what distinguishes today's AI from early AI and puppets—that triggers a specific reaction whose function is to prevent the transition from perceptual illusion toward cognitive illusion: an actual belief that AI possesses humanlike self-awareness. The reaction manifests as a dismissive attitude characteristic of the AI effect, instead of an enthusiastic, playful one characteristic of the ELIZA effect.

The lack of real intelligence in early ELIZA-like AI, which consisted of relatively simple algorithms based on a few crude preprogrammed responses to user input, enabled the ELIZA effect. Immersive user experience with simple, not truly intelligent chatbots assumed a form of pattern completion by tapping into and stimulating users' imagination insofar as the imagined psychic depth and real understanding on

the side of the chatbots was entirely the user's construction (Natale 2021, 107–25), much in the same way that people pattern-complete the actions of puppets on stage or animated characters on a TV screen. Viewers are shown only a hint—anthropomorphous bodies moving and human voices talking—and their imaginations fill in the rest.

In a reverse situation, the actual intelligence of the current DL AI disables users' immersion because it presents a risk that perceptual illusion will become a cognitive one. Users attempt to convince themselves that AI is not truly intelligent, for if it were, then the risk of cognitive illusion would not exist. Due to experiencing its own intelligence as self-awareness, the human mind cannot help but to automatically tie intelligence to self-awareness. However, self-awareness is the way that human intelligence works and is not binding for any other forms of intelligence. Current DL AI is a striking example of an actual intelligence that works without any recourse to self-awareness or subjective experience (Browning and LeCun 2022).

For human users, the situation is a precarious one. People cannot help but assign awareness to machines that they interact with (i.e., everyday animism), because assigning awareness is a foundational feature of human sociality: people interact with other people only if they assume that human others have minds and self-awareness as they do. However, human sociality does not stop at other humans, for people routinely assign awareness to toys, animals, characters in books and movies, and even machines as well. That tendency works flawlessly if humans are certain that the illusion that machines are self-aware is only a game, a playful as-if immersion, which is possible as long as machines are not really intelligent. However, once machines begin exhibiting actual intelligence, the routine assignment of awareness runs into trouble and breaks down.

When interacting with truly intelligent machines, it becomes far more challenging to keep the illusion of their awareness under control. The routine assignment of awareness works flawlessly in the case of unintelligent computer programs; users assign awareness to them but are at once certain that because they are not intelligent that they cannot possibly be aware. The illusion of machine awareness thus remains at the level of perceptual illusion. The reverse is the case when users confront AI that exhibits real intelligence. If they grant it intelligence, they risk triggering an automatic correspondence between intelligence and self-awareness and consequently a transition from merely a perceptual into a full cognitive illusion of machine awareness. At the same time, however, users can be quite certain that AI has no real self-awareness; thus, users confront cognitive dissonance. In that case, because AI is intelligent, it has to be self-aware, but it obviously cannot be self-aware. The solution is to deny that AI is intelligent at all, which is precisely the mechanism of the AI effect, as explored later in the section on AlphaGo.

3. The AI effect as an ethical issue

The question of AI's intelligence, because it involves human relations with others and ourselves, is not only an epistemological question but also, though not usually

presented as such, an ethical question in a Foucauldian sense—that is, not an application of transcendent morality but an immanent examination of how people relate to themselves and others (Foucault 1997/1994). This article attempts to show that any kind of genuine ethical relation to AI can be established only once humans resolve the tendency to posit their common sense understanding of human intelligence—how we think that we think—as a measure of any intelligence whatsoever (Bratton 2015, 74) and any deviation from it as a sign of inadequacy and a lack of intelligence on the side of AI. One of the problems of the AI effect is precisely that it uses human intelligence as a measure of any intelligence, which makes it not only an epistemological but an ethical issue as well.

The examination of human intelligence's often bigoted relationship to other, especially machine, intelligences constitutes this article's contribution to current discussions on AI ethics. Those discussion chiefly focus on the machine side of that relationship in the sense that they investigate how machines could be made to act in more ethical ways (Pasquale 2020) as well as how they can become more transparent to human understanding and accountable to human judgment (Diakopoulos 2020). In both cases, instead of implying a truly ethical relationship, the word *ethics* seems to be an euphemism for "machine submission" and "human control": that an often erratic, unpredictable machine behaviour has to be brought under strict human supervision and its opaque inner workings made readable by humans.

Because the explicability of increasingly complex DL AI is currently one of the most-discussed issues in conversations about AI ethics, this critique may represent a contribution to the very conditions of making AI understandable by highlighting certain weaknesses and blind spots in common sense assumptions about what would understanding AI would mean. In particular, humans tend to judge AI by the standards of human intelligence. As a consequence, they fail to account for differences between how human and artificial intelligences work. If people instead suspended their anthropocentrism, then a better understanding of AI would be possible but also reveal that the fault for AI's current opacity lies not entirely with AI—with its characteristic black box quality (Pasquale 2015)—but also with humans' flawed attempts to understand it thus far. In other words, significantly more effort would be required to cultivate a greater human understanding of AI because understanding AI is made difficult not only by its complexity but also by the lack of human readiness to understand it.

Although providing a complete theoretical framework for understanding AI is beyond this article's scope, it nevertheless attempts to provide at least one necessary starting point: a relaxation of the assumption that people with their current understanding can understand AI. It also requires a disclosure of certain problematic habits of thought that are ingrained in human attitudes toward AI and function as epistemological obstacles (Bachelard 2002/1938, 24–32) to understanding AI. Admitting that AI is indeed intelligent, just in a different way, might allow for a different, less judgmental, and thus more understanding approach to AI ethics (Kaluža 2023).

What makes the AI effect problematic from an ethical perspective is the tendency to extend the question of whether AI is self-aware to a negation of its intelligence. Although it is obvious that AI has no self-awareness, the question of whether it

exhibits actual intelligence is an altogether different, less unequivocal one (Agüera y Arcas 2022; Browning 2020). Off-hand claims that the intelligence of AI is not real intelligence are not just *non sequiturs*—that is, if something is not self-aware, then it does not logically follow that it is also not intelligent—but also contain unwarranted dismissiveness toward AI beyond what is necessary to establish that it is not self-aware. To show that AI is not self-aware does not require resorting to dismissiveness, and said dismissiveness therefore reveals something that says as much, if not more, about the faults in the human understanding of any intelligence, including our own, than it does about AI.

The anthropocentric (mis)understanding of (artificial) intelligence does not come first but is instead a derived resolution of the cognitive dissonance mentioned above, such that AI is disqualified from intelligence proper as collateral damage of the human need to keep the illusion of machine awareness under control. Positing human intelligence as being the only real form of intelligence comes afterward, because machines cannot be intelligent, for real intelligence can only be human intelligence. In that way, the intuitive perception of human intelligence is established as a measure of any intelligence not as a consequence of some prejudice against machine intelligence but as a pragmatic solution of a problem posed by the risk of a perceptual illusion developing into a cognitive one. As a consequence, the very possibility of any real machine intelligence is excluded a *priori*, which leads to the development of the AI effect—to wit, that anything AI does is not intelligence and intelligence is whatever AI cannot do. Because the standard of intelligence is posited in exclusively anthropocentric terms, the only way to evaluate AI is to measure it against an intuitive perception of human intelligence, namely by perceiving it as a lesser, ever-inadequate version of human intelligence that can never really live up to the original. Although "it would be wiser to separate 'intelligence' from 'consciousness' and 'sentience'" (Agüera y Arcas and Norvig 2023), such lenience toward machine intelligence is challenging to achieve due to the pervasiveness of the AI effect.

In contrast to other work on AI explainability (Larsson and Heintz 2020), this article contends not that AI is an opaque black box (only) due to its complexity but (also) due to the unwillingness and inability to understand it on the human side owing to the AI effect. AI is understood as the inadequate mimicking of human intelligence due to the pervasive human inability and refusal to even acknowledge the possibility of the existence of autonomous machine intelligence, much less to understand it as such. For example, AI critic Harry Collins (2018) has called artificial intelligence "artifictional intelligence," meaning an apparent, counterfeit intelligence because it is not the same as human intelligence (16–18). Another example from another prominent AI critic is that "neither deductive algorithms nor statistical techniques excel in mimicking human intelligence" (Pasquinelli 2023, 232). In such claims, the unquestioned assumption is that AI mimics human intelligence, and it is dismissed as falling short of it. Along the horizon of intelligibility constituted by the AI effect, any difference between human and machine intelligence is immediately recast as a lack or insufficiency with respect to the human norm. "Outing" AI (Bratton 2015) would thus mean perceiving it as an actual but different intelligence instead of it being perceived as a lesser version of human intelligence.

The ethical question is not whether humans can understand AI with their current thinking but instead how humans relate to AI and whether that relationship involves acknowledging AI's autonomous intelligence that might not exist (exclusively) for our understanding and functional use. In other words, although discussions on XAI have focused almost exclusively on the AI side, ethical issues exist on the human side as well. Consequently, human relationships with AI cannot be ethical as long as the AI effect is at play. Instead of making AI more understandable, the ethical task ahead may be making humans more understanding of AI not in the sense of completely comprehending AI, which may be impossible anyway, but its acceptance as another form of intelligence that may remain at least partly secretive (Amoore 2020, 133–53).

XAI has little to do with ethics but a lot to do with power, namely attempts to curb AI and make it subservient or "aligned": "The current wave of Artificial Intelligence Ethics Guidelines can be understood as desperate attempts to achieve social control over a technology that appears to be as autonomous as no other" (Héder 2021, 120). In that sense, XAI is not so much about genuine understanding as it is about surveillance (Héder 2020). At the same time, cultivating an ethical relationship with AI would also involve adopting a more understanding attitude toward it, even if it means dispensing with the obsessive urge to scrutinise and control everything about it. Thus, "Disobedient AI is not necessarily a threat" but rather "an opportunity to shape new human–technology relations that are not based on domination" (Hosseinpour 2020, 50).

Such an attitude would not substantially differ from how we ethically relate to other humans, because other humans are, in a way, black boxes as well. As in AI's case, humans lack access to the actual processes of intelligence in other humans and can only discern their surface traces. Even so, that does not preclude people from (sometimes) forming meaningful ethical relations with them. The only difference is that humans think that they have access to the inner workings of their own intelligence when they in fact do not. The explanation that people demand from AI is doubly anthropocentric: that AI makes itself explainable not only in human terms but in terms of a deceptive intuitive (mis)understanding of their own intelligence. The first step in making humans more AI understanding would subsequently mean humans' reengagement with their own intelligence.

4. AI Effect after AlphaGo's 2016 victory

The turn toward human intelligence and its common sense (mis)understandings is not a detour but a necessary step to explain the AI effect as a denial of AI's intelligence because the AI effect has the same structure as a denial of human intelligence. Of course, the AI effect does not mean that humans tend to deny that they are intelligent in the same way that they routinely deny that AI is. On the contrary, humans tend to deny that AI is intelligent while not only affirming that humans are but also positing human intelligence as a norm for AI to attain. However, what is the same in both the affirmation of human intelligence and the denial of AI's intelligence is a common sense misperception of how intelligence works.

Because humans are not and cannot be aware of the material, computational processes of human intelligence, their experiential conception of it involves reducing it to self-awareness as something immaterial in the human brain. Such a reductive experience of human intelligence as immaterial awareness prompts an aversion to any conception of intelligence as an (exclusively) material, computational process without awareness. That aversion, in turn, is involved in the AI affect—a conviction that non-aware intelligence is not real intelligence—and in dismissive attitudes toward AI that prevent the formation of any meaningful ethical relations with it.

An exemplary case of the contemporary AI effect was the denigration of AI that played the game go in 2016. Immediately before the landmark 2016 AlphaGo victory against Lee Sedol, computer scientists thought that such an achievement was at least a decade away due to the difficulties posed by the immense complexity of playing go that cannot be resolved via brute force computation (Levinovitz 2014). In other words, for AI to play go, it would need something akin to artistic intuition, which was precisely the reason why an AI victory against expert human go players before 2016 was conceived to be not only difficult but impossible (Du Sautoy 2019, 18–22, 25, 30–43). Before 2016, the AI effect was thus expressed as follows: Because playing go requires creative intuition, it is by definition impossible for machines to win given that they may be capable of lowly computation but never of higher intelligence functions such as creative intuition, as if intelligence is whatever machines cannot do. However, after the 2016 AI go triumph, the AI effect remained in place, only that instead of continuing to be the epitome of human creative intuition, the game go itself was instead demoted to mere math, which can be solved by computational processes done by machines. Again, the thinking was that whatever AI does is not intelligence.

For example, in his 2023 piece "The Stupidity of AI" in *The Guardia*n, influential AI critic Bridle (2023) grouped playing go together with playing chess as an example of a "narrow domain of puzzles" that he contrasted with "imagination and creativity." Thus, playing go after 2016 was no longer a puzzle of imagination and creativity but a simple puzzle not necessarily requiring imagination or creativity. In another example, Gray and Suri (2019) highlighted AlphaGo as an example of how "AI is simply not as smart as people hope or fear" (30). To preclude readers from being overly impressed by the victories of AlphaGo and later AlphaGo Zero, the authors remind readers that "the rules of go are fixed and fully formalized and it is played in a closed environment" (31). Those claims completely bypass the crucial point of the immense complexity of playing go due to the sheer number of possible moves within such a closed environment. As any game, go is based on rules, though their execution, due to the game's complexity, is far from straightforward and involves a great deal of imagination and creativity from humans, which is precisely why go used to be regarded as a game that requires more or higher intelligence than the simple execution of rules and why it required deep learning AI to finally master it. Similarly to Bridle and consistent with the AI effect, Gray and Suri (cherry-)picked a dimension of go that makes it similar to any other game while quietly sidelining the dimensions that set it apart and make it require special intelligence. They thus conclude their brief dismissal of go and AI playing go as follows: "Life is more complicated than a game

of go" (32). Although no doubt true, it hardly proves that winning go is not a display of real intelligence. Because AI's victories in go are undeniable, the somewhat predictable human move in response has been to deny the intelligence required to play go and, by extension, AI's intelligence.

Another influential AI critic, Crawford (2021), has countered the claims that AlphaGo exhibits "some kind of otherworldly intelligence" with an alternative explanation:

> AI game engines are designed to play millions of games, run statistical analyses to optimize for winning outcomes, and then play millions more. These programs produce surprising moves uncommon in human games for a straightforward reason: they can play and analyze far more games at a far greater speed than any human can. This is not magic; it is statistical analysis at scale (205).

Albeit factual—AlphaGo did indeed play millions of games and ran statistical analyses—the reiteration of the "it's just statistics" theme still fails to explain why running millions of games and running statistical analyses does not constitute intelligence. The idea that mere statistics cannot ever constitute real intelligence is only self-evident if it is assumed that intelligence cannot be a material computational process as a matter of principle. Crawford reduces AI's intelligence to a technical operation and suggests that if something is technical, then it cannot be truly intelligent. That claim is as unusual as it would be to maintain that because Lee Sedol's brain fired billions of neurons in complex patterns during a game of go, it is just neurochemical activity and thus not real intelligence. The fact that something involves a material computational process is not in itself proof that it is not intelligent. In short, go went from being *the* game, one requiring high-level intelligence before 2016, to being just *a* game and, as such, irrelevant for intelligence after 2016.

In perhaps the clearest example of an aversion to intelligence as a computational process—AI is inferior because it consists of "mere computation" without sentience—yet another influential AI critic Broussard (2018, 36) has argued that "AlphaGo is not an intelligent machine, however. It has no consciousness." That claim implies, however, that real intelligence requires humanlike consciousness. Broussard adds quite emphatically that computers are not and cannot be intelligent because they only execute orders and have no sentience or soul (11–12) and thus completely disregards the two key questions of current AI: whether DL is really a mere execution of orders and whether there can be an intelligence without awareness. Broussard's disavowal of those questions is a *non sequitur;* if something is not sentient, then it does not (necessarily) follow that it is merely executing orders. Although logically false, such an assertion still makes perfect sense in the context of the AI effect; if something merely executes orders, then it is not intelligent, and if it has no awareness, then it must be unintelligent (i.e., the inviolable axiom of the AI effect). Thus, AI is reduced to merely executing orders.

The case study presented here perfectly illustrates the way in which the AI effect works. It corresponds to the logic of moving goalposts, as already observed by

McCorduck (2004). In short, intelligence becomes whatever machines cannot do, and likewise, whatever machines can do does not constitute intelligence.

5. Conclusion

Many theories view AI as real intelligence that is merely different from human intelligence (Agüera y Arcas and Norvig 2023; Bratton and Agüera y Arcas 2023; Ernst 2021; Fazi 2021). Denials of AI's intelligence have also been critically investigated as a way of coping with a Copernican trauma that involves the decentering of the image of human intelligence as the norm and endgame on intelligence, triggered by the development of AI (Bratton 2024). This article's contribution to the discussion is an attempt to connect denials of AI's intelligence to certain deeply ingrained common sense assumptions that link intelligence to humanlike self-awareness. Moreover, it presents such denials as an ethical issue, with the only precursor known to be Amoore (2020), and provides a concrete case study of the AI effect in the case of AlphaGo after 2016. Considering all the above, it seems that denials of AI's intelligence, ubiquitous in current critical literature on AI, are not so much expressions of a refined intellectual reflection but rather of common sense epistemological obstacles.

Despite being a default mode of how humans relate to AI, the AI effect is not insurmountable. It is a spontaneous common sense reaction that can be rectified upon reflection in ways similar to negations of animal intelligence (Keim 2024), non-White intelligence (Allan 2002), and female intelligence (George 1915) in the past. At the same time, because it is a common sense epistemological obstacle deeply ingrained in everyday thinking, overcoming the AI effect will probably never be a *fait accompli* but a continuous process that is made ever more urgent as the pace of AI's development accelerates. Overcoming the AI effect is perhaps *the* ethical question regarding AI.

Along with allowing a genuine ethical relationship with AI as an actual, although different form of intelligence, to develop, overcoming the AI effect would also present an opportunity to learn more about ourselves. Such learning would need to involve accepting that what we imagine as our intelligence that supposedly makes us superior to any other forms of intelligence is in reality merely a self-misunderstanding—an insight that could provide an antidote to currently prevailing anthropocentric conceits in relation to AI.

References

Allan, Alexander. *Race in Mind: Race, IQ, and Other Racisms.* London: Palgrave Macmillan, 2002.

Amoore, Louise. *Cloud Ethics: Algorithms and Attributes of Ourselves and Others.* Durham, NC: Duke University Press, 2020.

Agüera y Arcas, Blaise. "Do Large Language Models Understand Us?" *Daedalus* 15, no. 2 (2022): 183–97. https://doi.org/10.1162/daed_a_01909

Agüera y Arcas, Blaise, and Peter Norvig. “Artificial General Intelligence Is Already Here.” *Noema,* October 10, 2023. https://www.noemamag.com/artificial-general-intelligence-is-already-here/

Bachelard, Gaston. *The Formation of the Scientific Mind: A Contribution to a Psychoanalysis of Objective Knowledge.* Manchester: Clinamen, 2002.

Bratton, Benjamin. “Outing Artificial Intelligence: Reckoning with Turing Tests.” In *Alleys of Your Mind: Augmented Intelligence and Its Traumas,* edited by Matteo Pasquinelli, 69–80. Lüneburg: Meson, 2015.

Bratton, Benjamin, and Blaise Agüera y Arcas. “The Model is the Message.” Noema, July 12, 2022. https://www.noemamag.com/the-model-is-the-message/

Bratton, Benjamin. “The Five Stages of AI Grief.” *Noema,* June 20, 2024. https://www.noemamag.com/the-five-stages-of-ai-grief/

Bridle, James. “The Stupidity of AI.” *The Guardian,* March 16, 2023. https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt

Broussard, Meredith. *Artificial Unintelligence: How Computers Misunderstand the World.* Cambridge, MA: MIT Press, 2018.

Browning, Jacob. “Learning Without Thinking.” Noema, December 29, 2020. https://www.noemamag.com/learning-without-thinking/

Browning, Jacob, and Yann LeCun. “What AI Can Tell Us About Intelligence.” *Noema,* June 16, 2022. https://www.noemamag.com/what-ai-can-tell-us-about-intelligence/

Burrell, Jenna. “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms.” *Big Data and Society* 3, no. 1 (2016). https://doi.org/10.1177/2053951715622512

Collins, Harry. *Artifictional Intelligence: Against Humanity’s Surrender to Computers.* Cambridge, MA: Polity, 2018.

Crawford, Kate. *Atlas of AI: Power, Politics and the Planetary costs of Artificial Intelligence.* New Haven: Yale University Press, 2021.

Damasio, Antonio. *Feeling and Knowing: Making Minds Conscious*. New York: Pantheon, 2021.

Diakopoulos, Nicholas. “Transparency.” In *The Oxford Handbook of Ethics of AI,* edited by Markus Dubber, Frank Pasquale, and Sunit Das, 197–213. Oxford: Oxford University Press, 2020.

Dreyfus, Hubert. *What Computers Still Can’t Do: A Critique of Artificial Reason.* Cambridge, MA: The MIT Press, 1992.

Du Sautoy, Marcus. *The Creativity Code: Art and Innovation in the Age of AI.* Cambridge, MA: Belknap, 2019.

Ernst, Wolfgang. *Technológos in Being: Radical Media Archaeology and the Computational Machine.* London: Bloomsbury, 2021.

Fazi, Beatrice. “Beyond Human: Deep Learning, Explainability and Representation.” *Theory, Culture and Society* 38, no. 7/8 (2021): 55–77. https://doi.org/10.1177/0263276420966386

Foucault, Michel. *Ethics: Subjectivity and Truth.* Translated by Robert Hurley and others. New York: The New Press, 1997/1994.

George, W. L. "Notes on the Intelligence of Woman." The Atlantic (December 1915). https://www.theatlantic.com/magazine/archive/1915/12/notes-on-the-intelligence-of-woman/304038/

Gray, Mary, and Siddharth Suri. Ghost Work: *How to Stop Silicon Valley from Building a New Global Underclass.* Boston: Houghton Mifflin Harcourt, 2019.

Graziano, Michael. *Consciousness and the Social Brain.* Oxford: Oxford University Press, 2013.

Graziano, Michael. *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. New York: W. W. Norton, 2019.

Hainlein, Michael, and Andreas Kaplan. "A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence." *California Management Review* 61, no. 4 (2019): 5–14. https://doi.org/10.1177/0008125619864925

Héder, Mihály. "A Criticism of AI Ethics Guidelines." *Információs Társadalom* 20, no. 4 (2020): 57–73. https://dx.doi.org/10.22503/inftars.XX.2020.4.5

Héder, Mihály. "AI and the Resurrection of Technological Determinism." *Információs Társadalom* 21, no. 2 (2021): 119–30. https://dx.doi.org/10.22503/inftars.XXI.2021.2.8

Hosseinpour, Hesam. "Disobedience of AI: Threat or Promise." *Információs Társadalom* 20, no. 4 (2020): 48–56. https://dx.doi.org/10.22503/inftars.XX.2020.4.4

Kaluža, Jernej. "Hume's Empiricism versus Kant's Critical Philosophy (in the Times of Artificial Intelligence and the Attention Economy)." *Információs Társadalom* 23, no. 2 (2023): 67–82. https://dx.doi.org/10.22503/inftars.XXIII.2023.2.4

Keim, Brandon. *Meet the Neighbors: Animal Mind and Life in a More-Than-Human World.* New York: W. W. Norton & Company, 2024.

Land, Nick. "Primordial Abstraction." *Jacobite,* April 3, 2019. https://www.scribd.com/document/808999040/Primordial-Abstraction-Jacobite

Larsson, Stefan, and Fredrik Heintz. "Transparency in Artificial Intelligence." *Internet Policy Review* 9, no. 2 (2020). https://dx.doi.org/10.14763/2020.2.1469

Ledoux, Joseph. *The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains*. New York: Viking, 2019.

Levinovitz, Alan. The Mystery of Go, the Ancient Game That Computers Still Can't Win. *Wired,* May 12, 2014. https://www.wired.com/2014/05/the-world-of-computer-go/

McCorduck, Pamela. *Machines Who Think*. 2nd edn. Natick, MA: A K Peters, 2004.

Mendon-Plasek, Aaron. "Mechanized Significance and Machine Learning: Why It Became Thinkable and Preferable to Teach Machines to Judge the World." In *The Cultural Life of Machine Learning*, edited by Jonathan Roberge and Michael Castelle, 31–78. Cham: Palgrave Macmillan, 2021.

Natale, Simone. *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test.* Oxford: Oxford University Press, 2021.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information.* Cambridge, MA: Harvard University Press, 2015.

Pasquale, Frank. *The New Laws of Robotics: Defending Human Expertise in the Age of AI.* Cambridge, MA: Belknap, 2020.

Pasquinelli, Matteo. The Eye of the Master. *A Social History of Artificial Intelligence.* London: Verso, 2023.

Weizenbaum, Joseph. *Computer Power and Human Reason: From Judgement to Calculation.* New York: W. H. Freeman, 1976.

Weizenbaum, Joseph. "ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine." *Communications of the ACM* 9, no. 1 (1966): 36–45. https://doi.org/10.1145/365153.365168