



Finding AI Faces in the Moon and Armies in the Clouds: Anthropomorphising Artificial Intelligence in Military Human-Machine Interactions

James Johnson

To cite this article: James Johnson (2023): Finding AI Faces in the Moon and Armies in the Clouds: Anthropomorphising Artificial Intelligence in Military Human-Machine Interactions, *Global Society*, DOI: [10.1080/13600826.2023.2205444](https://doi.org/10.1080/13600826.2023.2205444)

To link to this article: <https://doi.org/10.1080/13600826.2023.2205444>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 27 Apr 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Finding AI Faces in the Moon and Armies in the Clouds: Anthropomorphising Artificial Intelligence in Military Human-Machine Interactions

James Johnson

Department of Politics and International Relations, University of Aberdeen, Aberdeen, UK

ABSTRACT

Why are we likely to see anthropomorphisms in military artificial intelligence (AI) human-machine interactions (HMIs)? And what are the potential consequences of this phenomena? Since its inception, AI has been conceptualised in anthropomorphic terms, employing biomimicry to digitally map the human brain as analogies to human reasoning. Hybrid teams of human soldiers and autonomous agents controlled by AI are expected to play an increasingly more significant role in future military operations. The article argues that anthropomorphism will play a critical role in future human-machine interactions in tactical operations. The article identifies some potential epistemological, normative, and ethical consequences of humanising algorithms for the conduct of war. It also considers the possible impact of the AI-anthropomorphism phenomenon on the inversion of AI anthropomorphism and the dehumanisation of war.

ARTICLE HISTORY

Received 1 October 2022

Accepted 12 April 2023

KEYWORDS

Artificial intelligence;
autonomous weapons;
human-machine interaction;
military tactics;
anthropomorphism

Introduction

Scottish philosopher David Hume asserted that “there is a universal tendency among mankind to conceive all beings like themselves ... we find faces in the moon, armies in the clouds” (Hume 1957, 29). In recent years, the study of anthropomorphism – people’s propensity to attribute the traits of human agents to non-human ones – has become a multi-disciplinary phenomenon encompassing insights from social psychology and cognition, social science, the theory of mind, behavioural science, philosophy, and most relevant to this study, neurosciences (Bering 2006; Kwan and Fiske 2008; Duffy 2003).

How human warfighters perceive these interactions are critical to how they function as part of a hybrid team; trust, acceptance, tolerance, and social connection contribute to this interface’s scope, efficiency, and reliability. Understanding the various psychological mechanisms that undergird artificial intelligence (AI)-anthropomorphism is crucial in determining the potential impact of military human-machine interactions. This

CONTACT James Johnson  james.johnson@abdn.ac.uk  Department of Politics and International Relations, University of Aberdeen, King’s College, Aberdeen, UK

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

appreciation is critical to increasing the accuracy (predictive and explanatory behaviour), reliability (interpreting human goals and priorities), and efficacy (coordinating and planning tactical operations) of human-machine interactions (HMIs) in military operations (Cappuccio, Galliot, and Sandoval 2021a). The article addresses two related questions a) why are we likely to see anthropomorphisms in military AI HMIs? And b) what are the potential consequences of this phenomena? The article approaches these research puzzles primarily through empirical work conducted by the computer and behavioural science literature that considers HMIs in non-military settings. These findings are interpreted with a read-through for the military domain, for which non-classified studies are limited. The article's findings contribute to the IR and social sciences and the technical-scientific scholarship from where it originates epistemologically.

While much of the literature (Duffy 2003; Bartneck et al. 2009; Mutlu et al. 2011; Clark 2015) focuses on anthropomorphism's situational, developmental, or cultural determinants, less attention (exceptions include: Waytz, Cacioppo, and Epley 2010; Cappuccio, Galliot, and Sandoval 2021b; Spatola and Chaminade 2022; Johnson 2023) considers the potential consequences of human warfighters' *perceptions* of military human-machine interaction and, specifically, their tendency of attributing human traits to machines. This research speaks to the burgeoning literature on autonomous weapon systems (AWS), weaponized AI, and the proliferation of disruptive emerging technology in the military context (Bode and Huelss 2022; Johnson 2020, 2023; Payne 2016; Scharre 2019; Singer 2011). This article focuses on the impact and design of anthropomorphism in AI systems used in military HMIs – that is, the design and the use of computer technology and interfaces, and interaction between human users and machines in hybrid teams. Thus, it addresses an important epistemological and normative gap in our understanding of the nature and consequences of military HMIs (Card, Moran, and Newell 1983). It also elucidates the lightly researched (especially in the existing international relations literature) potential impact of the AI-anthropomorphism phenomenon on the inversion of AI-anthropomorphism and the dehumanisation of war.

The article argues that anthropomorphism will play a critical role in human-machine interactions in tactical operations. The article addresses the following research questions: What explains the psychological origin and persistence of anthropomorphism? What are the risks and opportunities associated with AI-anthropomorphism within AI agent-soldier teams? What are the possible consequences of AI-anthropomorphism of AI-enable military HMI? And in response, what are the most effective design solutions to maximise the advantages and minimise the risks in future HMI interfaces? The article will approach these research puzzles primarily through empirical work conducted by the computer and behavioural science literature that considers HMIs in non-military settings. These findings are interpreted with a read-through for the military domain, for which non-classified studies are limited. Given the paucity of relevant studies – that is, anthropomorphising tendencies of AI in military HMIs – and the embryonic nature of much of the technology discussed, much of the articles' discussion is necessarily conceptual and speculative.

The article is organised into two sections. The first traces the psychological origins, mechanisms, and persistence of anthropomorphism. This section is the article's empirical contribution, draws insights from the latest civilian social robotics and social cognition research to elucidate the impact and design of anthropomorphism in AI systems

used in military HMI in hybrid teams. Specifically, it considers when, why, and for whom anthropomorphism's effects are most likely to occur in military HMIs. Section two considers the potential problems such as ethical and moral, trust and responsibility, and social influence and unintended consequences of AI-enabled military HMIs operations. This section closes with a brief discussion on the potential implications of the inverse process of anthropomorphism, dehumanisation for AI-enabled military HMIs.

Conceptualising AI-anthropomorphism in military HMI operations

Cognitive and social psychologists, philosophers, and anthropologists have elucidated the origin of anthropomorphism as an evolutionary and cognitive adaptive trait, particularly concerning theistic religions (Ellis and Bjorklund 2004). Scholars speculate that for evolutionary reasons, early hominids (i.e. members of a family Hominidae, the great apes) interpreted ambiguous shapes as faces or bodies to improve their genetic fitness by making alliances with neighbouring tribes or by avoiding neighbouring outgroups threats and predatory animals (Guthrie 1995). Scholars have recently described people's propensity to turn non-human agents into human ones (Epley and Waytz 2013). Thus, the psychological and behavioural mechanisms intrinsic to the phenomenology of anthropomorphism are considered universal (across genders, race, and cultures), cognitively deep, innate, and developed in human's formative years (Duffy 2003; Dacey 2017).

Anthropomorphism, therefore, is a process of inference encompassing not only physical features but also perceiving an agent in a human-like form, thus imbuing it with mental capacities that humans consider uniquely human, such as emotions (e.g. empathy, revenge, shame, and guilt), and the capacity for conscious awareness, metacognition, and intentions. Moreover, anthropomorphism is the result not only of an agent's behaviour but also of the human perceiver's motivation, social background, gender, and age (Eyssel and Kuchenbrandt 2012; Eyssel et al. 2012; Hegel et al. 2012). In other words, anthropomorphism is very context dependent; different representations and judgments of the same non-human agent may be produced by various, and even the same, individuals (Spatola and Chaminade 2022).

Persistence of anthropomorphism

Epley et al. proposed a theory that determined three psychological factors as affecting when people anthropomorphise non-human agents. These variables, either independently or in combination, help us to elucidate the tendency of individuals and groups to anthropomorphise non-human agents in HMIs (Dawes and Mulford 1996). First, because people have a much richer knowledge of humans than non-human agents like AI, individuals are thus more likely to seek anthropomorphic explanations of non-human agents' actions to create mental models and heuristics. Second, when individuals are motivated to explain or understand an agent's behaviour – to reduce uncertainty and ambiguity and control one's environment and the need for cognitive closure – the tendency to anthropomorphise generally increases (Kruglanski and Webster 1996). Third, individuals who lack adequate levels of human social connection tend to compensate for this by treating non-human agents as if they were human.

The theory predicts that warfighters' predisposition to anthropomorphise machines is highest in situations where they are aware of the features and functions that justify human-machine analogies (i.e. accessible and applicable anthropocentric knowledge) when their survival depends on the cohesion and solidarity of their team members. In situations where users perceive machines as a threat or need to feel less isolated and alone (i.e. the desire for social contact and affiliation), they are more likely to anthropomorphise (i.e. the desire for social contact and affiliation) (Cappuccio, Galliot, and Sandoval 2021b). Because of the fuzzy nature of ML algorithmic logic – coupled with the high incentives for understanding and effectively interfacing with AI agents – the tendency to anthropomorphise the workings of many non-humans AI agents will likely be especially acute. In approaching this problem, AI designers must ensure that algorithmic decisions are explainable, reliable, and predictable – see below for possible ways to achieve this goal.

The perception that AI systems benefit from the projection of anthropomorphism from its users to, for instance, cope with information overload, promote acceptance, and foster trust and cooperation in HMIs has prompted developers to deliberately elicit this reaction to facilitate the utility of AI agents (Moreale and Watt 2004). Studies have demonstrated, for example, that humans judge robots that exhibit playful behaviour as more outgoing, and others that appeared more were considered easier to cooperate and work with (DiSalvo et al. 2002). In addition to the perception of efficiency, anthropomorphising non-human agents can foster close social connections, which despite being far less meaningful than human interactions, can make its users more cognitively more favourably disposed towards technological agents than might otherwise be the case (Airenti, Cruciano, and Plebe 2019).¹ In short, anthropomorphism is not simply a by-product of HMI but rather an intrinsic feature, embodying social cognitive features, potentially enabling mutual adaptation and coordination during intersubjective and complex decision-making (Cappuccio 2014).

Anthropomorphism in AI by design

From depictions of Alan Turing's early computational machines to AlphaZero's modern-day technological infamy, researchers often use human-like traits, concepts, and expertise when referring to AI systems to highlight the similarities of humans and AI algorithms (Salles, Evers, and Frisco 2020). Intellectual and emotional anthropomorphic manifestations are deliberately baked into AI systems by their designers, for the efficacious, control, and social cognitive reasons described, so they can be used in HMI. In this sense, the perception of human users interacting with the AI-system seems to partially being shaped by design choices.

Other possible driving forces underlying the humanisation of AI by designers include the intrinsic epistemic limitation and bias of AI researchers and a broader shift in science from the late 19th century from "eliminativism" (the belief that our understanding of the mind is wrong and that many of the mental states posited by common-sense do not exist in reality) and "psychophobia" (an irrational fear of the mind) to an emphasis on

¹The experience of social connection (e.g., emotional proximity) with non-human agents is not necessarily anthropomorphic.

“anthropocentric” (viewing humans as the central or most important element of existence) mental concepts and terms applied to inanimate non-human entities. The tendency of popular culture and media coverage to emphasise the human-like qualities (emotional, cognitive, sentience, consciousness, ethical, etc.) of AI and robots creates a limited understanding of the state of AI capabilities. It inadvertently expounds false notions about what AI can and cannot do, thus creating dystopian and utopian polarising expectations (Bartneck 2013).

The high-profile success of systems like these (e.g. Israel’s Harpy loitering munition, Russia’s stealth Volk-18 UAV, and the US’s Loyal Wingman drone) have contributed further to the public and scientific alacrity that the development of AI depends on the emulate the human brain – and thus also critical in achieving a better understanding of how the human brain works. Critics argue that these conceptualizations are misleading for the users and researchers of the system alike, understating the critical epistemological—how humans gain an understanding of the world through intuition, perception, introspection, memory, reason, and testimony—differences between human intelligence (and other attributes) and AI (McDermott 1976; Hassabis et al. 2017; and Ullman 2019). AI researcher David Watson writes: “It would be a mistake to say that these algorithms recreate human intelligence; instead, they introduce *some new mode of inference that outperforms us in some ways and falls short in others*” (emphasis added) (Watson 2019, 425). Whether the goal of future AI will be to replicate the human brain’s functional architecture (beliefs, desires, and intention models, etc.) or innovate an entirely novel approach to “intelligence,” is an open question that has profound epistemic consequences for trust, acceptance, and tolerance in HMIs that we explore below.

Military HMI in tactical hybrid teaming

This section presents empirical work from military HMI in hybrid teaming operations, drawing insights from social psychology, philosophers, and anthropology, to consider the impact and design of anthropomorphism in AI systems. The case studies in the scientific and social science literature on anthropomorphism in HMI’s military applications draw from and complement parallel research in civilian social robotics and social cognition (Carpenter 2013; Singer 2011; Galliot 2016). Civilian studies can highlight some of the potential unintentional consequences and risks of anthropomorphism in HMI and thus may provide novel and innovative ways of integrating AI in military hybrid teaming (Hoffman and Breazeal 2004). Specifically, social robotic studies demonstrate that a critical precondition in successful HMIs is how humans perceive non-human agents’ expertise, emotional engagement, and perceptual responses (Nass and Moon 2000). Thus, how AI agents are viewed by human military personnel will crucially influence the amount of trust, acceptance, and tolerance afforded to them, and thus the efficacious (i.e. function and the scope) hybrid teaming (Mutlu et al. 2011).

Conceptually speaking, several (combat and non-combat) physical (or dull, dirty, or dangerous) and decision-making military tasks, depending on the context and technology involved, could soon be delegated to or conducted with AI agents, including intelligence surveillance, and reconnaissance (ISR); selective target guidance and engagement; perimeter and border protection; shielding of military personnel and civilians; bomb disposal; handling of chemical, biological and nuclear materials; logistics and

transportation; “loyal wingman” drones to support manned fighter pilots; and medical and psychological assistance and training to the military.

Advances in bioelectric signals technology, such as electromyogram (EMG) and electroencephalogram (EEG) that reflect human internal states and intended actions, will soon enable new kinds of brain-computer interface (BCI) (Hayashi and Tsuji 2022). Thereby, allowing intuitive control of machines and connecting human neural functions to various command and control military systems, such as controlling drone swarms or even jetfighters (Charette 2018). US DARPA’s “ElectRx” programme, for example, is developing neural implants that interface directly with the nervous system to continuously assess the state of soldiers’ health, to regulate conditions such as depression, Crohn’s disease, and post-traumatic stress disorder (Otto and Bryant Webber 2013).² In HMIs, where the physical interaction is close and persistent, such as exoskeletons, human and machine behaviour forms a mutually dependent relationship, where both the goals and the physical effort applied for their efficacy are intertwined and must be co-jointly determined for a smooth and effective interaction (Giordano 2015).

The most potentially transformative effects of AI technology on HMI are less the use of biological implants in the bid for cyborg-like mergers, but rather the kinds of non-penetrative forms of augmentation that might transform the socio-technical problem-solving matrix with potentially profound human psychological implications (Clark 2005). Because the kinds of human-machine symbiosis we have discussed will likely depend on, for their ultimate success, intimate – however imperfect or superficial – technologized social interactions, some scholars worry that the effect of these interactions on moral responsibility and personal identity might adversely impact human-to-human interactions (Pickering 2001). For instance, it might cause humans to treat others on an equal moral footing with (or even below) AI agents, or worse still, ethically desensitising, or dehumanising human-to-human contact.

In a recent series of aerial combat simulations hosted by US DARPA as part of their “AlphaDogfight” project, AI agents were pitted against human F-16 fighter pilots in virtual dogfights – AI-powered fighters comprehensively defeated their human adversaries. In a separate project, a collaborative project with Boeing and the US Air Force is developing the “Loyal Wingman” project of supersonic autonomous combat drones capable of flying in formation with fifth-generation F-35 fighter jets – defending them from enemy attack and autonomously coordinating with on-board systems and pilots in joint attack missions (Tucker 2020). AI agents operate in physical and simulated domains. Adaptive algorithms control them, algorithms that change their behaviour when it is run based on information available and *a priori* defined reward mechanism, and machine learning (ML) systems can navigate and manipulate their environment and select optimum task-resolution strategies (Ferreira 2020). Tactical systems that use new-generation AI-enhanced aerial combat drones, such as the Loyal Wingman aircraft or AlphaDogfight simulations, in asymmetric offensive operations, AI systems could be trained – or eventually autonomously “learn” – to suppress specific anthropomorphic cues and traits or to use human-like cues and traits to generate false flag or other deception and disinformation operations.

²Studies demonstrate that human drone pilots display similar levels of PTSD and other mental health conditions compared to conventional pilots (Otto and Bryant Webber 2013).

Understanding the determinants and drivers of anthropomorphism can, therefore, help us to identify the conditions under which these effects will be most impactful. In short, the design of AI agents for hybrid teaming must embody both the positive and potentially negative psychological implications of anthropomorphism. AI agents must produce predictable, purposeful, and well-communicated behaviours, correctly identifying human intentions and the drivers of human behaviour – and in turn, relate to them. Identifying others' intentions is complicated when information is complex and overwhelming – which can also impair joint coordination – and when the nature of others' intentions is opaque because of deception, manipulation, bodily behaviour, emotional states, and cues obscured.

Strategies of deception and manipulation of information, signals, and intentions – to distract an adversary and delay or inhibit its ability to respond – can be replicated and magnified using AI technology (e.g. chatbots, digital avatars, deep-fake technology, and AI-augmented adversarial attacks and electromagnetic warfare) in ways that can make anthropomorphism more acute (Knight 2022). In tactical HMI's, the need for rapid decision-making in dynamic and contingent situations will complicate the challenge of accurately interpreting human bodily actions and subtle cues when AI agents (and machines and artificial tools generally) are used as a medium (Cappuccio, Galliot, and Sandoval 2021b). That is, interpreting the mental state of a combatant in close physical contact is generally easier than when they are using tools (drones, digital assistants, and other vehicles) that hide bodily expressions (Yong 2022). For example, using new-generation AI-enhanced aerial combat drones in asymmetric offensive operations, AI systems might be trained – or eventually autonomously “learn” to – to suppress specific anthropomorphic cues and traits or use human-like cues and traits to generate false flag or other disinformation operations (Cappuccio, Galliot, and Sandoval 2021a).

The consequences of AI-anthropomorphism

In a military context, individuals perceiving an AI agent to have human-like qualities (mind, intelligence, emotion, sentience, consciousness, etc.) have significant ethical, moral, and normative consequences for both the human perceiver and the AI agent perceived (Gray, Gray, and Wegner 2007). While some scholars contend that anthropomorphic projections (explicit or implicit) might expose soldiers in hybrid teams to physical and psychological risks (Scharre 2019). Others, in contrast, by understating the potential impact of anthropomorphism in AI on the performance of human operators, risk underplaying the tactical, ethical, and cognitive implications (Barnes and Evans 2010).

Ethical and moral

In addition to the epistemological problems described, the anthropomorphic rhetoric (or anthropomorphism in AI by design) surrounding the development of AI systems also has significant ethical consequences for HMIs. Perceiving an AI agent as conscious and possessing human-like intelligence implies that AI agents should be treated as “moral agents” with moral autonomy (the capacity to regulate one's actions through moral principles or ideals that shapes a persons' own narrative), and thus deserving of protection, empathy,

and rights such as autonomy and freedom (Tiku 2022). By anthropomorphising non-human agents, we are *ipso facto*, allowing them to be moralised.

Anthropomorphising terms like “ethical,” “intelligent,” and “responsible” in the context of machines can lead to false attributions and mythical tropes implying that inanimate AI agents are capable of moral reasoning, compassion, empathy, mercy, etc. and thus might perform more ethically and humanely than humans in warfare (Kurzweil 2000). Robotist Ronald Arkin’s research on developing autonomous battlefield robots with an artificial conscious and synthetic uploaded human ethics demonstrates what can happen when anthropomorphic tropes and perceptions of machine ethics are applied to make an equivalence with human morality in war. Arkin argues that “I am convinced that they [autonomous battlefield robots] can perform more ethically [and more humanely] than human soldiers are capable of” (Arkin 2009, 47–48).

Similarly, the European Remotely Piloted Aviation Systems (RPAS) Steering Group, in its report on drones, stated that “citizens will expect drones to have an *ethical behavior comparable to the human one*, respecting some commonly accepted rules” (emphasis added) (European RPAS Steering Group 2013, 44). The shift from viewing technology as tools to support military operations to becoming an integral team member, or even a source of moral authority, rests on an anthropomorphic expectation that machines as moral agents can exhibit human-like rationally, dispassionately, and ethnically in the conduct of war. Some fear that greater levels of automation and intelligence in AI systems may further entrench the authoritative status of technology in war, such that machines become “a science of imaginary technical solutions to the problem of war legitimization,” and in turn, further dehumanise warfare (discussed below) (Roderick 2010, 228).

Some scholars describe the semantical problem of how we conceptualise ethics and machines in war; namely, the distinction between machines behaving ethnically – assuming that machines have sufficient agency and cognition to make moral decisions, and machines being used (i.e. by humans) ethically in operational contexts. In a recent report on the role of autonomous weapons, the US Defense Advisory Board alluded to this problem, concluding that “treating unmanned systems as if they had *sufficient independent agency to reason about morality* distracts from designing appropriate rules of engagement and *ensuring operational morality*” (emphasis added) (US Task Force Report, 2012, 48). Using anthropomorphic language to conflate human ethics and reasoning with machine logical inductive statistical reasoning – on the false premise that machine and human ethical reasoning in war are similar. In short, humanising AI is not ethically or morally neutral; instead, it presents a critical barrier to conceptualising the many challenges AI poses as an emerging technology (Floridi and Sanders 2004).

Trust & responsibility

An AI agent with human intelligence capable of intentional action would presumably be worthy of human “trust” and thus held legally and morally responsible for its actions (Taddeo 2010).³ To be sure, it is highly speculative whether machines will ever be

³Human trust is a psychologically nebulous and sticky notion, which cannot be modeled by algorithms or adequately formalized by rational choice theory.

endowed with the sorts of agency (or “human intelligence”) to merit legal and moral culpability. Were military personnel to perceive AI agents as more capable and intelligent than they are (or “automation bias”), they may become more predisposed to “social loafing” (or complacency) in tasks that require human and machine collaboration, such as target acquisition, intelligence gathering, or battlefield situation awareness assessments (Skitka, Mosier, and Burdick 1999). In other words, the anthropomorphic tendency of people to conflate a technological capacity for accuracy and speed with tactical competency means that AI agents are more likely, for better or worse, to be judged as responsible and thus trustworthy in the conduct of war – and other safety-critical HMI collaborative domains such as robotic surgery (Verger 2021). Whether these responses are simply the result of anthropomorphism, and how using AI might affect radiologists’ decisions, and if so, whether this creates new risks, has yet to be empirically tested (Kiros 2022).

People tend to mistakenly infer an inherent connection between these human traits to machines when their performance matches or surpasses humans (Florida 2017). Moreover, people are more likely to feel *less responsible* for the success or failures of tasks that use human-like HRIs and treat anthropomorphic interactions with AI agents as scapegoats when the technology malfunctions. Paradoxically, advances in autonomy and machine intelligence will require more (rather than less) contributions from the human operator to cope with the inevitable unexpected contingencies that fall outside of an algorithm’s training parameters or fail in some way (Gray, Gray, and Wegner 2007). Overconfidence in the abilities and trust (mis)placed in AI agents, coupled with the abdication of responsibility, might result in the proliferation of these technologies (to state and non-state actors), lower the threshold for war, and make inadvertent and accidental war more likely (Duffy 2003).

Studies demonstrate that individuals more willingly punish an agent they consider intelligent and conscious of legal and moral violations (Gray, Gray, and Wegner 2007). Moreover, people are more likely to hold groups (militaries, corporations, governments, etc.) comprising single personified agents more legally culpable for moral violations than those representing collectives of disparate individuals (French 1986). Furthermore, if an AI agent is deemed responsible for their actions, then the humans controlling or collaborating with AI agents may *ipso facto* consider themselves *less* responsible for the actions resulting from hybrid teaming decisions. If war crimes, for instance, treated AI agents as “moral agents” would complicate the attribution of responsibility (Arkin 2009), which has become a key point of contention in international debates on lethal autonomous weapons (LAWS) (Bode and Huelss 2022). Debates about diffusing safety-critical moral and legal responsibility to AI-powered decision-support systems are also evident in the medical domain (Bleher and Braun 2022). If the decisions and actions of AI agents during combat *appear* “human-like,” does this necessarily decrease the perceived responsibility of the humans who designed the algorithms or collaborated with AI agents in hybrid teaming?

Inverting AI anthropomorphism and the dehumanisation of war

Whereas anthropomorphism is the process of perceiving non-human agents to possess human-like qualities, dehumanisation represents the inverse process (Haslam 2006;

Johnson 2023). Just as increasing levels of similarity to humans can invoke the tendency to anthropomorphise a non-human agent, so can decreased similarity increase the tendency to dehumanise other humans (Harris and Fiske 2006). Humanness exists on a continuum; how we perceive others is inextricably connected to how we perceive non-humans. The psychological mechanisms that make people likely to attribute human-like qualities can also increase our understanding of when and why people do the opposite. Using this theoretical inversion, we can draw insights to understand better the potential consequences of anthropomorphism in AI-enabled military HMIs and for general dehumanising war.⁴ As a counterpoint (discussed below), the increasing remoteness of warfare through the use of military AI and autonomous systems may also account for – or at least act as an aggravator – of this phenomenon. Moreover, the literature on autonomous weapons systems reveals that dehumanising behaviour may also derive from the algorithmic processing of people (Waytz, Cacioppo, and Epley 2010). In this case, the dehumanisation of warfare may occur in the absence of anthropomorphism. Evidence indicates, for example, that unmanned remote drones do not dehumanise warfare in the way people expect. Counterintuitively, rather than treating combat as a video game, human drone pilots often form deep emotional bonds with their targets, and in this war at a distance, many pilots suffer long-term mental health issues similar to traditional combat experience (Saini, Raju, and Chail 2021).

If technologies like AI and autonomous weapons draw warfighters further away from the battlefield, they risk becoming conditioned to view the enemy as inanimate objects “neither base nor evil, but also things devoid of inherent worth” (Brough 2007). Although the “emotional disengagement” associated with a mechanistically dehumanised enemy is considered conducive for combat efficiency and tactical decision-making, the production of controlled and banal socio-technical interactions devoted to moral emotions is ethically and morally lamentable (French and Jack 2015). As political philosopher Hannah Arendt warned: the development of *robot soldiers* ... would *eliminate the human factor* and, conceivably, permit one man with a push button to destroy whomever he pleases (emphasis added) (Arendt 1970, 50).

The tendency to anthropomorphise when people are motivated to explain or understand an agent’s (human or non-human) behaviour described earlier should exhibit the inverse dehumanising proclivity when individuals are motivated to *reduce* levels of interaction with others – and thus are not motivated or desire to understand, develop social connections, or empathise with (Waytz, Cacioppo, and Epley 2010). Power and influence over others are crucial determinants for increasing an individual’s independence, thus decreasing the need for effective interaction with others (Bandura, Underwood, and Fromson 1975). In a recent social psychology study, for instance, people in a position of power increased the propensity to objectify subordinates, regarding them as a means to end and neglecting their essentially human qualities (Gruenfeld et al. 2008). As a corollary, soldiers in anthropomorphised hybrid teaming might a) come to view their inanimate machine “team-members” as deserving of more protection and care

⁴The literature on autonomous weapons systems has shown that dehumanising behaviour may also derive from the algorithmic processing of people. In this case, the dehumanisation of warfare may occur in the absence of anthropomorphism (Bode and Huelss 2022).

than their human adversary,⁵ or b) soldiers intoxicated by the power over an adversary may (especially in an asymmetrical conflict) become more predisposed to dehumanise the enemy (the out-group), justifying past wrongdoings, and excessive and potentially immoral acts of aggression.

Conclusion

This article considers the role of anthropomorphism in military human-machine interaction augmented by AI technology. It advances explanations for the psychological origin and persistence of anthropomorphism, the risks and opportunities associated with AI-anthropomorphism within AI agent-soldier teams, possible consequences of AI-anthropomorphism of AI-enabled military HMI, and in response, possible design solutions to maximise the advantages and minimise the risks in future HMI interfaces.

The article's key findings can be summarised as follows. First, understanding the various psychological mechanisms that undergird the phenomenology of AI-anthropomorphism in military HMI is a critical step in a) determining the potential positive and negative impact of military human-machine interactions, and thus, b) optimising the accuracy, reliability, and efficacy of HMIs in military operations. A key finding of the article is that the tendency to anthropomorphise AI agents in military HMIs would likely be especially acute because of machine-learning algorithmic logic coupled with the high incentives for understanding and effectively interfacing with AI agents.

Second, anthropomorphism can create a sense of efficacy and competence in interacting with AI agents in military HMIs. This perception has prompted developers to deliberately and explicitly elicit this reaction to optimise the integration of AI agents into hybrid team operations. Moreover, anthropomorphising AI agents, particularly in time-pressured and stressful war conditions, may encourage social bonding in HMIs, making its users more cognitively disposed towards technological agents (i.e. caring about their well-being) than might otherwise be the case (Xie and Pentina 2022). This disposition may prompt soldiers in anthropomorphised hybrid teaming to either view their AI "team-mates" as deserving of more protection and care than their human adversary or make them more likely to dehumanise the enemy, justifying excessive and potentially immoral acts of aggression. More empirical work is needed, however, on the psychological impact of the perception, whether accurate or otherwise, of machine efficacy and competence on HMIs.

Third, a significant worry with anthropomorphic language and popular tropes in describing AI in military HMIs is that it overlooks the intrinsic limitations of AI technology (i.e. brittle, inefficient, vulnerable, and myopic), thus creating a false equivalence between human and machine intelligence – which are ontologically, epistemologically, and metaphysically very different. Once AI systems are anthropomorphised, their statistical probabilistic outputs may be treated as equivalent to human judgments, decisions, and "functional" ethics in war, which risks abdicating control over human ethical

⁵As a counterargument, teams comprising solely of humans on both sides might equally come to view their teammates as deserving of more protection and care than their (human) adversaries. This would, of course, be less morally problematic than holding this view for an inanimate machine. The author thanks the anonymous reviewers for this suggestion.

decision-making to machines. Relatedly, the anthropomorphic tendency of people to conflate a technological capacity for accuracy and speed with tactical competency, exacerbated by “automation bias”, means that AI agents are more likely to be judged as responsible and thus trustworthy in the conduct of war. By anthropomorphising AI, we risk affording machines a level of unwarranted agency that exaggerates its capabilities and may reduce human autonomy and sense of agency.

Finally, in military HMI design, an essential precondition for success is how humans perceive an AI agent’s expertise, emotional engagement, and perceptual responses – influencing the trust, acceptance, and tolerance placed in machine team members. AI agents must produce predictable, purposeful, and well-communicated behaviours, correctly identifying human intentions and the drivers of human behaviour, and in turn, relate to them. In tactical HMIs, the need for rapid decision-making will likely complicate the challenge of accurately interpreting human bodily actions and subtle cues when AI agents are used as a medium.

The phenomenology of AI-anthropomorphism and its impact on HMIs in military hybrid collaboration needs to be acknowledged and understood by the AI and defense research community, its users, and the broader constituents of the socio-technical ecosystem if they desire to realistically anticipate both the opportunities, challenges, and risks associated with hybrid tactical teamwork. To date, while the risks associated with dysfunctional AI in HMIs highlighted in this paper should not be underestimated, the evidence suggesting that anthropomorphism in HMIs leads to more risky behaviour and accidents is anecdotal. It does not justify the prohibition of anthropomorphic AI design. Some argue that many of the risks associated with anthropomorphic tendencies in HMIs could be mitigated and controlled through appropriate monitoring, design, training, and force structuring.

Possible policy measures designed to maximise the advantages and minimise the risks in future HMI interfaces identified in this article that policy-makers, designers, and users might consider include, *inter alia*: (1) the design of AI-driven systems to monitor biases, errors, adversarial behaviour, potential anthropomorphic risk, and incorporating “human” ethical principles and norms in AI systems while retaining the role of humans as moral agents and keeping humans in the loop as fail-safes (Singer 2009); (2) training that emphasises “meaningful human control” – in accordance with human designs and legal and ethical constraints moral responsibility – and a culture of collective vigilance against automation bias and complacency in hybrid teaming (Hagerott 2014); (3) educating both combatants and support staff about the possible benefits and risks of anthropomorphising AI-agents; (4) regulating human-machine interfaces to counteract the potential impact of dehumanisation, groupthink, and other concerns related to diffused moral responsibility; and; (5) closely coordinating force structuring decisions with training exercises to maximise human-machine communications, particularly when communications are restricted or compromised. These efforts should be coordinated and implemented to optimise human-machine communication and establish appropriate levels of trust, acceptance, and tolerance in human-machine interactions.

Future empirical studies should examine: the prevalence of anthropomorphism in military HMIs to validate the mostly anecdotal claims about the risks of anthropomorphism in warfare; the extent to which there are differences between how various groups of

actors in the military anthropomorphise; the risks and benefits associated with anthropomorphism within AI-agent-human soldier teams; the effects of anthropomorphism on the interactions between hybrid military teams and external entities; and the optimum design solutions to maximise the benefits and minimise the risks in the future human-machine (and human-human) interactions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Airenti, Gabriella, Marco Cruciano, and Alessio Plebe. 2019. *The Cognitive Underpinnings of Anthropomorphism*. Lausanne: Frontiers Media.
- Arendt, Hannah. 1970. *On Violence*. New York: Harcourt Publishing Co.
- Arkin, Ronald. 2009. *Governing Lethal Behavior in Autonomous Robots*. Boca Raton, FL: Chapman, pp. 47-48.
- Bandura, Albert, Bill Underwood, and Michael Fromson. 1975. "Disinhibition of Aggression Through Diffusion of Responsibility and Dehumanization of Victims." *Journal of Research in Personality* 9 (4): 253–269. doi:10.1016/0092-6566(75)90001-X.
- Barnes, Michael, and William Evans. 2010. "Soldier–Robot Teaming: An Overview." In *Human-robot Interactions in Future Military Operations*, edited by Michael Barnes, and Florian Jentsch, 9–31. Farnham: Ashgate Publishing.
- Bartneck, Christoph. 2013. "Robots in the Theatre and the Media." *Proceedings of the Design and Semantics of Form and Movement, Wuxi*, 64–70.
- Bartneck, Christoph, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots." *International Journal of Social Robotics* 1 (1): 71–81. doi:10.1007/s12369-008-0001-3.
- Bering, Jesse. 2006. "The Folk Psychology of Souls." *Behavioral and Brain Sciences* 29 (5): 453–462. doi:10.1017/S0140525X06009101.
- Bleher, Hannah, and Matthias Braun. 2022. "Diffused Responsibility: Attributions of Responsibility in the use of AI-Driven Clinical Decision Support Systems." *AI and Ethics*, doi:10.1007/s43681-022-00135-x.
- Bode, Ingvild, and Hendrik Huels. 2022. *Autonomous Weapons Systems, and International Norms*. Queensland: McGill-Queen's University Press.
- Brough, Michael. 2007. "Dehumanization of the Enemy and the Moral Equality of Soldiers." In *Rethinking the Just War Tradition*, edited by Michael Brough, John Lango, and Harry van der Linden, 160–161. New York: SUNY Press.
- Cappuccio, Massimiliano. 2014. "Inference or Familiarity? The Embodied Roots of Social Cognition." *Synthesis Philosophica* 29 (2): 253–272. UDC 159.9.01:165.242.2.
- Cappuccio, Massimiliano, Jai Galliot, and Eduardo Sandoval. 2021a. "Mapping Meaning and Purpose in Human-Robot Teams: Anthropomorphic Agents in Military Operations." *The Philosophical Journal of Conflict and Violence* 5 (1): 86. doi:10.22618/TP.PJCV.20215.1.139005.
- Cappuccio, Massimiliano, Jai Galliot, and Eduardo Sandoval. 2021b. "Saving Private Robot: Risks and Advantages of Anthropomorphism in Agent-Soldier Teams." *International Journal of Social Robotics*, doi:10.1007/s12369-021-00755-z.
- Card, Stuart, Thomas Moran, and Allen Newell. 1983. *The Psychology of Human–Computer Interaction*. London: CRC Press.
- Carpenter, Julie. 2013. "Just Doesn't Look Right: Exploring the Impact of Humanoid Robot Integration Into Explosive Ordnance Disposal Teams." In *Handbook of Research on*

- Technoself: Identity in a Technological Society*, edited by Rocci Luppicini, 609–636. Hershey: IGI Global.
- Charette, Robert. “Automated to Death,” *IEEE Spectrum*, December 15, 2009, <https://spectrum.ieee.org/automated-to-death>.
- Clark, Andy. 2005. *Natural Born Cyborgs*. Oxford: Oxford University Press, pp. 13–34.
- Clark, Andy. 2015. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. New York: Oxford University Press.
- Dacey, Mike. 2017. “Anthropomorphism as Cognitive Bias.” *Philosophy of Science* 84 (5): 1152–1164. doi:10.1086/694039.
- Dawes, Robyn, and Matthew Mulford. 1996. “The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in how we Study Judgment?” *Organizational Behavior and Human Decision Processes* 65 (3): 201–211. doi:10.1006/obhd.1996.0020.
- DiSalvo, C., F. Gemperle, J. Forlizzi, and S. Kiesler. 2002. “All Robots are Not Created Equal: Design and the Perception of Humanness in Robot Heads.” Paper presented at the DIS2002 Conference Proceedings, London, 2002.
- Duffy, Brian. 2003. “Anthropomorphism and the Social Robot.” *Robotics and Autonomous Systems* 42 (3): 177–190. doi:10.1016/S0921-8890(02)00374-3.
- Ellis, Bruce, and David Bjorklund. 2004. *Origins of the Social Mind: Evolutionary Psychology and Child Development*. New York: (The Guildford Press.
- Epley, Nicholas, and Adam Waytz. 2013. “Mind Perception.” In *The Handbook of Social Psychology (5th ed.,)*, edited by Gilbert Fiske, and Lindzey, 498–454. New York: Wiley.
- European RPAS Steering Group. June 2013. *Roadmap for the Integration of Civil Remotely-Piloted Aircraft Systems Into the European Aviation System*. Brussels: European Commission.
- Eyssel, Friederike, and Dieta Kuchenbrandt. 2012. “Social Categorization of Social Robots: Anthropomorphism as a Function of Robot Group Membership.” *British Journal of Social Psychology* 51 (4): 724–731. doi:10.1111/j.2044-8309.2011.02082.x.
- Eyssel, Friederike, Dieta Kuchenbrandt, Simon Bobinger, Laura De Ruiter, and Frank Hegel. 2012. “If you Sound Like me, you Must be More Human: On the Interplay of Robot and User Features on Human-Robot Acceptance and Anthropomorphism.” In: *HRI’12: Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*, 125–126.
- Ferreira, Raul. “Machine Learning in a Nonlinear World: A Linear Explanation Through the Domain of the Autonomous Vehicles,” *European Training Network for Safer Autonomous Systems*, January 9, 2020, <https://etn-sas.eu/2020/09/01/machine-learning-in-a-nonlinear-world-a-linear-explanation-through-the-domain-of-the-autonomous-vehicles/>.
- Florida, Luciano. 2017. “Digital’s Cleaving Power and its Consequences.” *Philosophy & Technology* 30 (2): 123–129. doi:10.1007/s13347-017-0259-1.
- Floridi, Luciano, and J. W. Sanders. 2004. “On the Morality of Artificial Agents.” *Minds and Machines* 14 (3): 349–379. doi:10.1023/B:MIND.0000035461.63578.9d.
- French, Peter. 1986. “Principles of Responsibility, Shame, and the Corporation.” In *Shame, Responsibility, and the Corporation*, edited by H. Curtler, 17–55. New York: Haven.
- French, Shannon, and Anthony Jack. 2015. “Dehumanizing the Enemy: The Intersection of Neuroethics and Military Ethics.” In *Responsibilities to Protect: Perspectives in Theory and Practice*, edited by David Whetham, and Bradley Strawser, 165–195. Leiden: Brill.
- Galliot, Jai. 2016. “Defending Australia in the Digital age: Toward Full Spectrum Defence.” *Defence Studies* 16 (2): 157–175. doi:10.1080/14702436.2016.1165596.
- Giordano, James. 2015. “Neurotechnology, Global Relations, and National Security: Shifting Contexts and Neuroethical Demands.” In *Neurotechnology in National Security and Defense: Practical Considerations, Neuroethical Concerns*, edited by James Giordano, 1–10. BocaRaton: CRC Press.
- Gray, Kurt, Heather Gray, and Daniel Wegner. 2007. “Dimensions of Mind Perception.” *Science* 315 (5812): 619. doi:10.1126/science.1134475.
- Gruenfeld, Deborah, Ena Inesi, Joe Magee, and Adam Galinsky. 2008. “Power and the objectification of social targets.” *Journal of Personality and Social Psychology* 95 (1): 111–127. doi:10.1037/0022-3514.95.1.111.

- Guthrie, Stuart. 1995. *Faces in the Clouds: A new Theory of Religion*. Oxford: Oxford University Press.
- Hagerott, Mark. "Lethal Autonomous Weapons Systems from a Military Officer's Perspective: This Time is Different: Offering a Framework and Suggestions." Paper Presented at the United Nations Informal Meeting of Experts at the Convention on Conventional Weapons, May 15, 2014, Geneva, Switzerland.
- Harris, Lasana, and Susan Fiske. 2006. "Dehumanizing the lowest of the low: Neuroimaging responses to extreme outgroups." *Psychological Science* 17 (10): 847–853. doi:10.1111/j.1467-9280.2006.01793.x.
- Haslam, Nick. 2006. "Dehumanization: An Integrative Review." *Personality and Social Psychology Review* 10 (3): 252–264. doi:10.1207/s15327957pspr1003_4.
- Hassabis, Demis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick. 2017. "Neuroscience-inspired Artificial Intelligence." *Neuron* 95 (2): 245–258. doi:10.1016/j.neuron.2017.06.011.
- Hayashi, Hideaki, and Toshio Tsuji. 2022. "Human–Machine Interfaces Based on Bioelectric Signals: A Narrative Review with a Novel System Proposal." *IEEJ Transactions on Electrical and Electronic Engineering* 17 (11). doi:10.1002/tee.23646.
- Hegel, Frank, Sebastian Gieselmann, Annika Peters, Patrick Holthaus, and Britta Wrede. 2012. "Towards a Typology of Meaningful Signals and Cues in Social Robotics." In: *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, 72–78.
- Hoffman, Guy, and Cynthia Breazeal. 2004. "Collaboration in Human-Robot Teams." In: AIAA 1st Intelligent Systems Technical Conference. Infotech Aerospace Conferences, American Institute of Aeronautics and Astronautics.
- Hume, David. 1957. *The Natural History of Religion*. Stanford, CA: Stanford University Press.
- Johnson, James. 2020. "Artificial Intelligence in Nuclear Warfare: A Perfect Storm of Instability?" *The Washington Quarterly* 43 (2): 197–211. doi:10.1080/0163660X.2020.1770968.
- Johnson, James. 2023. *AI and the Bomb: Nuclear Strategy and Risk in the Digital Age*. Oxford: Oxford University Press.
- Kiros, Hana. 2022. "Doctors Using AI Catch Breast Cancer More Often Than Either Does Alone," *MIT Technology Review*, July 11, 2022, https://www.technologyreview.com/2022/07/11/1055677/ai-diagnose-breast-cancer-mammograms/?utm_source=substack&utm_medium=email.
- Knight, Will. 2022. "As Russia Plots Its Next Move, an AI Listens to the Chatter," *Wired*, April 4, 2022, <https://www.wired.com/story/russia-ukraine-war-ai-surveillance/>.
- Kruglanski, Arie, and Donna Webster. 1996. "Motivated Closing of the Mind: "seizing" and "Freezing."." *Psychological Review* 103 (2): 263–283. doi:10.1037/0033-295X.103.2.263.
- Kurzweil, Ray. 2000. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York, NY: Penguin.
- Kwan, Virginia, and Susan Fiske. 2008. "Missing Links in Social Cognition: The Continuum from non-Human Agents to Dehumanized Humans." *Social Cognition* 26 (2): 125–128. doi:10.1521/soco.2008.26.2.125.
- McDermott, Drew. 1976. "Artificial Intelligence Meets Natural Stupidity." *ACM Sigart Bulletin* 57 (4): 4–9. doi:10.1145/1045339.1045340.
- Moreale, Emanuela, and Stuart Watt. 2004. "An Agent-Based Approach to Mailing List Knowledge Management. Agent-Mediated Knowledge Management." *Lecture Notes in Artificial Intelligence* 2926: 118–129. doi:10.1007/978-3-540-24612-1_8.
- Mutlu, Bilge, Christoph Bartneck, Jaap Ham, Vanessa Evers, and Takayuki Kanda, eds. 2011. *Social Robotics: Third International Conference on Social Robotics, ICSR 2011, Amsterdam, The Netherlands, November 24-25, 2011. Proceedings* (New York: Springer).
- Nass, Clifford, and Youngme Moon. 2000. "Machines and Mindlessness: Social Responses to Computers." *Journal of Social Issues* 56 (1): 81–103. doi:10.1111/0022-4537.00153.
- Otto, Jean L., and J. Bryant Webber. 2013. "Mental Health Diagnoses and Counseling among Pilots of Remotely Piloted Aircraft in the United States Air Force." *MSMS* 20 (3): 3–8. PMID: 23550927.

- Payne, Kenneth. 2016. "Strategy in the Age of Artificial Intelligence." *Journal of Strategic Studies* 39 (5–6): 793–819. doi:10.1080/01402390.2015.1088838.
- Pickering, John. 2001. "Human Identity in the Age of Software Agents." In *Cognitive Technology: Instruments of the Mind: Proceedings of the 4th International Conference on Cognitive Technology*, edited by M. Beynon, C. Nahniv, and K. Duatenhahn, 442–451. Berlin: Springer.
- Roderick, Ian. 2010. "Mil-Bot Fetishism: The Pataphysics of Military Robots." *Topia* 23 (4): 228. doi:10.3138/topia.23-24.286.
- Saini, Rajiv, V. K. Raju, and Amitr Chail. 2021. "Cry in the Sky: Psychological Impact on Drone Operators." *Industrial Psychiatry Journal* 30 (1): 15–19. doi:10.4103/0972-6748.328782.
- Salles, Arleen, Kathinka Evers, and Michele Frisco. 2020. "Anthropomorphism in AI." *AJOB Neuroscience* 11 (2): 91–92. doi:10.1080/21507740.2020.1740350.
- Scharre, Paul. 2019. *Army of None: Autonomous Weapons and the Future of war*. New York: WW Norton & Company.
- Singer, Peter W. 2009. "In the Loop? Armed Robots and the Future of War," Brookings, January 28, 2009, <https://www.brookings.edu/articles/in-the-loop-armed-robots-and-the-future-of-war/>.
- Singer, Peter. 2011. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. New York: Penguin.
- Skitka, Linda, Kathleen Mosier, and Mark Burdick. 1999. "Does Automation Bias Decision-Making?" *International Journal of Human-Computer Studies* 51 (5): 991–1006. doi:10.1006/ijhc.1999.0252.
- Spatola, Nicolas, and Thierry Chaminade. 2022. "Cognitive Load Increases Anthropomorphism of Humanoid Robot. The Automatic Path of Anthropomorphism." *International Journal of Human-Computer Studies* 167: 1–15. doi:10.1016/j.ijhcs.2022.102884.
- Taddeo, Mariarosaria. 2010. "Modelling Trust in Artificial Agents, a First Step Toward the Analysis of e-Trust." *Minds and Machines* 20 (2): 243–257. doi:10.1007/s11023-010-9201-3.
- Tiku, Nitasha. "The Google Engineer who Thinks the Company's AI has Come to Life," *Washington Post*, June 11, 2022, <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lambda-blake-lemoine/>.
- Tucker, Patrick. "An AI Just Beat a Human F-16 Pilot in a Dogfight – Again," *Defense One*, August 20, 2020, <https://www.defenseone.com/technology/2020/08/ai-just-beat-human-f-16-pilot-dogfight-again/167872/>.
- Ullman, Shimon. 2019. "Using Neuroscience to Develop Artificial Intelligence." *Science* 363 (6428): 692–693. doi:10.1126/science.aau6595.
- Verger, Rob. "Two Fighter Pilots, Passed out Over Nevada Last Year. The Software Saved Them Both," *Popular Science*, February 22, 2021, <https://www.popsci.com/story/technology/agcas-technology-saves-pilots-lives/>.
- Watson, David. 2019. "The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence." *Minds and Machines* 29: 417–440. doi:10.1007/s11023-019-09506-6.
- Waytz, Adam, John Cacioppo, and Nicholas Epley. 2010. "Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism." *Perspectives on Psychological Science* 5 (3): 219–232. doi:10.1177/1745691610369336.
- Xie, Tianling, and Iryna Pentina. 2022. "Attachment Theory as a Framework to Understand Relationships with Social Chatbots: A Case Study of Replika." In: *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2046–2055.
- Yong, Ed. 2022. *An Immense World: How Animals Senses Reveal the Hidden Realms Around us*. New York: Knopf Canada.