

ETICA DELL'INGANNO E INTELLIGENZA ARTIFICIALE: IL CASO DELLA ROBOTICA SOCIALE

FABIO FOSSA

Dipartimento di Ingegneria Meccanica

Politecnico di Milano

fabio.fossa@polimi.it

ABSTRACT

Deception is not only a complex human phenomenon that has always interested philosophers and social scientists. It is also an opportunity widely exploited in many areas of technological design – and, in particular, in Artificial Intelligence and social robotics. This paper offers an ethical discussion on design practices aimed at maximising the quality of interactions between human and robots by deceiving users into perceiving robots as social actors. Deontological and consequentialist approaches are discussed, and their respective strengths and weaknesses are underlined. Finally, the ethical admissibility of self-deception – i.e., of consciously engaging in fictional social relationships with robots – is examined.

KEYWORDS

Ethics; Social Robotics; Deception; Self-deception; Design.

1. INTRODUZIONE

Scopo di questo saggio è di discutere il problema dell'inganno robotico da una prospettiva morale, facendo soprattutto riferimento all'ambito della robotica sociale.

Che l'inganno giochi un ruolo rilevante nella relazione tra robot – o, più in generale, sistemi di Intelligenza Artificiale – ed esseri umani è una tesi sempre più riconosciuta.¹ Nel contesto della robotica sociale, dove le interazioni tra sistema tecnologico e utente umano assumono un'importanza fondamentale, la dimensione dell'inganno è ancora più centrale.² Il fatto che le relazioni tra umani e robot comportino inevitabilmente un elemento di inganno non può che sollevare dubbi circa la loro legittimità morale. La

¹ Su ciò si veda per es. S. Natale, *Macchine ingannevoli*, Einaudi, Torino 2021; e M. Musial, *Enchanting Robots. Intimacy, Magic, and Technology*. Palgrave Macmillan, Basingstoke 2019.

² Non è possibile, per motivi di spazio, aggiungere di più sulla robotica sociale e sull'inganno come suo elemento costitutivo. Su questo e sui temi toccati solo di passaggio nei paragrafi successivi mi sia permesso rimandare a F. Fossa, *Robotica sociale: persuasione, inganno e etica del design*, in F. Fossa, V. Schiaffonati, G. Tamburrini (a cura di), *Automi e persone. Introduzione all'etica dell'intelligenza artificiale e della robotica*, Carocci, Roma 2021, pp. 175-193. Mi limito anche ad assumere che le nozioni di inganno e manipolazione, pur nelle loro differenze, trovino un'applicazione sensata nel contesto delle interazioni tra umani mediate da robot sociali e seguo il dibattito nel riconoscerne la possibile problematicità morale. Una discussione critica di questi assunti, per quanto rilevante, esula dalle intenzioni del presente lavoro.

presenza dell'inganno come cifra ineludibile delle interazioni tra umani e robot ne richiede un'attenta valutazione etica.

La questione dell'inganno in robotica sociale non può prescindere dal fatto che il veicolo della finzione sia un artefatto tecnologico le cui caratteristiche sono il risultato di un processo di progettazione. L'aspetto esteriore e le modalità di funzionamento di un robot sociale giocano un ruolo imprescindibile nelle dinamiche ingannevoli che l'artefatto suscita e alimenta. Sono le scelte – consapevoli o meno – di ingegneri e designer a determinare il potenziale d'inganno dell'artefatto. Saper dosare con efficacia questo potenziale è essenziale per costruire prodotti con cui sia facile e piacevole interagire, il che rappresenta il fine ultimo della robotica sociale.

A prima vista, il carattere ingannevole della robotica sociale potrebbe apparire angosciante da un punto di vista etico. Sarebbe però affrettato giudicare immorale qualsiasi inganno robotico suscitato *by design*. Il che vale senz'altro in alcuni casi – primo fra tutti, quando l'inganno robotico è volto ad approfittarsi delle vulnerabilità o delle reazioni istintive degli utenti a loro danno.³ È chiaro che l'inganno sia intollerabile quando le intenzioni dell'ingannatore sono malevoli e mirano a danneggiare o circuire.⁴ Come scrivono Isaacs e Bridewell, l'inganno implica “un obiettivo nascosto, un secondo fine”⁵ rispetto a quello che un'azione comunica. Quando questi contrastano con il benessere e i diritti degli utenti, è evidente che l'inganno sia inaccettabile.

La questione si fa più complessa quando si suppone che l'inganno sia a fin di bene – che, cioè, produca benefici eticamente rilevanti per gli ingannati. Ammesso che l'inganno malevolo sia moralmente spregevole, l'inganno benevolo potrebbe essere non solo legittimo, ma anche doveroso. Nel nostro caso, quindi, si tratta di chiarire se e in che misura siano eticamente accettabili inganni robotici volti a promuovere il benessere degli utenti.

L'indagine riformula così la *vexata quaestio* della filosofia morale e politica relativa alla nobile menzogna e alle condizioni in cui sia eticamente giustificabile ingannare altri esseri umani – dove però il veicolo dell'inganno è ora l'oggetto robotico e gli ingannati sono i suoi utenti. Se fosse possibile dimostrare che inganni mediati tramite robot promuovano il benessere degli utenti o della società in senso lato, argomenta ad esempio Arkin,⁶ non bisognerebbe perlomeno considerarne l'eventuale legittimità morale?

Sulla base di queste osservazioni preliminari, le prossime sezioni prendono in considerazione la questione dell'inganno robotico benevolo così come è stata affrontata nel dibattito – ovvero, secondo prospettive pressoché riconducibili ad approcci

³ Si veda per es. il caso della *hidden state deception* in J. Danaher, *Robot Betrayal: a Guide to the Ethics of Robotic Deception*. “Ethics and Information Technology”, 22 (2020), pp. 117-128, p. 121.

⁴ Cfr., e.g., W. Wallach, C. Allen, *Moral Machines. Teaching Robots Right from Wrong*. Oxford University Press, Oxford 2009, pp. 44-45; E. Adar, D.S. Tan, J. Teevan, *Benevolent Deception in Human Computer Interaction*, in *Proceedings of CHI 2013*, ACM press, New York 2013, pp. 1863-1872, p. 1865.

⁵ A.M.C. Isaac, W. Bridewell, *White Lies on Silver Tongues. Why Robots need to Deceive (and How)*, in P. Lin, R. Jenkins, K. Abney (eds), *Robot Ethics 2.0. New Challenges in Philosophy, Law, and Society*, Oxford University Press, New York 2017, pp. 157-172, p. 158. Tutte le traduzioni di brani tratti da testi in inglese sono mie.

⁶ R.C. Arkin, *Ethics of Robotic Deception [Opinion]*, “IEEE Technology and Society Magazine”, 37, n. 3, pp. 18-19.

deontologici e consequenzialisti.⁷ L'analisi ha due obiettivi principali. Innanzitutto, mostrare come la robotica sociale offra un'opportunità significativa per tornare ad esplorare un problema tanto antico quanto sempre rilevante per il pensiero filosofico. In seconda battuta, contribuire alla discussione sottolineando come le due prospettive intendano l'utente umano soprattutto in senso passivo - cioè, come vittima ignara o impotente dell'inganno mediato dai robot sociali. Che l'utente possa solo subire l'inganno robotico, tuttavia, è un'assunzione che merita di essere discussa. Come primo passo in questa direzione, l'ultima sezione del saggio considera da un punto di vista morale il caso in cui siano gli utenti a prestarsi all'inganno occasionato dall'oggetto robotico.

2. NON INGANNARE! PROSPETTIVE DEONTOLOGICHE

Il principio secondo cui inganni robotici possano essere moralmente giustificabili in ragione dei benefici che procurano all'ingannato o alla società in senso lato non è così pacifico come potrebbe sembrare. Al contrario, si può insistere sul fatto che l'inganno sia di per sé una pratica immorale, indipendentemente da ciò che ne consegua. In quanto forma di manipolazione, l'inganno - non importa a quale fine - può essere ritenuto un affronto nei confronti dell'ingannato. Più precisamente, l'inganno robotico può essere inteso come una violazione della dignità degli utenti, e quindi un atto intrinsecamente immorale.⁸

Simili posizioni fanno perno sulla componente manipolatoria dell'inganno e i suoi effetti negativi sull'esercizio dell'autonomia personale, la quale è strettamente connessa all'ideale della dignità umana.⁹ Manipolare le credenze e il comportamento degli utenti tramite robot ingannevoli è considerata una pratica degradante, che viola il diritto individuale a vivere liberamente la propria vita secondo regole e valori propri, per quanto possibile. Sebbene la capacità di autodeterminazione sia variamente circoscritta da condizioni fisiche, biologiche, psicologiche e sociali che ne descrivono l'ambito di esercizio, l'interferenza causata da robot non rientra in queste condizioni; costituisce, invece, un'intrusione illegittima in un campo di azione che merita di essere tutelato, in quanto esprime caratteri umani - autodeterminazione, autonomia, libertà, dignità - che esigono rispetto incondizionato. Qualsiasi conseguenza dell'atto ingannatorio giunge troppo tardi per essere moralmente rilevante.

⁷ La questione dell'inganno è discussa principalmente in riferimento a queste due famiglie di teorie etiche, ma il dibattito morale sulla robotica sociale adotta anche altre prospettive, tra cui spiccano l'etica della virtù e della cura. Si vedano ad esempio il recente *Social Robotics and the Good Life. The Normative Side of Forming Emotional Bonds with Machines*, a cura di J. Loh e W. Loh, transcript, Bielefeld 2023; e i lavori di A. van Wynsberghe, tra cui *Designing Robots for Care: Care Centered Value-Sensitive Design*. "Science and Engineering Ethics", 19 (2013), pp. 407-433.

⁸ Si veda ad es. A. Bertolini, *Human-Robot Interaction and Deception*, "Osservatorio del Diritto Civile e Commerciale" 2 (2018), pp. 645-659.

⁹ Per la discussione dei complessi effetti dei sistemi di IA su autonomia, identità e dignità personale si rimanda a A. Rubel, C. Castro, A. Pham, *Algorithms and Autonomy. The Ethics of Automated Decision Systems*, Cambridge University Press, Cambridge 2021; S. Tiribelli, *Identità personale e algoritmi. Una questione di filosofia morale*, Carocci, Roma 2023.

Il lavoro di Rob Sparrow ha contribuito in modo eminente ad articolare questa posizione. Tra le sue prime formulazioni spicca uno studio dedicato alle interazioni tra persone anziane e robot che simulano l'aspetto e il comportamento di animali domestici. Sparrow contesta "l'idea che tali dispositivi possano svolgere un ruolo apprezzabile nel soddisfare le esigenze di una popolazione sempre più anziana e isolata"¹⁰ facendo leva proprio sulla dimensione ingannevole delle relazioni che si instaurano tra i due termini. Infatti, spiega il filosofo, i benefici che si suppone seguano all'impiego di robot animaloidi negli ospedali e nelle strutture di assistenza per anziani, peraltro piuttosto difficili da dimostrare con certezza,¹¹ sono frutto di un design volutamente ingannevole – il che getta una luce obliqua sull'intera pratica.

Come molti altri robot sociali (umanoidi in particolare), i robot animaloidi sono esplicitamente progettati per suscitare reazioni emotive negli utenti atte a rafforzare la piacevolezza e l'accettabilità del prodotto.¹² In questo senso, scrive Henrik Skaug Sætra, esibiscono una natura parassitaria: "sfruttano e si approfittano di processi e meccanismi umani sviluppatisi nel corso dell'evoluzione per scopi del tutto alieni all'interazione" tra umani e macchine.¹³ I supposti benefici che ne possono conseguire risultano da un inganno mirato a manipolare gli stati emotivi degli utenti e, conseguentemente, la loro capacità di formare un modello mentale dell'artefatto tecnologico che sia adeguato alle effettive caratteristiche dell'artefatto stesso. L'inganno serve il secondo fine di spingere gli utenti a esperire i robot come se fossero esseri viventi e agenti sociali – e, quindi, a proiettare su di essi stati mentali, emozioni, intenzioni, e affetti. L'utente, così, è incoraggiato a stringere un profondo legame emotivo con il prodotto, il che ne assicura la fidelizzazione.¹⁴

La natura falsa, fittizia, o simulata della socialità robotica, tuttavia, ne fa un sostituto solo parziale di quella animale o umana. Ciò, secondo Sparrow, mette in dubbio la reale consistenza di eventuali effetti positivi. Ad esempio, solo gli aspetti più esterni della relazione tra umani e animali domestici si verificano anche nella relazione analoga tra

¹⁰ R. Sparrow, *The March of the Robot Dogs*, "Ethics and Information Technology" 4 (2002), pp. 305-318, p. 305.

¹¹ Sulle difficoltà intrinseche nel dimostrare effetti eticamente positivi dell'impiego di robot sociali con utenti vulnerabili, si vedano Sparrow, *The March of the Robot Dogs*, pp. 309-316; A. Sharkey, N. Sharkey, *The Crying Shame of Robot Nannies. An Ethical Appraisal*, "Interaction Studies" 112, n. 2 (2010), pp. 161-190, pp. 178-180; A. Sharkey, *Should We Welcome Robot Teachers?* "Ethics and Information Technology" 18 (2016), pp. 283-297, p. 286; C. Berridge, Y. Zhou, J.M. Robillard, J. Kaye, *Companion Robots to Mitigate Loneliness in Older Adults: Perceptions of Benefit and Possible Deception*, *Frontiers in Psychology* 14 (2023): 1106633, pp. 1-9.

¹² Per una discussione più approfondita su questo tema, cfr. F. Fossa, I. Sucameli, *Gender Bias and Conversational Agents: An Ethical Perspective on Social Robotics*, "Science and Engineering Ethics" 28 (2022), 23.

¹³ H.S. Skætra, *The Parasitic Nature of Social AI: Sharing Minds with the Mindless*, "Integrative Psychological and Behavioural Science" 54 (2020), pp. 308-326, p. 309.

¹⁴ Sulla dimensione commerciale della questione si veda per es. M. Scheutz, *The Inherent Dangers of Unidirectional Emotional Bonds Between Humans and Social Robots*, in *Robot Ethics. The Ethical and Social Implications of Robotics*, a cura di P. Lin, K. Abney, G. Bekey, MIT Press, Cambridge-London 2012, pp. 205-221. Nel prosieguo l'attenzione cadrà soprattutto su *care robots*, ma simili considerazioni valgono anche per *sex e companion robots*. Si veda per es. S. Nyholm, L.E. Frank, *It Loves Me, It Loves Me Not: Is It Morally Problematic to Design Sex Robots that Appear to Love Their Owners?*, "Techné: Research in Philosophy and Technology" 23, n. 3 (2019), pp. 402-424.

utenti e robot animaloidi, la quale è del tutto priva di reciprocità e pregnanza morale.¹⁵ La finzione in cui le persone anziane vengono irretite, e che le induce a trattare i robot animaloidi come se provassero emozioni, avessero desideri e presentassero una personalità distinta, è troppo superficiale per essere considerata una valida sostituzione del suo archetipo. La solitudine degli anziani non sarebbe così veramente mitigata, ma solo illusoriamente coperta.¹⁶

Basandosi l'intera pratica su utenti “che confondono, a livello conscio o inconscio, il robot con un animale vero” o “si illudono sistematicamente riguardo alla vera natura della loro relazione” con il prodotto tecnologico, l'inganno robotico “viola il dovere (minore) che abbiamo verso noi stessi di comprendere il mondo con precisione”¹⁷ e supporta un'illusione degradante per l'individuo surrettiziamente spinto a indugiare in essa. “Un simile inganno”, scrivono Rob e Linda Sparrow,

è riprovevole perché è improbabile che illusioni incontrino le nostre esigenze, servano i nostri interessi, o supportino il nostro benessere. (...) Noi desideriamo non tanto che le nostre credenze sul mondo, o le nostre esperienze del mondo, siano in un certo modo; ma che sia il mondo reale ad essere in un certo modo.¹⁸

Come evidenziano anche Sherry Turkle e colleghi,¹⁹ l'autenticità è cruciale nel caso di relazioni sociali moralmente significative. Robot progettati per apparire ciò che non sono possono sì rendere le interazioni più facili e fluide, ma non ambire all'autenticità e alla pregnanza che contraddistinguono le relazioni sociali. Ciò non può che andare a discapito della dignità degli utenti, indotti a un “sentimentalismo moralmente deprecabile”²⁰ che ricorda casi di infatuazione, ossessione o allucinazione. “Nella misura in cui possono rendere le persone più felici solo quando vengono ingannate su ciò che realmente essi sono”, concludono Rob e Linda Sparrow, “i robot non possono offrire miglioramenti reali al benessere degli utenti”.²¹

Questa prospettiva non lascia spazio alcuno alla considerazione delle conseguenze come elemento rilevante per la valutazione morale. Qualsiasi stato di cose risulti dalla relazione è già contaminato da un inganno moralmente inaccettabile. L'idea stessa che il contatto con animali o esseri umani possa essere fruttuosamente simulato tramite robot sociali è illusoria, e solo tramite inganni può ottenere una parvenza di realtà. Alimentare

¹⁵ R. Sparrow, L. Sparrow, *In the Hands of Machines? The Future of Aged Care*, “Minds and Machines” 16 (2006), pp. 141-161. Su ciò, e in relazione alla robotica sociale in generale, insiste anche A. van Wynsberghe, *Social Robots and the Risks to Reciprocity*. “AI & Society” 37 (2022), pp. 479-485.

¹⁶ Una simile argomentazione si trova per es. anche in J.P. Sullins, *Robots, Love, and Sex: the Ethics of Building a Love Machine*, “IEEE Transactions on Affective Computing” 3, n. 4 (2012), pp.398-409.

¹⁷ Sparrow, *The March of the Robot Dogs*, p. 306.

¹⁸ Sparrow, Sparrow, *In the Hands of Machines?*, p. 155.

¹⁹ S. Turkle, W. Taggart, C.D. Kidd, O. Dasté, *Relational Artifacts with Children and Elders: the Complexity of Cybercompanionship*, “Connection Science” 18, n. 4 (2006), pp. 347-361, p. 360.

²⁰ Sparrow, *The March of the Robot Dogs*, p. 315.

²¹ Sparrow, Sparrow, *In the Hands of Machines?*, p. 155. Simili argomenti si ritrovano anche nel recente E. Portacolone, J. Halpern, J. Luxenberg, K.L. Harrison, K.E. Covinsky, *Ethical issues raised by the introduction of artificial companions to older adults with cognitive impairment: a call for interdisciplinary collaborations*, “Journal of Alzheimer's Disease” 76 (2020), pp. 445-455.

l'illusione *by design*, dunque, è eticamente discutibile, quali che siano gli effetti che si suppone debbano seguirne.

Alla luce di quanto detto, appare evidente che posizioni come quelle di Sparrow e Turkle escludano l'ammissibilità morale di strategie di progettazione volte a ingannare gli utenti circa le effettive capacità sociali dei robot. Anche se sarebbe eccessivo decretare l'assoluta immoralità di intrattenere relazioni simil-sociali con robot, si può invece concludere che accrescere la piacevolezza dei robot sfruttando le tendenze zoomorfiche degli utenti sia irrispettoso della loro dignità e, come tale, rappresenti una pratica condannabile dal punto di vista etico e nociva dal punto di vista sociale.²² Ciò vale in particolare qualora interessi soggetti vulnerabili – come, appunto, persone anziane fortemente bisognose di contatto umano e, di conseguenza, più esposte al potenziale ingannatore della robotica.

L'insistenza sul dovere di rispettare l'autonomia, e quindi la dignità,²³ degli utenti della robotica sociale contro ogni inganno rivela la radice deontologica che accomuna questi approcci. La matrice kantiana dell'argomentazione è d'altronde evidente, e a volte esplicita. Nel riassumere i punti chiave dell'argomentazione, Rob e Linda Sparrow scrivono che

ingannare intenzionalmente altre persone, fosse anche per il loro stesso bene, è immorale” in quanto “corrisponde a trattarli come oggetti da manipolare per servire i nostri fini – anche nel caso in cui il nostro fine sia la loro felicità; il che viola il dovere kantiano fondamentale di trattare gli altri con rispetto, in quanto ‘fini in se stessi’.”²⁴

Il richiamo al noto imperativo categorico di Kant²⁵ è più di una semplice citazione. Piuttosto, rivela lo sfondo filosofico dell'argomentazione. Il ragionamento riposa sulla tesi kantiana secondo cui ogni essere razionale è, in quanto tale, un essere autonomo – la cui volontà è cioè capace di determinarsi liberamente secondo la rappresentazione dei principi pratici della ragione. Gli esseri razionali, in quanto portatori della capacità di autonomia morale, hanno valore in sé e come tali esigono di essere trattati – condizione a cui si riferisce il concetto di dignità personale. Nelle parole di Kant, “l'*autonomia* è dunque il fondamento della dignità della natura umana e di ogni natura razionale”²⁶, e la dignità personale fonda il diritto e il dovere corrispondente del rispetto: questo “principio dell'umanità e di ogni natura razionale in generale *come fine in sé stesso*” vale da “suprema condizione limitativa della libertà delle azioni”²⁷ di ogni essere umano. Condizione limite che vale anche nel caso degli utenti della robotica sociale, ma viene violata da strategie di design che ne disconoscono l'autonomia e li trattano alla maniera di mezzi.

Prima di passare alla considerazione di posizioni opposte, è interessante notare come la discussione circa la liceità morale dell'inganno a fin di bene abbia sempre rappresentato

²² H.S. Sætra, *Social Robot Deception and the Culture of Trust*, Paladyn 12 (2021), pp. 276-286

²³ Si veda anche H.S. Sætra, *First, They Came for the Old and Demented: Care and Relations in the Age of Artificial Intelligence and Social Robots*, “Human Arenas” 5 (2022), pp. 25-43.

²⁴ Sparrow, Sparrow, *In the Hands of Machines?*, pp. 155-156.

²⁵ I. Kant, *Fondazione della metafisica dei costumi*, Laterza, Roma-Bari 2007, p. 91.

²⁶ *Ivi*, p. 105.

²⁷ *Ivi*, p. 95.

una questione spinosa per l'etica di ispirazione kantiana.²⁸ La questione era chiara a Kant stesso, che in alcune pagine molto note prese posizione sulla tesi che mentire²⁹ o fare una promessa falsa³⁰ potesse essere moralmente lecito qualora evitasse conseguenze deprecabili o per motivi filantropici. Formulando la questione nei termini dell'imperativo categorico, ingannare altri soggetti morali dando loro ad intendere che un dato stato di cose non sia come effettivamente è implica servirsene come meri mezzi a nostro fine e, dunque, venire meno al dovere assoluto di rispettare la loro natura razionale – oltre che degradare se stessi in quanto agenti morali.³¹ L'inganno lede inevitabilmente l'autonomia dell'ingannato, e dunque la sua dignità: nessuna supposta conseguenza benefica può riparare a un simile danno.

3. IL BENE DELL'INGANNATO: PROSPETTIVE CONSEQUENZIALISTE

Una linea di pensiero tanto rigida non può che suscitare reazioni polemiche volte a superarne il carattere “estremo”³², se non addirittura “puritano”.³³ Emerge così la necessità di elaborare un approccio all'etica dell'inganno robotico che sia meno pregiudizievole, più in grado di cogliere le diverse sfaccettature del fenomeno. Il roboticista Ron Arkin,³⁴ ad esempio, è convinto che l'arte di creare illusioni per mezzo di robot sociali rappresenti una risorsa chiave per la realizzazione di sistemi in grado di portare benefici sostanziali a chi li impiega o usa. Sebbene ammetta che alcune forme di progettazione ingannevole siano inequivocabilmente immorali, allo stesso tempo reputa che altre forme possano apportare più vantaggi che danni agli utenti. Sarebbe miope chiudere gli occhi di fronte a ciò che l'inganno robotico potrebbe renderci capaci di conseguire. Lo spazio per un uso virtuoso dell'inganno merita di essere preservato.³⁵

D'altronde, una posizione tanto rigida sul rapporto tra elementi di progettazione e illusioni zoomorfiche sembra peccare di astrattezza. Si considerino, ad esempio, robot sociali pensati per la cura dei bambini. È certamente vero che la robotica può contare su un insieme senza precedenti di mezzi per suscitare illusioni di socialità – caratteri corporei, linguaggio parlato, riconoscimento delle espressioni e delle emozioni, e così via. Tuttavia, è anche vero che tendenze zoomorfiche appartengono naturalmente alla psicologia

²⁸ Cfr. per es. A. Isenberg, *Deontology and the Ethics of Lying*, “Philosophy and Phenomenological Research” 24, n. 4 (1964), pp. 463-480.

²⁹ I. Kant, *Su un preteso diritto di mentire per amore degli uomini*, in Id., *Scritti di storia, politica e diritto*, Laterza, Roma-Bari 2007, pp. 209-214.

³⁰ Kant, *Fondazione*, p. 93.

³¹ Come afferma J. Vuillemin, *On Lying: Kant and Benjamin Constant*, “Kant Studien” 73, n. 1-4 (1982), pp. 413-424, p. 414, mentire “distrugge la propria dignità” “qualsiasi risultato esterno possa produrre”.

³² A. Sharkey, N. Sharkey, *Children, the Elderly, and Interactive Robots. Anthropomorphism and Deception in Robot Care Companionship*, “IEEE Robotics and Automation Magazine”, March 2011, pp. 32-38, p. 34.

³³ Wallach, Allen, *Moral Machines*, p. 44.

³⁴ A.R. Wagner, R.C. Arkin, *Acting Deceptively: Providing Robots with the Capacity for Deception*, “International Journal of Social Robotics” 3 (2011), pp. 5-26.

³⁵ Per una discussione della questione della menzogna virtuosa si veda ad es. W. Simon, *Virtuous Lying: A Critique of Quasi-Categorical Moralism*, “Georgetown Journal of Legal Ethics” 12(1999), pp. 432-463.

dell'infanzia e, in certa misura, rappresentano disposizioni innocue, se non produttive. “Se è il naturale antropomorfismo dei bambini a ingannarli”, scrivono Amanda e Noel Sharkey, “non dovrebbero sorgere preoccupazioni morali circa le scelte di roboticisti o produttori”.³⁶ Se poi si considera come queste tendenze costituiscano un tratto profondamente radicato nella psicologia umana in generale,³⁷ appare chiaro quanto contrastarle possa essere vano, se non controproducente. L'inganno non è un ingrediente che possa essere aggiunto o sottratto a piacimento dalla relazione tra utenti e robot. Ne è una componente “strutturale”.³⁸

Dal momento che prendono per una variabile quella che è una condizione delle relazioni tra umani e robot, posizioni deontologiche risultano inadeguate a impostare un discorso rigoroso sull'inganno in robotica sociale. Amanda e Noel Sharkey riconoscono quanto sia “arduo adottare un approccio etico assolutista per discutere questioni di robot e inganni”.³⁹ “Piuttosto che opporsi a tutti questi usi”, osservano, “ha più senso focalizzare la nostra attenzione morale su quelle situazioni in cui è probabile che design antropomorfi generino effetti negativi per il benessere umano”.⁴⁰ Se l'inganno non può essere evitato, la sua legittimità etica deve essere decisa in funzione di benefici e danni potenziali: ciò che conta in definitiva è “l'impatto probabile su individui e società, indipendentemente dalle intenzioni benefiche o malevoli di chi inganna”.⁴¹

Ammettere l'inganno come condizione costitutiva delle relazioni tra umani e robot pone il dovere di gestirne gli effetti responsabilmente, il che inclina verso approcci di carattere consequenzialista.⁴² In questo quadro, l'idea di ingannare per un bene maggiore o per il bene dell'ingannato è meno problematica. Legittimando azioni o regole pratiche sulla base della loro capacità di massimizzare uno o più valori predefiniti (per esempio, il benessere umano), il pensiero consequenzialista apre a pratiche ingannevoli che concorrano a massimizzare quei valori. Se fosse dimostrabile che migliorare le interazioni tra umani e robot sociali ricorrendo a strategie di design ingannevoli abbia effetti positivi sul benessere umano, approcci consequenzialisti ne accetterebbero la legittimità. Condannerebbero, invece, casi in cui l'inganno comporti l'emergere di situazioni in cui il benessere umano è ridotto.

³⁶ Sharkey, Sharkey, *The Crying Shame...*, p. 172.

³⁷ Per simili considerazioni nell'ambito di studi sulla robotica si vedano ad es. S.R. Fussel, S. Kiesler, L.D. Setlock, V. Yew, *How People Anthropomorphize Robots*, in *HRI'08, March 12-15, 2008*, ACM, Amsterdam (2008), pp. 145-152; J. Fink, *Anthropomorphism and Human Likeness in the Design of Robots and Human Robot Interaction*, in *ICSR 2012, LNAI 7621*, a cura di S.S. Ge et al., Springer Verlag, Berlin Heidelberg 2012, pp. 199-208; B.R. Duffy, *Anthropomorphism and the social robot*, “Robotics and Autonomous Systems” 42 (2003), pp. 177-190.

³⁸ Natale, *Macchine ingannevoli*, p. 71.

³⁹ Sharkey, Sharkey, *The Crying Shame...*, p. 173.

⁴⁰ Sharkey, Sharkey, *Children, the Elderly...*, p. 35; si veda anche A. Matthias, *Robot Lies in Healthcare: When is Deception Morally Permissible?*, “Kennedy Institute of Ethics Journal”, 25, n. 2 (2015), pp. 183-186.

⁴¹ A. Sharkey, N. Sharkey, *We Need to Talk About Deception in Social Robotics!*, “Ethics and Information Technology” 23 (2021), pp. 309-316, p. 311.

⁴² Si vedano ad es. i già citati Wagner, Arkin, *Acting Deceptively*, e Isaac, Bridewell, *White Lies...*

Senza dubbio, gli approcci consequenzialisti reagiscono meglio alle sollecitazioni della robotica. Adar e colleghi⁴³ sottolineano come inganni privi di intenti malevoli siano ampiamente impiegati nel design di molti prodotti tecnologici. Scelte di progettazione ingannevoli mirate a facilitare le interazioni tra umani e tecnologie sono così diffuse che una loro condanna morale sarebbe assai complessa da applicare. Nel caso della robotica sociale, imitare comportamenti ingannevoli ma prosociali come “scambiare convenevoli in modo insincero”, ricorrere a “metafore e iperboli”, o “controllare le impressioni da suscitare”⁴⁴ potrebbe agevolare notevolmente l'integrazione sociale dei robot senza provocare danni morali evidenti. Nella misura in cui l'utente stesso potrebbe preferire “l'esperienza offerta dall'interfaccia ingannevole rispetto a quella offerta dall'interfaccia ‘onesta’”, sembra opportuno lasciare che l'inganno faccia parte delle risorse a disposizione dei roboticisti.⁴⁵

Se molti apprezzano l'apertura che contraddistingue gli approcci consequenzialisti, non si dà però consenso sui valori da massimizzare, i loro rapporti gerarchici, e le modalità della loro misurazione – il che può portare a conclusioni contrarie a quelle augurate. Secondo Sparrow,⁴⁶ ad esempio, l'analisi degli effetti dell'impiego di robot ingannevoli in strutture di cura per anziani evidenzia benefici limitati e rischi significativi se si considerano i valori della dignità e dell'autonomia. Posizioni più sfumate sono però abbastanza comuni. Tra gli altri, Matthias contesta rigide associazioni tra inganno e violazioni di dignità e autonomia, invitando a valutare caso per caso. Difatti, alcuni pazienti potrebbero preferire a un assistente robotico che “dice sempre e solo la verità”, un sistema che “fornisca conforto e sostenga il benessere e il processo di guarigione del paziente” – il che non comporterebbe alcuna “violazione della fiducia nella relazione tra paziente e assistente robotico”.⁴⁷ Nel delicato contesto della cura, piccole bugie a fin di bene rappresentano uno strumento relazionale comune.⁴⁸ Robot destinati a essere impiegati in questo contesto non potrebbero abitarlo correttamente, se non fossero in grado di ingannare per il bene dei pazienti. In modo simile, inganni robotici che limitano l'autonomia degli utenti non

⁴³ Adar et al., *Benevolent Deception...* Per una disamina critica delle diverse forme di consequenzialismo che possono essere esplorate per giustificare l'inganno robotico si veda Sætra, *First, They Came for the Old and Demented*.

⁴⁴ Isaac, Bridewell, *White Lies...*, p. 165.

⁴⁵ Adar et al., *Benevolent Deception...*, p. 1865. Si veda su questo anche M.M.A. de Graaf, *An Ethical Evaluation of Human-Robot Relationships*, “International Journal of Social Robotics” 8, n. 4 (2016), pp. 589-598.

⁴⁶ Sparrow, *The March of the Robot Dogs*; Sparrow, Sparrow, *In the Hands of Machines?*.

⁴⁷ Matthias, *Robot Lies...*, p. 179. Su questo si veda anche, ad es., R.N. van Wyk, *When is Lying Morally Permissible? Casuistical Reflections on the Game Analogy, Self-defense, Social Contract Ethics, and Ideals*, “The Journal of Value Inquiry” 24 (1990), 155-168.

⁴⁸ Secondo Matthias, *Robot Lies...*, p. 182, “se un certo grado di inganno è necessario per salvaguardare il benessere di chi si affida al medico, il primato del suo benessere sulla veridicità giustificerebbe l'inganno e, entro i limiti strettamente necessari a mantenerne il benessere, impedirebbe a un tale inganno di costituire una violazione della fiducia.” Per ulteriori discussioni su veridicità, cura e robotica si veda S. Segers, *Robot Technology for the Elderly and the Value of Veracity: Disruptive Technology or Reinvigorating Entrenched Principles?*, “Science and Engineering Ethics” 28 (2022), 64, pp. 1-14 e M. Coeckelbergh, *Care Robots and the Future of ICT-mediated Elderly Care: a Response to Doom Scenarios*, “AI & Society” 31 (2016), pp. 455-462.

possono essere dichiarati senz'altro immorali, in quanto potrebbero massimizzare beni gerarchicamente più rilevanti o complessi, come il loro benessere individuale. Se l'inganno robotico mira alla massimizzazione di valori più rilevanti dell'autonomia, una sua contestuale (e parziale) limitazione non sembra sufficiente per decretarne la condanna.

La flessibilità degli approcci consequenzialisti ne determina anche la complessità. Come dimostrano i numerosi contributi di Amanda e Noel Sharkey, casi diversi chiamano in causa valori diversi, richiedono considerazioni diverse, e motivano valutazioni diverse. Sebbene i due autori apprezzino il potenziale etico dell'inganno robotico, negli scritti sulla robotica assistenziale per bambini e anziani i rischi sembrano loro più preoccupanti dei possibili benefici. Ciò avviene principalmente per due motivi: la vulnerabilità dei suddetti utenti e le aspettative che il pubblico è portato a sviluppare circa le capacità della tecnologia. Ad esempio, la possibilità di lasciare i bambini nelle mani di dispositivi robotici che siano davvero capaci di “sorveglianza costante, intrattenimento e compagnia”⁴⁹ è ancora lontana, ma potrebbe essere suggerita ingannevolmente da scelte di design dedicate – il che potrebbe causare danni eticamente significativi. E se anche il mezzo tecnico fosse adeguato, la pratica in sé non sarebbe scevra da rischi: si pensi ad esempio a potenziali violazioni della privacy, eventuali coercizioni fisiche,⁵⁰ e agli impatti ancora ignoti sullo sviluppo psicologico dei bambini.

In conclusione, l'apertura propria dell'approccio consequenzialista a pratiche di inganno robotico sembra valere tanto quanto la sua capacità di discriminare tra usi benefici e dannosi. La legittimità morale di inganni robotici è da valutare caso per caso. Quando strategie di design ingannevoli vengono messe in atto per “indurre un bambino” o altri utenti “a stringere una relazione falsa con un robot”, a credere che questo “si prenda cura di loro”, tanto da sentirsi “ansiosi o tristi quando il robot è assente” o da “trascorrere tempo con il robot preferendolo ai loro coetanei umani”⁵¹, la pratica dovrebbe essere considerata immorale sia in quanto motivata da intenti malevoli, sia per i probabili danni psicologici che ne seguirebbero. Quando le stesse strategie servono il benessere degli utenti, per quanto ingannati, non dovrebbero essere considerate moralmente riprovevoli. Se gestite responsabilmente, al contrario, possono rappresentare risorse morali importanti. Ad esempio, l'impiego di robot sociali con persone anziane, saggiamente bilanciato con occasioni di cura e contatto umano, può accrescerne significativamente il benessere e costituire un caso di inganno buono, offrendo preziose occasioni per esercitare abilità mentali, linguistiche e sociali.⁵² Per quanto ponderare rischi e benefici sia un compito molto difficile, è anche un compito che, se condotto con attenzione, può aiutare a sfruttare appieno il potenziale benefico della robotica sociale.

Sebbene approcci consequenzialisti offrano una prospettiva più articolata del problema qui in discussione, non sono certamente privi di criticità. Come per altre specificazioni di questa famiglia di dottrine morali, risulta difficile fornire sia metodologie oggettive, quasi-

⁴⁹ Sharkey, Sharkey, *The Crying Shame...*, p. 163.

⁵⁰ *Ivi*, p. 166.

⁵¹ *Ivi*, p. 173.

⁵² Sharkey, Sharkey, *Granny and the Robots: Ethical Issues in Robot Care for the Elderly*, “Ethics and Information Technology” 12 (2012), pp. 27-40.

quantitative, per misurare l'impatto delle diverse tecnologie, sia quadri per pesare le conseguenze l'una contro l'altra – conseguenze la cui previsione affidabile è di per se stessa un problema non indifferente, ulteriormente aggravato dall'incertezza tipica dei processi di innovazione tecnologica. Ma soprattutto rimane difficile comprendere come siano da determinare concetti generali quali il benessere degli utenti o i loro interessi – e, soprattutto, a chi spetti determinarli.⁵³ Difatti, simili posizioni sembrano implicare una legittimazione morale di pratiche a prima vista paternalistiche.⁵⁴

Queste ultime considerazioni suggeriscono che l'utente finale non sia ritenuto occupare una posizione tale per cui il perseguimento del proprio benessere per mezzo dell'oggetto robotico sia alla sua portata. La figura dell'utente è avvolta nei panni passivi dell'ingannato contro la propria volontà e privo di consapevolezza. Mettere in discussione l'idea di utente che sembra sottesa non solo alle considerazioni consequenzialiste, ma anche a quelle deontologiche, può però aiutare a gettare una luce più adeguata sulle dinamiche dell'inganno robotico e la relativa indagine morale.

4. SOSPENSIONE VOLONTARIA DELL'INCREDELITÀ: ROBOTICA E AUTOINGANNO

Le prospettive prese in esame finora condividono l'assunzione che l'illusione imbastita dalla robotica sociale riservi un ruolo puramente passivo agli utenti. La necessità di una riflessione morale sull'inganno robotico nasce dalla constatazione che gli utenti ne sono inconsapevolmente esposti, il che rende necessario uno sforzo di tutela dei loro diritti e del loro benessere.

Sarebbe però limitante ridurre l'inganno robotico alla celata manipolazione delle scelte e dei comportamenti altrui, sia questa malevola o benevola. Non si deve ignorare che in molte occasioni gli esseri umani si lasciano volontariamente coinvolgere in pratiche ingannevoli o illusorie di carattere estetico o ludico.⁵⁵ Come nel gioco l'immaginazione dei bambini vivifica attivamente oggetti inanimati, si potrebbe sostenere, l'immaginazione degli utenti può lasciarsi inconsapevolmente coinvolgere in gratificanti illusioni ludiche occasionate da robot sociali.⁵⁶ Così intese, strategie di design ingannevoli offrirebbero lo spunto per sperimentare relazioni immaginarie o fittizie con oggetti robotici ed esplorarne il potenziale cognitivo e emotivo. Non mirerebbero, quindi, ad esercitare un effetto coercitivo su un utente passivo e ignaro, ma a stimolarne la fantasia e la capacità di finzione ludica.

⁵³ Si veda per es. Isenberg, *Deontology and the Ethics of Lying*, B. Whitby, *Do You Want a Robot Lover?*, in *Robot Ethics*, a cura di P. Lin et al., pp. 233-248.

⁵⁴ Sul problema del paternalismo in etica della tecnologia si vedano per es. C. Sunstein, *The Ethics of Nudging*, "Yale Journal on Regulation" 32, pp. 413-450; J. Millar, *Technology as moral proxy: Autonomy and paternalism by design*, in *2014 IEEE International Symposium on Ethics in Science, Technology and Engineering*, IEEE Press, Chicago 2014, pp. 1-7.

⁵⁵ Sull'aspetto ludico delle interazioni tra umani e tecnologie medialti insiste giustamente Natale, *Macchine ingannevoli*.

⁵⁶ Sharkey, Sharkey, *Granny and the Robots*, p. 36.

Intendere l'utente come parte attiva dell'inganno robotico conduce alla questione dell'autoinganno. Con il termine autoinganno si intende qui un'esperienza in cui l'ingannato è consapevole a livello cognitivo della sua natura illusoria, ma sceglie deliberatamente di sospendere questa consapevolezza e lasciarsi coinvolgere in vista di una potenziale gratificazione. In robotica, l'adozione di strategie di design ingannevoli atte a originare e supportare esperienze illusorie di socialità potrebbe anche essere compresa secondo queste linee generali – cioè, come la predisposizione di inviti ad atti di “collusione volontaria”⁵⁷ da parte degli utenti che si situano al di là di manipolazioni furtive e inganni in senso forte.

Come quest'ultima considerazione suggerisce, sembra ora opportuno riformulare il problema in termini meno aspri, spostando il baricentro dall'inganno comunemente inteso (cioè, come atto d'offesa) alla dimensione più sfumata della finzione.⁵⁸ Dal momento che il soggetto s'impegna attivamente in un processo di “sospensione volontaria dell'incredulità”⁵⁹ che costituisce un dominio d'esperienza attiguo a quello reale e regolato da logiche proprie, parlare di inganno vero e proprio sembra inadeguato.⁶⁰ “Più che costituire un inganno,” scrive Matthias, lasciarsi coinvolgere nell'illusione della robotica sociale dà corso a un'esperienza analoga alla “‘credenza’ condizionale e fortemente limitata nei ‘fatti’ di un universo di finzione”.⁶¹

Ciò non significa ovviamente che indagini etiche non siano necessarie. Per quanto richieda considerazioni diverse, anche l'autoinganno occasionato dai robot sociali solleva questioni importanti. A partire dal fatto che non è sempre facile distinguere tra autoinganno consapevole e inganno effettivo. La consapevolezza del soggetto che fa esperienza dell'illusione è fondamentale perché questo non si tramuti in inganno. Che si dia effettivamente la possibilità di mantenere attiva la consapevolezza della finzione durante l'interazione con i robot sociali, senza però dissiparne la magia, è oggetto di disaccordo.

In verità, il dato della consapevolezza risulta cruciale per dirimere problemi di inganno e finzione sin dai tempi della filosofia classica greca.⁶² In questo senso, il contesto della robotica sociale richiama linee argomentative che hanno dato forma alla diatriba sul valore dell'arte mimetica nella *polis*. Anche in questo caso è stato riconosciuto come determinante il fatto che l'illusione eserciti la sua influenza a livello inconscio o conscio. Nella Repubblica di Platone,⁶³ com'è noto, l'attività mimetica – e in particolare l'arte – è

⁵⁷ *Ibid.*

⁵⁸ Si veda su ciò il recentissimo P. Sweeney, *Social Robots: A Fictional Dualism Model*, Rowman & Littlefield, London-New York 2023.

⁵⁹ L'espressione traduce la formula inglese *willing suspension of disbelief* introdotta da S.T. Coleridge, *Biographia Literaria*, Rest Fenner, Londra 1817 e variamente ripresa nel dibattito su robotica sociale e inganno.

⁶⁰ Cfr. van Wyk, *When is Lying...*, pp. 156-157; D. Fallis, *Lying and Deception*, “Philosopher's Imprint” 10, n. 11 (2010), pp. 1-22.

⁶¹ Matthias, *Robot Lies...*, p. 177.

⁶² Per una discussione più approfondita su come rileggere Platone e Aristotele possa aiutare a chiarire problemi legati a inganno e nuove tecnologie si veda M. Coeckelbergh, *Moved by Machines. Performance Metaphors and Philosophy of Technology*, Routledge, New York 2019.

⁶³ Plato, *Resp.* III, X.

fonte di sospetto in quanto spaccia per vera una copia sfocata (l'opera d'arte) della copia sfocata (il mondo materiale) di ciò che è veramente vero (le forme), e in definitiva, avvolge i cittadini in una rete di falsità e illusioni inautentiche. Prendere consapevolezza dell'inganno della *mimesis* è arduo, ma una volta che lo si fa e se ne realizzano i rischi, l'intera pratica sembra perdere qualsiasi attrattiva. E tuttavia, Platone sembra pensare che si possa dare un uso buono dell'illusione mimetica, qualora sia ancorata ad archetipi degni di imitazione e contribuisca a raggiungere i fini della città – l'uso che Platone stesso ne ha fatto, e che sembra allineato alle note tesi sulla nobile menzogna.⁶⁴ In ogni caso, o l'arte inganna inconsapevolmente, o la si usa per ingannare consapevolmente.

Al contrario, nelle pagine dedicate all'arte tragica Aristotele si impegna a distinguere tra finzione e inganno, insistendo sul fatto che la *mimesis* non ambisce alla duplicazione ingannevole del reale, ma a una sua rappresentazione drammatica secondo il possibile e il necessario che lo spettatore è invitato a esperire come se fosse reale pur rimanendo consapevole del suo carattere di finzione.⁶⁵ La natura mimetica delle rappresentazioni artistiche occasiona un'attività sociale, intenzionale e basata sull'autoinganno che permette l'esplorazione innocua di passioni dirompenti – da cui l'effetto catartico che il filosofo ha riconosciuto al dramma tragico.⁶⁶ Finché la consapevolezza della finzione rimane salda nello spettatore, sembra suggerire Aristotele, non c'è ragione di parlare di inganno in senso forte e moralmente negativo. L'illusione consapevole offre possibilità di esperienza valide e meritevoli di essere perseguite in vista dei benefici individuali e sociali che ne possono seguire.

Le discussioni intorno all'autoinganno in robotica sociale ricordano per molti versi il dibattito antico sul valore morale e sociale delle esperienze fittizie. La convergenza di diffuse mitologie sulle macchine senzienti⁶⁷ ed efficaci strategie di design ingannevoli può mettere a dura prova la consapevolezza della finzione zoomorfica – specialmente in casi in cui varie condizioni di fragilità cognitiva ed emotiva spingano ad abbandonarsi pienamente ad essa.⁶⁸ La sospensione dell'incredulità implicata nell'impegnarsi attivamente in pratiche illusorie è pur sempre un atto cognitivo delicato. Ad esempio, la capacità di un robot sociale di funzionare “anche quando il bambino è da solo con esso”⁶⁹ potrebbe rompere la sottile membrana della finzione e indurre il bambino a sviluppare relazioni zoomorfiche unilaterali – e, perciò, dannose – con entità robotiche. Il già ricordato radicamento psicologico delle tendenze zoomorfiche mette in dubbio la possibilità effettiva di intrattenere illusioni consapevoli nelle interazioni con robot sociali, per cui ad essere illusoria sarebbe in verità la pretesa di rimanere in controllo

⁶⁴ Plato, *Resp.* 414c-415d.

⁶⁵ Arist., *Ret.* 1451a38.

⁶⁶ Arist., *Ret.* 1449b25-30.

⁶⁷ Si veda tra gli altri il bel libro di A. Mayor, *Gods and Robots. Myths, Machines, and Ancient Dreams of Technology*, Princeton University Press, Princeton-Oxford 2018.

⁶⁸ Per es., Sharkey, Sharkey, *Granny and the Robots*, p. 35; Sharkey, Sharkey, *We Need to Talk About Deception...*

⁶⁹ Sharkey, Sharkey, *The Crying Shame...*, p. 167.

dell'illusione.⁷⁰ Sebbene non sia da escludere che “alcune forme di autoinganno possano promuovere i nostri interessi e forse persino essere virtuose”, l’instaurazione di legami sociali con oggetti robotici sembra presupporre l’influenza di “credenze erranee a livello inconscio”⁷¹ che non possono essere pienamente controllate in modo consapevole.

Se la sospensione volontaria dell’incredulità rappresenti o meno una possibilità effettiva nelle interazioni tra umani e robot sociali è in ultima analisi una questione empirica – per quanto, ovviamente, fornire prove in questo senso vada incontro a difficoltà considerevoli. Sembra però ragionevole ammettere la possibilità che un’esposizione strutturata a robot sociali progettati con cura possa supportare gli utenti nella costituzione di esperienze illusorie ludiche che non minino al contempo la loro consapevolezza dei caratteri reali della tecnologia – in cui, cioè, la natura fittizia della socialità robotica sia allo stesso tempo chiara ma anche facile da sospendere.

Data la pervasività dell’inganno nelle interazioni tra umani e tecnologie, approfondire lo studio delle condizioni che rendano possibile la sospensione volontaria dell’incredulità e il godimento dei benefici che l’interazione con robot sociali può offrire, pur nella sua differenza e specificità rispetto al contatto umano, rappresenta una fruttuosa direzione di indagine per l’etica della robotica sociale, la quale a sua volta può gettare nuova luce su un dibattito che appartiene a pieno titolo alla tradizione del pensiero filosofico. L’inganno robotico, così, rappresenta sia un’opportunità che una sfida per la filosofia morale – non solo per continuare a riflettere su alcuni dei suoi problemi più discussi, ma anche per trattenere il pensiero presso la dimensione concreta del fenomeno etico, a cui in definitiva appartiene e che deve abitare.

⁷⁰ Di questo parere sembra essere Sætra (*Social Robot Deception...*, p. 282; *The Parasitic Nature...*, p. 320) quando scrive che il ruolo dell’inconscio nelle relazioni tra umani e robot è troppo rilevante per ammettere senz’altro esperienze comuni di autoinganno consapevole e volontario.

⁷¹ Sparrow, *The March of the Robot Dogs*, p. 315.