The Lovelace effect: Perceptions of creativity in machines

DOI:10.1177/14614448221077278

(Article begins on next page)

The Lovelace Effect: Perceptions of Creativity in Machines

Journal:	*New Media and Society*
Manuscript ID	NMS-20-1531.R1
Manuscript Type:	Original Manuscript
Keywords:	Artificial Intelligence, Computational Creativity, Ada Lovelace, computer-generated art, Human-Machine Communication, Turing Test, Lovelace Objection, artificial creativity, AICAN

Abstract:	This article proposes the notion of the 'Lovelace Effect' as an analytical tool to identify situations in which the behavior of computing systems is perceived by users as original and creative. It contrasts the Lovelace Effect with the more commonly known 'Lovelace objection', which claims that computers cannot originate or create anything, but only do what their programmers instruct them to do. By analysing the case study of AICAN - an AI art-generating system - we argue for the need for approaches in computational creativity to shift focus from what computers are able to do in ontological terms to the perceptions of human users who enter into interactions with them. The case study illuminates how the Lovelace effect can be facilitated through technical but also through representational means, such as the situations and cultural contexts in which users are invited to interact with the AI.

The Lovelace Effect: Perceptions of Creativity in Machines

In computer science circles, the phrase 'Lovelace objection' indicates the claim that computers cannot originate or create anything, but only do what their programmers instruct them to do (Abramson, 2008). Today, most computer scientists dismiss this objection: the complexity of contemporary systems and advances in areas like machine learning have proven that computer software can develop in ways that cannot be always anticipated by programmers (Kelleher, 2019). Yet the Lovelace objection is routinely mentioned in discussions of computational creativity and artificial intelligence (AI) (e.g. Gunkel, 2021). Some, however, have argued that the objection was never actually made by its supposed source and namesake Ada Lovelace, but originated from a misreading of her argument (Ward 2020; Green, 2001). This paper draws from this ongoing discussion to put forth a different assessment and application of Lovelace's contribution. It proposes the notion of the 'Lovelace Effect' to describe situations in which the behavior of computing systems is *perceived by users* as original and creative. The concept of the Lovelace effect recalibrates Lovelace's contribution by shifting the focus from what computers are able to do in ontological terms to the reactions and perceptions of human users who enter into interactions with them.

Our proposal feeds into recent conversations in areas such as computational creativity and human-machine communication (HMC) that highlight the role of human users and observers in framing computer behaviors. In the field of computational creativity, scholars have increasingly refused an ontological definition of creativity, focusing instead on the goal of programming computing systems that observers deem creative (Boden, 2004). In HMC, social and cultural studies of communication between humans and artificial systems have highlighted that meaning is attributed by human users who project agency, feelings, and creativity onto machines (Guzman, 2018). Such emerging approaches provide a theoretical framework for this paper, as well as a powerful incitation to reconsider the so-called Lovelace objection through a new lens. Ongoing work in AI and computation, however, often fail to consider creativity from this perspective. Creativity is still often discussed, in popular as well as in academic circles, as a quality of the internal functioning of the machine (Elgammal, 2018), rather than as an attribution of human users that might be stimulated also through non-technical elements such as context, cultural expectations, and social dynamics (Natale, 2021). Tackling this problem, the Lovelace effect will provide a powerful analytical tool to help recognize such attributions as such and identify the circumstances that lead users to perceive creativity in the outputs of AI and computational systems. A key problem, indeed, is that we do not yet have enough understanding of the circumstances that lead users to attribute creativity to machines. For AI systems in particular, such understanding is vital, since users' perceptions of what a system is capable of directly informs their trust in that system (Fossa, 2019). For this reason, analytical tools for identifying and grouping situations that lead to attributions of creativity and originality are needed. This is a gap that the Lovelace effect aims to fill. By applying a blend of scholarship about humans' artistic, societal, and historical interactions with computers to a recent case of proclaimed computational capability, a software trained to produce images that can be assessed as artworks, we demonstrate the value of the Lovelace effect as one such analytical tool. Our discussions of this example shows that the emergence of the Lovelace effect is the result not only of the technical and material functioning of AI software, but also of the contextual elements of its representation (Seaver, 2015). These elements may include aspects such as the visual appearance of the hardware and software interface, as well as the location and/or space in which users are invited to interact with the AI or its outputs, and the mises-en-scène through which these are presented. Ultimately, we argue that the analysis of the Lovelace effect requires a dual approach that takes into account the weights of both technical and non-technical - i.e. representational - means, which stimulate users' attributions of creativity to machines. With its recognition of the technical-representational hybridity of AI systems and output, the Lovelace effect shows how AI systems also activate the mechanisms of a cultural technique, highlighting technology's ability to co-shape new ontologies of communication and meaning with their human developers and users (Geoghegan, 2013; Karppi, 2018).

Proposing the concept of the Lovelace effect has two main objectives. First, the concept serves as a practical analytical tool for helping both critics and designers assess and plan interactive AI systems. While some efforts to study users' perceptions of AI and their impact on human-computer interactions have focused on if - or how - users attribute a sense of humanity to machines (e.g. Go and Sundar, 2019), researchers have underlined that understandings of attributions and feelings projected onto AI systems are much more diverse and nuanced than a simple human/machine dichotomy implies (Beattie et al., 2020; Guzman and Lewis, 2019). Conceptual tools that categorise different kinds of responses, such as attributions of creativity, provide ways for analysts to identify recurring dynamics of interaction, and for designers to anticipate possible outcomes of human-machine encounters. Second, the Lovelace effect contributes to reassessment of the computational contributions of Ada Lovelace, whose legacy is often misunderstood and whose work has too often been interpreted indirectly, dependent upon secondhand quotes and paraphrased prose (Ward, 2020). In this way, the Lovelace effect feeds into ongoing efforts to more systematically recognise the crucial contributions of women in the history of computing (Hicks, 2017). Additionally, it provides a necessary next step for HMC research: one that draws attention to the often theatrical contexts of system presentation that may inform the reception of these systems and their resulting output.

1. Human-Machine Communication, computational creativity and the subjective component of AI

Since AI's origins, debate about the possibility of 'strong AI' has been ongoing: could computers one day reach or surpass human intelligence? This controversy has led to an interpretive tension between those who believe that human-level intelligence could potentially be replicated through mechanical means (e.g. Minsky, 1961) and those who have pointed to essential limitations in computational proficiency, markedly distinguishing computers from humans (e.g. Dreyfus, 1972). Yet both sides of this controversy tend to agree on one point: at the crux of this debate is the internal functioning of computing machines. In other words, according to both supporters and critics of strong AI, a resolution can only be achieved through comparisons of computers' objective capabilities to those of the human mind (Ekbia, 2008).

An alternative perspective, however, has also emerged. Rather than examine what happens inside the machine's 'brain,' researchers working through what Russell and Norvig call a ‘behavioral approach’ (2002: 2-3) aim to develop computing systems that are able to *exhibit* rather than replicate intelligence. Echoing the pioneering work of Alan Turing, which assigned the responsibility of assessing AI to a human interrogator in what is now called the Turing Test, this approach focuses on the perspectives of human users. Thus, while AI has been broadly defined in objective terms as the apparent expression of intelligence by machines, it may too be defined through attention to the subjective points of view of those human users and observers who may attribute intelligence and agency to machines (Natale, 2021). By adopting this more holistic approach to understanding what AI is, we may better understand the limits, capacities, implications, and possibilities of AI technologies for the people who develop, use, and are impacted by them (Uricchio, 2017: 137). Through deliberate consideration of the human interpretations of machine processes and output, we may assess the new dynamics of human-computer interactions.

From a social sciences perspective, the implications of this approach have been explored by the Computers Are Social Actors (CASA) paradigm. The CASA paradigm originated in Reeves and Nass’ early work on the ‘media equation’ (1996), which argued that human users tend to replicate aspects of social behaviors usually applied to interpersonal interactions in their interactions with digital media. Studies conducted within this paradigm have illuminated how people’s perception of computers and robots inform the outcomes of human-computer interactions. The more recent human-machine communication (HMC) research agenda, by contrast, emerged from communication and media studies. While communication studies has primarily focused on mediated communication between humans, HMC widens this focus to accommodate communication between humans and machines, applying relevant frameworks and theories previously used in studies of human communication (Guzman, 2018). HMC investigates fundamental similarities between interpersonal communication and human-machine communication, especially when the latter involves communicative AI systems like voice assistants (Guzman, 2016; Gunkel, 2020; Hepp, 2020; Nah et al., 2020). A similar intellectual trajectory, which moves away from ontological definitions towards more relational and subjective understandings of AI, has also characterized recent discussion in computational creativity. Computational creativity as an area of study encompasses theoretical and practical inquiry into supposedly creative behaviors demonstrated by computing technologies (Gunkel, 2021). In this context, a longstanding debate in computer science revolves around the question of computers 'creating' beyond what their programmers intended or expected.

The discussion of computational creativity is traditionally traced back to the mathematician who is often credited as the first computer programmer, Ada Lovelace, who intuited as early as the mid-nineteenth century that calculators could be used not only to compute numbers but also to compose music, produce graphics, and advance science (Hollings et al., 2018; Dasgupta, 2014). Yet even earlier than Lovelace's work, reviewed below, speculation about machine capability was underway. Recall, for instance, the Mechanical Turk of 1770, wherein a life-sized automaton – the Turk – challenged participants to chess games, with the Turk masterfully completing each of its turns. In actuality, though, a chess master sat hidden within the automaton, guiding its movements and gameplay. The Turk was in effect an oversized puppet manipulated by humans for the sake of theatrical performance. This performance depended on simultaneous manifestations of alterity; the Turk was not only ostensibly non-human, but also exoticized through, for example, its donning of a turban and robe (Geoghegan, 2020). These visual cues supported the positioning of the Turk as distinct from its viewers, contributing to a non-human/human dichotomy. However, these cues also supported the implicit connections back to familiar stereotypes related to orientalism that situate the Turk within human social contexts. The Turk was therefore framed as both non-human and human in such a way that granted viewers leeway to believe in its autonomy and agency.

While the issue has been debated passionately, today few deny software's capacity to transcend programmer agency. Given the complexity of contemporary software systems and the opacity of the operations performed by neural networks, programmers can often not anticipate or even understand software output (Ekbia, 2008: 52). But there is an additional reason why software cannot be reduced to the agency of programmers. Once released, software is adapted to different platforms, situated in different contexts, embraced by different users and communities, and 'performed' in concrete, temporal interactions (Manovich, 2012). It is applied or recycled to fulfil very different goals than originally intended. It is integrated into social and cultural environments different from the ones of its origin. It is perceived, described and represented through cultural constructions of creativity that also change throughout time.

Margaret Boden, in this regard, considers the social and creative implications of AI systems throughout her germinal *Creative Mind* (2004). Also inspired by Lovelace's insights, Boden argues that focus should not be on whether or not computers are *really* creative, but if they can *appear* as such. In this way, Boden shifts from an emphasis on computational capacity understood in ontological terms towards an emphasis on human perceptions stimulated by computational operations. Boden's perspective reflects and responds to broader discussions of creativity that refuse absolute definitions under the belief that creativity is abstract rather than objective truth. Creativity, in such discussions, depends not only on the characteristics of products or actors, but also on subjective factors like the previous conceptions and cultural biases of those who are attributing creativity (Kaufman and Sternberg, 2010; Colton et al. 2014: 3; Riedl, 2014; Bringsjord et al., 2001).

In this context, the concept of the Lovelace effect aims to identify and describe specific situations in which users or observers attribute creativity to computers. While studies in HMC are increasingly emphasizing the nuance of human perception, analytical tools for categorizing specific reactions and perceptions of humans in their interactions with machines are still lacking, despite their potential value for both HMC and computational creativity research. Before further addressing how the Lovelace effect may contribute to such analyses, though, it is useful to consider the past: the contributions of Ada Lovelace, how ideas surrounding her 'objection' emerged, and why other interpretations of her work should be privileged.

2. Ada Lovelace and her 'objection'

Despite having lived in the first half of the nineteenth century, long before the introduction of electronic computers, British mathematician Ada Lovelace is rightly celebrated as one of the pioneers of modern computing (Isaacson, 2014). After collaborating with Charles Babbage - who designed, but never completed, a project that would have resulted in the first general-purpose mechanical computer, the Analytical Engine (Spufford and Uglow, 1996) - she shared groundbreaking insight related to the possibilities of programming machines not just for mathematical operations, but also for a broader range of applications (Fuegi and Francis, 2003).

In her own time, Lovelace was not so able to freely express her intellectual views as her male counterparts (Stein, 1985). Consequently, her contributions to computing were partially hidden: for example, disguised in the notes appended to her translation of a French article on Babbage's analytical engine by Italian mathematician Luigi Federico Menabrea. The translation was published anonymously, and her notes credited her with the initials 'A.A.L.' (Lovelace, 1843). Yet in such notes historians have found ample evidence for crediting Lovelace with a vision of the relationship between computation and symbolic processing that was more expansive and imaginative than Babbage's (Abbate, 2012; Hollings et al., 2018; Hammerman and Russell, 2015). Working more than a century before the digital age, Lovelace's intuitions included anticipating some aspects of the idea of software and envisioning that computing operations could be used to generate poems and music (Carlucci Aiello, 2016).
One particular passage from these notes has been widely cited and discussed:

> The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform. It can follow analysis, but it has no power of anticipating any analytical relations or truths (Lovelace, 1843: 722).

The most influential interpretation of this passage is that of Alan Turing in 'Computing Machinery and Intelligence' (1950). In this paper, a milestone for the then-budding field of AI, Turing outlines the proposal for his Imitation Game, now widely known as the Turing Test. In the remainder of the paper, Turing notes potential objections to the Test, one of which is the 'Lovelace objection': that a computer has no pretensions to originate anything. Turing's interpretation of Lovelace's alleged 'objection' spurred a longstanding debate in computer science about whether or not a computer can originate anything not anticipated by the programmer (Abramson, 2008). As Megan Ward has recently observed, however, the so-called Lovelace objection is 'typically invoked erroneously: incomplete, decontextualized, and attributed to Turing's citation of Lovelace rather than to the original source' (2020: 146). In fact, 'Lovelace's contribution to the field might be more accurately stated not as scepticism but as an invitation to develop the machine's capacity for originality and the human's role within this radical new field of possibility' (148).

Ward (2020) has also noted the advantage of reframing Lovelace's statement as having to do more with circuits of meaning created and perpetuated by human-machine interactions than with what machines might achieve in isolation. As shown by the passage of Lovelace's notes preceding the above-cited statement, the Victorian mathematician was evidently concerned not only with what the Analytical Engine could do, but also with what people might think of it:

> It is desirable to guard against the possibility of exaggerated ideas that might arise as to the powers of the Analytical Engine. In considering any new subject, there is frequently a tendency, first, to overrate what we find to be already interesting or remarkable; and, secondly, by a sort of natural reaction, to undervalue the true state of the case, when we do discover that our notions have surpassed those that were really tenable (Lovelace, 1843: 722).

Considering that Lovelace's subsequent comment about the capacity to originate anything referred specifically to the analytical engine and not to computing in general, Lovelace suggests not skepticism about the potential of computing as such, but rather instinctive awareness that evaluations of computing machines are always subjective: humans assess computational output based on their own assumptions and perceptions of computational capacity. Although this passage has been given relatively little attention, it indicates that Lovelace intuited the extent to which computing machines could inspire exaggerated claims and projections about their functionality and outcomes (Cave et al., 2020).

Our aim here is not so much to reach a definite interpretation of Lovelace's text. Rather, we propose reconsideration of Lovelace's pioneering intuitions that is sympathetic with recent calls to study AI from a communication and media studies perspective (Guzman, 2018; Gunkel, 2020; Hepp, 2020), and that explicitly recognises these intuitions' relevance to discussions of computational creativity (Boden, 2004). We believe that it is necessary to examine AI technologies not only in terms of technical capability, but also in terms of their social constructions and deployments (Uricchio, 2017), as well as how they are inserted within communicative circuits wherein machine output depends on and informs the interpretive work of participating humans.

3. Identifying the Lovelace Effect: The case of AICAN

In contrast with the Lovelace objection, the Lovelace effect represents a reevaluation of Lovelace's contribution that responds to more recent efforts in areas like computational creativity that emphasize the role of human users in attributing qualities such as intelligence, agency, and originality to machines. The Lovelace effect, in this regard, puts forth a concrete analytical tool for mobilizing these efforts by describing situations in which the behavior of computing systems *is perceived by users* as original and creative.

To better show what the Lovelace effect entails and how it can be applied to the analysis of concrete situations and specific technologies, we selected the case study of AICAN: 'the first and only AI artist trained on 100,000 of the greater works in art history' (AICAN, n.d.). While AI-generated art is increasingly common, only a few studies have considered audience reception and evaluation. Such studies indicate that audience members are often unable to distinguish between AI-generated and human-created artworks (Gangadharbatla 2021). Further, audience members approach AI-generated art with particular expectations of what it should look like (e.g. abstract), and may have more favourable views of such art when it meets these expectations (Chamberlain et al., 2018). The case study of AICAN and its output serves as a useful example for showing the utility of the Lovelace effect because it comprises publicly-documented stages of production and reception that are relatively distinct; such step-by-step documentation is an uncommon luxury for scholars of AI systems. Drawing from this documentation, we reflect upon how AICAN's output has been framed as 'art', progressing through the discussion with the awareness that some scholars have stressed that AI might contribute to shifting existing conventions and meanings of art (Notaro, 2020). Following an introduction to AICAN and its functionality, we consider the promotional and curatorial decisions manifest in AICAN's first solo exhibition, 'Faceless Portraits Transcending Time'. These decisions have contributed to the emergence of the Lovelace effect amongst exhibition attendees, with AICAN being positioned as a semi-autonomous creative agent through both technical and representational means.

3.1 An exhibition in search of an author

In 2019, the HG Contemporary Gallery in New York City opened a new exhibition featuring paintings produced through AICAN (Artificial Intelligence Creative Adversarial Network), an AI system designed by Rutgers University computer scientist Ahmed Elgammal and his collaborators. Heralded by the media as 'AI's first art exhibit' that is 'filled with art made entirely by AI' (Mashable, 2020), the event - entitled 'Faceless Portraits Transcending Time' - represented a compelling example of how AI developers mobilize technical as well as communicative means to create the impression that their systems are endowed with originality and creativity.

AICAN is based on a variant of Generative Adversarial Networks (GANs). A GAN comprises two networks: a generator, which produces output, and a discriminator, which evaluates how that output compares to a training set. The GAN has succeeded when the discriminator cannot distinguish between the generated examples and the examples from the training set. Humans may similarly struggle to discriminate between generated output and 'real' content; a recent *New York Times* article shows numerous GAN-generated facial portraits and asks 'Do These People Look Real to You?' (Hill and White 2020). AICAN's developers stressed that the works were well received at all of the art venues in which they were exhibited, and that viewers who had no previous knowledge of AICAN could often not tell the difference between the computer-generated and human-produced artworks (Mazzone and Elgammal, 2019: 5).

For 'Faceless Portraits', Elgammal trained AICAN on a database of 80,000 images - representing, in his view, the Western art canon of the last five centuries - and activated the generator and discriminator with two aims: to learn the aesthetics of the canon, but also to produce output that does not mimic those aesthetics too closely. This latter aim was inspired by the work of cognitive psychologist Colin Martindale, who has controversially argued that artistic change is not attributable to political, religious or social forces, but mainly to a constant pressure for novelty (Martindale, 1990).

In his study, Martindale used quantitative methods to support his claims about the evolution of art and the role of the drive to novelty, which has motivated the development of earlier AI art generators (Krzeczkowska et al., 2010). As neural networks such as GANs also mobilize the power of statistics, drawing on Martindale as a theoretical foundation suited the kinds of tools available to Elgammal. As one commentator observed, the idea was ‘a convenient take, given that any machine-learning technique has to base its work on a specific training set’ (Bogost, 2019). While the theory might have provided Elgammal with an appropriate justification of AICAN’s computational approach, however, it did not encounter the favor of observers from an art history background, who have often responded negatively to Elgammal’s enterprise (see, among others, Mansell, 2021; Notaro, 2020). One of the problems highlighted, for instance, is that AICAN’s neural networks cannot mobilize the symbolic, allegorical, and cultural meanings that are embedded in the artworks used to train them, so that the images produced by AICAN are equivalent to abstract painting even when if they are purportedly framed within the genre of portraiture (Bogost, 2019).

Although Elgammal (2018) asserts that ‘using our prior work on quantifying creativity, AICAN can judge how creative its individual pieces are,’ it proves difficult for the Rutgers University computer scientist to make a convincing claim that its technical procedure provided a reliable method to mechanically produce art. As Ian Bogost observed, ‘the whole of 20th-century art was predicated on the idea that putting something in a gallery or museum makes it art, rather than the opposite’ (Bogost, 2019). In fact, the exhibition’s mise-en-scene suggests that the circumstances of the pieces’ reception were at least as important as Elgammal’s technical choices. The decision to present the AI-generated images in an art gallery bolstered the claim that the products of AICAN were of an artistic character; Elgammal himself noted that AICAN-generated works had not been so readily hailed as artworks when they were presented at technological exhibitions and venues (Future Blink, 2019). As established by studies in art history (e.g. Fyfe, 2000), artistic institutions including galleries are instrumental in providing authority to specific works, facilitating definitions of what values as art and what is more likely to be accepted by the public. According to Mazzone and Elgammal, whenever the images were exhibited in artistic venues 'the reception of works was overwhelmingly positive on the part of viewers who had no prior knowledge that the art shown was generated using AI' (Mazzone and Elgammal, 2019). Thus, the presentation of AICAN's outputs within an art gallery likely played a role in orienting reception from the public. The choice of a gallery located in Chelsea, the epicentre of New York City's art world, may have further strengthened this effect.

The arrangement of the exhibition - comprising two series of 'faceless portraits', with the first imitating Renaissance portraits and the second delving further into abstraction (Oduber, 2019) - also contributed to the sense that the images reflected artistic intent. The installation followed established conventions in art venues, presenting the pictures as individual pieces hung so that viewers could observe them frontally and individually. Aligning with the name of the exhibition - 'Faceless Portraits Transcending Time' - individual pieces were labelled with such titles as 'Faceless Portrait of a Merchant.' Studies in the psychology of art have shown that the presence and content of titles inform people's assessments and recognition of artworks (Franklin et al., 1993; Russell and Milne, 1997). The titles chosen for AICAN's output draw from a long tradition of titles using the 'Portrait of a [x]' format. By using such titles, as well as a database of images from a specific genre of the figurative arts (i.e. portraiture), known forms are evoked, signalling to viewers that they are in an 'art' setting that necessitates an 'art' mindset. Moreover, wooden picture frames were used for several pictures, reiterating conventions of how artworks are presented and exhibited - although no mention of this was made in publications by Elgammal and collaborators (Mazzone and Elgammal, 2019).

The art exhibition setting also provided opportunities for curatorial interventions that contributed to shape receptions of the objects exhibited: interventions that both art and development communities recognize as imperative for positive AI-generated art reception (Spratt, 2018: 41). Gangadharbatla (2021) has experimentally shown that the information provided to viewers about AI-generated pictures impacts on their perception. Textual materials that accompanied the exhibition described AICAN as a 'technical artist' that 'lives at the Art and AI Lab at Rutgers'; it also explicitly mentioned that AICAN was programmed to be creative (HG

Contemporary, 2019). The exhibition's 'Artist Statement' mobilized language and themes familiar to art circles by stressing, for instance, how the 'uncanny, dream-like imagery' generated by AICAN excavate 'the ageless themes of mortality and representation of the human figure,' or by underlining that 'the ability of algorithms to generate imagery from their 'imagination' causes a severance between imagery and reality altogether' (AICAN and Elgammal, 2019). Such exhibition materials provided a textual framing that invited visitors to understand and read the prints displayed at the exhibition from a specific viewpoint that emphasised the purported artistic value of Elgammal and HG Contemporary's operation.

Finally, the exhibition was promoted by mobilizing established conventions that define artistic creativity through attribution of authorship. Promotional materials explicitly assigned co-authorship to Elgammal and AICAN, and introduced the exhibition as 'a collaboration between an artificial intelligence named AICAN and its creator' (AICAN and Elgammal, 2019). In this formulation, AICAN and Elgammel are seemingly placed at the same level. It is ironic, then, that Elgammal claims to not consider himself an artist (Elgammal, 2018) but proposes that his 'collaborator' should be regarded as such.

Elgammel's self-positioning as a collaborator may exemplify a kind of 'shifting sense of social presence' (Guzman, 2019: 349) in which AICAN has been afforded a sense of artistic agency usually reserved for humans. Research on the variables that inform humans' perception of AI-generated objects as artworks have shown how people tend to be biased against the possibility that computers generate art, but that elements that contribute to the anthropomorphization of the system - such as, in AICAN's case, references to the human programmer and the software acting as "collaborators" - may counteract such bias, facilitating consideration of the outputs as artworks (Chamberlain et al., 2018; Lu, 2005). For all Algammal's emphasis on the technical solution and the theoretical approach taken from Martindale, AICAN's effectiveness was largely measured by pointing to viewers' reactions: in other words, by pointing to its Lovelace effect. To demonstrate that AICAN was capable of artistic creation, the developers devised a 'visual Turing test to register how people would react to the generated images and whether they could tell the difference between AICAN- or human-created art' (Mazzone and Elgammal, 2019: 4). Although the visual Turing test was conducted at an art fair rather than in a laboratory setting and the human-authored artworks that acted as controls in this test were selected by the AICAN team, Elgammal and collaborators did not acknowledge that the effects AICAN's works had on viewers were also informed by aspects such as the context and the modality of the works' presentation.

As art institutions such as galleries and museums construct objects as aesthetic artworks, they also contribute to assign them material value within the art market. Prints of AICAN-generated images were sold at the HG exhibitions for between 6,000 and 18,000 US dollars (Bogost, 2019). As Mansell (2021) notes, the commodification of artworks is distinct from their production, as it entails art objects being integrated within institutional, representational and legal systems that guarantee that the objects can become a property and can be assigned monetary value. Within digital art, a range of procedures have been developed for museums, galleries and private collectors to 'own' something that escapes traditional understandings of art objects, such as a website or a software piece: for instance, conserving not only the software code but also the machinery used to run it (Dominguez Rubio, 2020). The complexity of these procedures, which provide material substance to objects that often escape fixed materiality, do not coincide with the act of artistic creation but more aptly to the needs and conventions of art institutions; in the same vein, the exhibition of AICAN-generated pictures and their elevation as (marketable) artworks was separated from the act of algorithmic creation of the images (Christie, 2018; Notaro, 2020). In this sense, it could be argued that it wasn't AICAN who generated the 'artworks' but rather the institutional frame that surrounded its presentation to the public; or, more aptly, that the fact that AICAN pictures were treated by some as artworks is not indicative of the creativity of AICAN in itself but rather of the particular circumstances and activities that prepared and foregrounded attributions of AICAN's creativity.

While some scholars deny art generation systems agency, intent, and authorship (McCormack et al., 2019; Chamberlain et al. 2018), others argue for the inherently creative nature of some systems (Cook and Colton 2011; Mazzone and Elgammal 2019: 4). Ultimately, though, the AICAN case illustrates that attributions of creativity can be facilitated through both representational and technical means, with the 'disaggregation of artistic process from execution' (Uricchio, 2017: 129) being only one piece of a much larger puzzle of public reception. For all AICAN's developer's focus on technical choices, viewer reactions cannot be explained solely and even predominantly through AICAN's internal functioning. Scholars in visual anthropology have shown that all artifacts, including artworks, are always embedded in social and cultural circuits of meanings that direct the attributions given to them (Gell, 1998; Dominguez Rubio, 2020). Every consideration of the circumstances through which AI-generated artworks result in a Lovelace effect should take therefore into account the particular social and cultural frame in which the technology's outputs were embedded.

The case of AICAN, therefore, reminds us that the emergence of the Lovelace effect never happens in a vacuum, but can only be explained by attentively considering the context of reception of the digital resource under consideration. Attributions of creativity never depend on the technical configurations alone, since there is always a cultural and social context in which the technology is immersed (Natale, 2021). In the case of AICAN, this context included institutional infrastructures, such as: museums, art galleries, and fairs, that contribute to reinforce claims of artistic values for the software's outputs; curatorial texts, such as the printed exhibition catalogue or the online description of an exhibition; material props such as the wooden frames in which the artworks were mounted; and, finally, the public's expectations of what is meant by art and creativity, which can vary significantly in social and spatial context, since the visitors of an art gallery in New York City will differ from those attending an art fair in Asia or from other potential publics in different parts of the world.

Conclusion

As shown by our examination of the AICAN case, attributions of creativity do not depend exclusively on the technical functioning of a computing resource, but on the complex interactions between technical, cultural, and social features that prepare the conditions for the emergence of the Lovelace effect. While the 'Lovelace objection' has posed the question of creativity from a technical viewpoint, it is necessary to recalibrate perceptions of creativity in machines towards the perspectives of the users who attribute such creativity.

Although our analysis has focused on the case of AICAN, a similar dynamic can be observed in other AI resources that have proven able to stimulate the emergence of the

Lovelace Effect. Notwithstanding their technical limitations, for instance, responses given by voice assistants such as Alexa, Siri, and Google Assistant have proven able to appear ‘smart’ to users and stimulate the emergence of the Lovelace effect when embedded in established conventions of storytelling and wit (Thorne 2020). Similarly to what has been observed in the AICAN case, the perception of creativity stimulated by voice assistants also depends on the existing social conventions that are mobilized by teams of creative writers who work for companies such as Amazon, Apple and Google to script statements that might be considered insightful or creative by users (Young, 2019). Another useful example of the recurrent presence of this dynamic in different contexts is that of computer games. Human-machine confrontations in games such as computer chess or go encouraged lively debates about the possibility of machines originating something original and creative (Rasskin-Gutman, 2009). The capacity of the systems to surprise human opponents and observers stemmed not only from the internal functioning of the software, but also from the complex mises-en-scène through which these confrontations were staged. For instance, in the celebrated duel between AlphaGo, a Google-funded computer program that plays the Chinese game Go, and South Korean professional player Lee Sedol, AlphaGo’s moves were ‘simulated’ by a person who physically performed the moves on the material board, following system instructions. This person was just a human proxy - a medium between machine and human opponent - but his presence helped corroborate the interpretative framework that pointed to the interchangeability between the human and the software. This human’s overt presence - in direct contrast to the hidden human presence in the Mechanical Turk of 1770 - reflects a new form of mediation that aims to recognise not just computational capability, but also human likeness. The presence of this person bolstered the comparison of AlphaGo to a human player, in turn contributing to a sense of awe for AlphaGo’s creativity in its gameplay (Bory, 2019). These two instances help show how the dynamics observed in the case of AICAN are not just specific to this case, but can be generalized as characteristic of the ways in which computational resources come to generate the Lovelace effect.

The implications of such dynamics are not irrelevant; perceptions of an AI software as capable of originality and creativity, in fact, may enhance the authority of, and trust in, these systems and those who have created them. More broadly, being able to predict how a user may perceive and use a product allows producers to create and adapt this project to suit the expressed needs of a maximum consumer base. As Eitan Wilf observes in his article about computer-generated jazz, ‘computerized algorithms in consumer-centered production derive their profitability from their ability to tap into each consumer’s distinct patterns or styles of consumer behavior’ (2013: 717).

Considering how designers may develop strategies that maximize their software’s potential to create a Lovelace effect does not deny the active role of users, though. Although software is always constructed with action in mind (Bucher, 2018), users’ reactions and behaviors may subvert or nullify the intentions and expectations of designers (Gunkel, 2020). The Lovelace effect acknowledges that computational capacity is always informed by individual and subjective understandings. Each user experience results in a system’s unique perceptive and reactive effects, which may or not have been predicted by developers (Colton et al., 2014). Taina Bucher (2018) uses the term ‘algorithmic imaginary’ to refer to ‘the way in which people imagine, perceive and experience algorithms and what these imaginations make possible.’ Other scholars have further argued that algorithms should be understood within the circuits of imagination that surround them (Finn, 2017), or have called for a hermeneutics of algorithms that affirms the centrality of user interpretation (Andersen, 2020).

The Lovelace effect mediates actual software functionality with how individuals conceptualise and interpret that software, reminding us that all outcomes of interactions between humans and machines represent constant implicit and indirect negotiation between programmer intention and user experience. Considering perceptions of algorithms and algorithmic output also draws attention to ingrained assumptions about computational agency, developmental transparency, and control (Ziewitz, 2016). Future studies may elaborate upon how the Lovelace effect might be more subtly provoked to prompt particular user experiences and behaviours in commercial and/or political contexts, drawing upon case studies from a wider range of circumstances than time has permitted here.

From trawling the Web to producing artistic works, systems embedded with AI are wide-reaching and diverse. In this article, we propose a conceptual tool for understanding how AI may attract attributions of creativity that might affect outcomes of human-machine interactions. One potential criticism of the Lovelace effect might be that it does not address creativity on an ontological level: by shifting the emphasis to user perception, it renounces any one definition of creativity. Discussions of computational creativity, however, have largely established that 'creativity' is relative, historial, and subjective (Gunkel, 2021). Any attempt to define creativity in ontological terms would actually obscure the fact that creativity cannot be assigned as a quality of specific computing systems, but can only be attributed by users in specific situations. The Lovelace effect, in this sense, moves assessments of computational creativity from the level of the machine (how AI functions and what it does) to the level of reception (how users and observers attribute meaning to AI). In other words, the Lovelace effect advances a relational rather than ontological approach to computational creativity and, more broadly, AI. Building upon the extended debates about the Lovelace objective, the Lovelace effect advocates an alternative way of understanding computational creativity that places humans at the center of analysis.

References

Abbate J (2012) *Recoding gender: Women's changing participation in computing*. Cambridge, Mass.: MIT Press.

Abramson D (2008) Turing's responses to two objections. *Minds and Machines* 18(2): 147-167.

Andersen J (2020) Understanding and interpreting algorithms: toward a hermeneutics of algorithms. *Media, Culture & Society*, published online before print 28 April 2020, doi: 10.1177/0163443720919373.

AICAN. Available at: https://aican.io (accessed 21 September 2021).

AICAN and Elgammal A (2019) Artist's statement. Available at: http://www.hgcontemporary.com/artists/aican-ahmed-elgammal (accessed 20 November 2020).

Beattie A, Edwards AP and Edwards C (2020) A bot and a smile: Interpersonal impressions of chatbots and humans using emoji in computer-mediated communication. *Communication Studies* 71(3): 409–427.

Boden MA (2004) *The creative mind: Myths and mechanisms*. London: Routledge.

Bogost I (2019) The AI-Art Gold Rush is Here. *The Atlantic*, available at https://www.theatlantic.com/technology/archive/2019/03/ai-createdart-invades-chelsea-gallery-scene/584134/ (accessed 16 September 2021).

Bory P (2019) Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo. *Convergence* 25(4): 627–642.

Bringsjord S, Bello P and Ferrucci D (2001) Creativity, the Turing test, and the (better) Lovelace test. *Minds and Machines* 11: 3–27.

Bucher T (2018) *If... Then: Algorithmic power and politics*. Oxford: Oxford University Press.

Carlucci Aiello L (2016) The multifaceted impact of Ada Lovelace in the digital age. *Artificial Intelligence 235*: 58–62.

Cave S, Dihal K and Dillon S, eds. (2020) *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford: Oxford University Press.

Chamberlain R, Mullin C, Scheerlinck B, and Wagemans J (2018) Putting the art in artificial: Aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts* 12(2): 177-192.

Christie's (2018) Is artificial intelligence set to become art's next medium? Available at: https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx (accessed 22 November 2020).

Colton S, Cook M, Hepworth R, and Pease A (2014) On acid drops and teardrops: Observer issues in computational creativity. In: *Proceedings of the 7th AISB Symposium on Computing and Philosophy.* Available at: http://ccg.doc.gold.ac.uk/wp-content/uploads/2016/03/colton_aisb14c.pdf (accessed 12 May 2020).

Cook M and Colton S (2011) Automated collage generation – With more intent. In: Ventura D, Gervás P, Harrell DF, Maher ML, Pease A, and Wiggins G (Eds.), *Proceedings of the Second International Conference on Computational Creativity*. México City: ICCC, 1-3.

Copeland J (2004) *The essential Turing*. Oxford: Oxford University Press.

Dasgupta S (2014) *It began with Babbage: The genesis of computer science*. Oxford: Oxford University Press.

Day M, Turner G, and Drozdiak N (2019) Thousands of Amazon workers listen to Alexa users' conversation. Available at: https://time.com/5568815/amazon-workers-listen-to-alexa (accessed 17 May 2020).

Dominguez Rubio F (2020). *Still life: Ecologies of the modern imagination at the art museum*. University of Chicago Press.

Dreyfus HL (1972) *What computers can't do: A critique of artificial reason*. New York: Harper & Row.

Ekbia HR (2008) *Artificial dreams: The quest for non-biological intelligence*. Cambridge: Cambridge University Press.

Elgammal A (2018) Meet AICAN, a machine that operates as an autonomous artist. The Conversation, 17 October 2018. Available at: https://theconversation.com/meet-aican-a-machine-that-operates-as-an-autonomous-artist-104381 (accessed 20 November 2020).

Finn E (2017) *What algorithms want: Imagination in the age of computing*. Cambridge, Mass.: MIT Press.

Fossa F (2019) "I don't trust you, you faker!": On trust, reliance, and artificial agency. *Teoria* 39(1): 63-80.

Franklin MB, Becklen RC and Doyle CL (1993). The Influence of Titles on How Paintings Are Seen. *Leonardo*, 26(2), 103–108.

Fuegi J and Francis J (2003) Lovelace & Babbage and the creation of the 1843 "notes." *IEEE Annals of the History of Computing 25*(4): 16–26.

Future Blink (2019) We attended an AI's first art exhibit in NYC. Available at: https://mashable.com/video/ai-art-exhibit-nyc/ (accessed 20 November 2020).

Fyfe G (2000) *Art, power and modernity: English art institutions, 1750-1950*. London: Leicester University Press.

Gangadharbatla, H. (2021). The Role of AI Attribution Knowledge in the Evaluation of Artwork. *Empirical Studies of the Arts*, published online before print 16 February 2021, https://doi.org/10.1177/0276237421994697.

Gell A (1998) *Art and agency: An anthropological theory*. Oxford: Oxford University Press.

Geoghegan BD (2013) After Kittler: On the cultural techniques of recent German media theory. *Theory, Culture & Society* 30(6): 66-82.

Geoghegan BD (2020) Orientalism and informatics: Alterity from the chess-playing Turk to Amazon's Mechanical Turk. *Ex-position* 43: 45-90.

Go E and Sundar SS (2019) Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97: 304–316.

Green CD (2001) Charles Babbage, the Analytical Engine, and the possibility of a 19th-century cognitive science. In: Green CD, ShoreM and Teo T (Eds.), *The transformation of psychology: Influences of 19th-century philosophy, technology, and natural science*. Washington, DC: American Psychological Association, pp. 133–152.

Gunkel DJ (2021) Computational creativity. In: Navas E, Gallagher O and Burrough H (Eds.), *The Routledge Handbook of Remix Studies and Digital Humanities*. London: Routledge.

Gunkel DJ (2020) *An introduction to communication and Artificial Intelligence*. Cambridge, UK: Polity.

Guzman AL, ed. (2018) *Human-machine communication: Rethinking communication, technology, and ourselves*. New York: Peter Lang.

Guzman AL (2016) Making AI safe for humans: A conversation with Siri. In: Gehl RW and Bakardjieva M (eds.), *Socialbots and their friends: Digital media and the automation of sociality*. London: Routledge, pp. 69–82.

Guzman AL (2019) Voices *in* and *of* the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behaviour* 90: 343-350.

Guzman AL and Lewis SC (2019) Artificial intelligence and communication: A Human–Machine Communication research agenda. *New Media & Society* 22(1): 70–86.

Hammerman R and Russell AL (2015) *Ada's Legacy: Cultures of Computing from the Victorian to the Digital Age*. London: Morgan & Claypool.

Hepp A (2020) Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies. *Media, Culture & Society* 42(7-8): 1410-1426.

HG Contemporary (2019) What does a machine know about mortality? Available at_ http://www.hgcontemporary.com/news/what-does-a-machine-know-about-mortality (accessed 16 September 2021).

Hicks M (2017) *Programmed inequality: How Britain discarded women technologists and lost its edge in computing*. Cambridge, Mass.: MIT Press.

Hill K and White J (2020) Designed to deceive: Do these people look real to you? *The New York Times*. Available at: https://www.nytimes.com/interactive/2020/11/21/science/artificial-intelligence-fake-people-faces.html?smtyp=cur&smid=tw-nytimes (accessed 23 November 2020).

Hollings C, Martin U and Rice A (2018) *Ada Lovelace: The making of a computer scientist*. Oxford: Bodleian Library.

Isaacson, W (2014) *The innovators: How a group of inventors, hackers, geniuses and geeks created the digital revolution*. New York: Simon and Schuster.

Karppi T, Böhlen M and Granata Y (2018) Killer Robots as cultural techniques. *International Journal of Cultural Studies* 21(2): 107-123.

Kaufman JC and Sternberg RJ, eds. (2010) *The Cambridge handbook of creativity.* Cambridge: Cambridge University Press.

Kelleher JD (2019) *Deep Learning*. Cambridge, Mass.: MIT Press.

Krzeczkowska, A, El-Hage, J., Colton, S., & Clark, S. (2010). Automated Collage Generation – With Intent. In D. Ventura, A. Pease, R. Pérez y Pérez, G. Ritchie, & T. Veale (Eds.), *Proceedings of the First International Conference on Computational Creativity*. Lisbon: ICCC, 36-40.

Lovelace AA (1843) Notes by the translator. *Taylor's Scientific Memoirs 3*: 666–731.

Lu L (2005). Pre-service art teacher negative attitudes and perceptions of computer-generated art imagery: Recommendations for pre-service art education programs. *Visual Arts Research* 31(1), 89–102.

Manovich L (2012) How to follow software studies. Available at: http://manovich.net/content/04-projects/075-how-to-follow-software-users/72_article_2012.pdf (accessed 7 November 2020).

Mansell, R. (2021). Enclosing or Democratising the AI Artwork World. *Cambridge Journal of Law, Politics, and Art*, 1, 247–251.

Martindale C (1990) *The clockwork muse: The predictability of artistic change*. London: Basic Books.

Mashable (2020) This exhibit is filled with art made entirely by AI. Available at: https://twitter.com/mashable/status/1327603942532943872?s=20 (Accessed 14 November 2020).

Mazzone M and Elgammal A (2019) Art, creativity, and the potential of Artificial Intelligence. *Arts* 8(1): 1-9.

McCormack J, Gifford T and Hutchings P (2019) Autonomy, authenticity, authorship and intention in computer generated art. In Ekárt A, Liapis A and Castro Pena ML (eds.), *Computational intelligence in music, sound, art and design (EvoMUSART 2019 Conference Proceedings)*. Leipzig: Springer, pp. 35-50.

Minsky M (1961) Steps toward artificial intelligence. *Proceedings of the IRE* 49(1): 8–30.

Nah S, McNealy J, Kim JH and Joo J (2020) Communicating Artificial Intelligence (AI): Theory, research, and practice. *Communication Studies* 71(3): 369–372.

Natale S (2021). *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test*. Oxford: Oxford University Press.

Notaro A (2020). State-of-the-art: AI through the (artificial) Artist's Eye. In: *Proceedings of EVA London 2020: Electronic Visualisation and the Arts*. London: EVA, pp. 322-328.

Oduber A (2019) AICAN: Facebook Portraits Transcending Time. *Medium*. Available at: https://medium.com/lotus-fruit/aican-faceless-portraits-transcending-time-89ea045f3811 (accessed 21 September 2021).

Poe EA (1836) Maelzel's chess player. *Southern Literary Messenger* 2(5): 318-326.

Rasskin-Gutman D (2009) *Chess metaphors: Artificial Intelligence and the human mind*. Cambridge, Mass.: MIT Press.

Reeves B and Nass C (1996) *The media equation: How people treat computers, television, and new media like real people and places*. Stanford, Calif.: CSLI Publications.

Riedl MO (2014) The Lovelace 2.0 Test of artificial creativity and intelligence. Available at: https://arxiv.org/pdf/1410.6142.pdf (accessed 15 April 2020).

Russell PA and Milne S (1997). Meaningfulness and hedonic value of paintings: Effects of titles. *Empirical Studies of the Arts* 15(1): 61–73.

Russell SJ and Norvig P (2002) *Artificial Intelligence: A modern approach*. Englewood Cliffs: Prentice Hall.

Seaver N (2015) The nice thing about context is that everyone has it. *Media, Culture & Society* 37(7): 1101-1109.

Spratt E (2018) Creation, Curation, and Classification: Mario Klingemann and Emily L. Spratt in conversation. *XRDS* 24(3): 34-43.

Spufford F and Uglow JS, eds. (1996). *Cultural Babbage: Technology, time and invention*. London: Faber.

Stein D (1985) *Ada: A life and a legacy*. Cambridge, Mass.: MIT Press.

Thorne S (2020) Hey Siri, tell me a story: Digital storytelling and AI authorship. *Convergence*, published online before print 15 April 2020. Available at: https://journals.sagepub.com/doi/full/10.1177/1354856520913866 (accessed 23 April 2020).

Turing AM (1950) Computing machinery and intelligence. *Mind* 59(236): 433-460.

Uricchio W (2017) Data, culture and the ambivalence of algorithms. In: Schäfer MT and van Es K (eds.), *The datafied society: Studying culture through data*. Amsterdam: Amsterdam University Press, pp. 125-137.

Ward M (2020) Victorian fictions of computational creativity. In: Cave S, Dihal K and Dillon S (eds.), *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford: Oxford University Press, pp. 144-64.

Wilf E (2013) Toward an anthropology of computer-mediated, algorithmic forms of sociality. *Current Anthropology* 54(6): 716-739.

Young L (2019) ‘I’m a cloud of infinitesimal data computation’. When machines talk back: An interview with Deborah Harrison, one of the personality designers of Microsoft’s Cortana AI. *Architectural Design* 89(1): 112–117.

Ziewitz M (2016) Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*, *41*(1): 3–16.