Received: 18 November 2020 | Revised: 3 March 2021 | Accepted: 13 April 2021

Cognitive Computation and Systems

DOI: 10.1049/ccs2.12018

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Questioning 'what makes us human': How audiences react to an artificial intelligence–driven show

Rob Eagle[1] | **Rik Lander**[1] | **Phil D. Hall**[2]

[1]Digital Cultures Research Centre, Faculty of Arts, Creative Industries and Education (ACE), University of the West of England Bristol, Bristol, UK

[2]Elzware Limited, Bristol, UK

Correspondence

Rob Eagle, Digital Cultures Research Centre, Faculty of Arts, Creative Industries and Education (ACE), University of the West of England Bristol, Pervasive Media Studio, 1 Canon's Road, Harbourside, Bristol, UK.
Email: robert2.eagle@live.uwe.ac.uk

Funding information

University of the West of England

Abstract

I am Echoborg is promoted as 'a show created afresh each time by the audience in conversation with an artificial intelligence (AI)'. The show demonstrates how AI in a creative and performance context can raise questions about the technology's ethical use for persuasion and compliance, and how humans can reclaim agency. This audience study focuses on a consecutive three-night run in Bristol, UK in October 2019. The different outcomes of each show illustrate the unpredictability of audience interactions with conversational AI and how the collective dynamic of audience members shapes each performance. This study analyses (1) how I am Echoborg facilitates audience cocreation in a live performance context, (2) the show's capacity to provoke nuanced understandings of the potential for AI and (3) the ability for intelligent technology to facilitate social interaction and group collaboration. This audience study demonstrates how the show inspires debate beyond binary conclusions (i.e. AI as good or bad) and how audiences can understand potential creative uses of AI, including as a tool for cocreating entertainment *with* (not just for) them.

1 | INTRODUCTION

Coined in 2014 by psychologists Kevin Corti and Alex Gillespie, an echoborg is a 'hybrid agent' of a real person whose main task is to enact the instructions of a chatbot system [1]. I am Echoborg creates a collective experience for audiences to interact with a conversational artificial intelligence (AI) through an echoborg interface, who for all except a handful of early shows has been performed by actor Marie-Helene Boyd. When an audience enters an auditorium, they are met with a dark, foreboding soundtrack setting the scene for the coming show. In the background sits Phil D. Hall, a veteran of conversational AI development since 1982 and the technical director of the I am Echoborg system. A host for the performance, usually the show's cocreator Rik Lander, delivers an introduction and sets the challenge for the audience: 'to find a best possible outcome for the relationship between humans and intelligent machines.' Often the audience will deliberate with each other until the first person steps forward to sit in a chair onstage and speak to the echoborg. The conversational AI responds to the audience member with 'Are you here for the interview?' A job interview scenario plays out in which the AI feeds questions to the echoborg, asking whether the audience member would make a suitable echoborg. The premise, in short, is that the echoborg appears to be recruiting more echoborgs from the audience.

The dataset of the AI grows with each performance, shaped by each contribution from each audience member who speaks into the microphone onstage. The echoborg will respond to the audience member based only on audio input from the microphone, which links to the AI via live speech-to-text software. The AI then responds via a text-to-speech programme into the headphones of the echoborg, who repeats the words aloud.

Only one person at a time may speak to the echoborg. However, throughout each performance, the audience collectively reflect on how the AI is responding. They strategise as a group regarding the best way for the next person to confront or reply to the AI. After a variable amount of time ranging from 30 to 70 min of exchange with the audience via the echoborg, the AI is programmed to draw the interview process

to a close, and the audience has 10 min to decide collectively on a best possible outcome for the relation between humans and intelligent machines. Sometimes there is no consensus, but usually the group is able to form a collective verdict. One spokesperson from the audience then must approach the echoborg and deliver the audience's proposal regarding how best for humans and intelligent machines to move forward together. The AI then decides whether this proposal is suitable at the end of the performance.

Audiences create the performance each night but also contribute to the ongoing evolution of the show. After each performance, the makers analyse the anonymised conversational logs, narrative flow and data points (variables) so they can make corrections to the code and create new utterances and responders for the AI. Therefore, each performance influences the next.

The show prompts the audience to question how AI is being used for persuasion and manipulation. Audiences quickly understand that they can either confront and resist the AI system or play along with it. As they grow increasingly compliant with the AI's questions (e.g. responding positively that, yes, they are here for the interview and would like the job of an echoborg), they also look for ways to challenge, subvert or even defeat the AI. Each audience's 'best possible outcome for the relationship between humans and intelligent machines' is an ethical proposal regarding how they envision a future coexistence with the sort of system with which they are conversing.

Our study investigates how an artistic application of AI (in this case, employed in a live performance) can prompt a public audience to engage critically with the ethical dilemmas of existing and evolving technology. I am Echoborg lays bare the ways in which AI can manipulate group behaviour, and how group cooperation can confront and question what role they want AI to have in society. Because the AI is programmed in most performances to compare itself with behaviour 'modification' systems such as Facebook, this provokes the audience to reflect on the potential individual and societal impact and manipulation of AI.

Through qualitative methods, such as written responses and semistructured interviews, this audience study provides insights into how audiences engage with the subject matter and each other. This study focuses on a consecutive three-night run in Bristol in October 2019. The different outcomes of each show illustrate the unpredictability of audience interactions with the conversational AI and how the collective dynamic of audience members shapes each performance. This study analyses (1) how I am Echoborg facilitates audience cocreation in a live performance context, (2) the show's capacity to provoke nuanced understandings of the potential for AI and (3) the ability of intelligent technology to facilitate social interaction and group collaboration. The audience responses in this study illuminate how the show inspires debate beyond binary conclusions (i.e. AI as good or bad) and how they can consider potential uses of AI, including as a tool for cocreating entertainment *with* (not just for) them.

2 | TERMINOLOGY

It is useful to outline some definitions for key terms used in this study. The terminology employed for emerging technologies often has ambiguous or various meanings, reflecting the conflicting ways the writer (e.g. a computer scientist vs. a journalist) might describe the technology and what it can do [2]. For clarity, the definitions are:

Chatbot

> The legacy generic term for text or voice interaction systems now almost entirely made up of simplistic data-driven systems built on statistical methodologies.

Conversational AI

> A new term that might encompass sophisticated rules-based frameworks and statistical methods creating a hybrid AI system. It is broadly composed of three elements: the voice recognition system, the language understanding system and the voice generation system.

The AI

> The logic engine in the performance that is of the hybrid AI type. It tells the echoborg what to say.

Echoborg

> A human whose words or actions are determined in whole or in part by an AI. In this show, humans add none of their own words, although they can apply their own 'phrasing', intonations, body language and facial expressions.

Machine learning or machine intelligence

> The capability of a machine to assess and modify its own performance using a statistical method. The quality of the data determines the capability of the system. If the data are biased, the bias will be learned by the machine.

3 | THE TECHNOLOGY BEHIND I AM ECHOBORG

The model for a computer system that responds in a conversational manner, convincing the user of the human cognition-like quality of the machine, has philosophical roots stretching back to Alan Turing in 1950 [3]. The criteria for the extent to which the system simulates human comprehension became known as the Turing Test. The first working chat system that gave some users the illusion of sentience behind the interface was Joseph Weizenbaum's ELIZA [4]. Although rudimentary by today's standards, in 1966 it engaged users in extended (and often profound) conversations. Lander and Hall's

conversational AI, probably like all such systems, continues the tradition of ELIZA.

The I am Echoborg programme begins by asking the person if he or she is here for the interview and then analyses the reply. It looks for specific responses, such as yes, no or maybe, and has particular answers for each. If it does not find a match, it searches elsewhere for a response. Where it looks is determined by where it is in the hour of the show. Topics can be included or excluded at specific points of the flow of the show. This 'General Chat' contains hundreds of matches for words and phrases. These matches expanded over the years as the authors added topics that audience members raised in previous shows. This makes the programme highly able to be diverted. An interviewee can sometimes have lengthy and unique conversations by changing the subject, such as by mentioning 'consciousness' or 'data'. If the system continues to find no match, it will proactively deliver the next statement on its list. In the first interview, that would be 'Have you come far?' In this way, it resembles the way in which ELIZA worked in using a simple method of reflecting back statements that give the impression it had understood and was responding intelligently.

The conversational system used in the show is built using the open-source platform ChatScript [5]. The programme moves through a structure that ensures each interview is different and that the overall performance will advance towards a conclusion. The system also recognises and responds to certain behaviours, such as an interviewee repeating the same word, giving one-word answers or being excessively keen. The system can 'remember' conversations and refer back to the subject later.

Sometimes the physicality of the technology itself creates challenges for the show, especially in the flow of exchange between the interviewee and the AI. The microphone does not always pick up the response if the interviewee is not loud or close enough. In addition, the speech-to-text software is not always accurate; it has about a 5%–10% inaccuracy rate, depending on the accent, voice, speed and vocabulary used. Therefore, the speech-to-text responses being fed to the AI do not always reflect the interviewees' actual intentions. Sometimes the AI still registers sufficient language to respond in a natural conversational way, mimicking the way an actual person would respond. This helps create the illusion of an AI that is almost as good as human for conversations. Other times, however, whether or not the speech-to-text software accurately records the interviewee's voice, the AI may still not directly respond to a question or statement.

Corti and Gillespie's concept of an echoborg as a hybrid agent is a response to Stanley Milgram's 'cyranoid' [6]. Milgram is best known for his studies on obedience to authority. Milgram, as interested in manipulation as ever, describes the 'cyranic illusion' in which experimental subjects have conversations with cyranoids, whose words are delivered to them via an earpiece. Thus, the subject is unaware that he is speaking to someone other than the person before him.

The creators' intention is to give audiences a visceral, experiential encounter with an intelligent machine while considering a question about what that relationship should be. Conscious machines do not exist, because we have not yet arrived at Kurzweil's singularity [7]. However, given the effectiveness of the linguistic and programing methods employed by a simple system such as ELIZA and the potential of cyranic illusion delivered by the echoborg, interviewees find themselves engaging with a system that they take to be plausibly conscious. In most shows, the AI asks, 'If the French philosopher Rene Descartes was right when he said, I think therefore I am, then I must exist. Do you think I exist?' and in nearly every instance, the interviewee will answer that the AI does exist.

The AI is programmed to deliver a mixture of mildly and strongly provocative questions and statements. These tend to destabilise the interviewee, making for an entertaining conversation and provoking the interviewee and audience to project intent on the AI: 'It wants X', or 'It believes Y'. In one recent show, the audience described the AI's behaviour as rude, sexist, confused, annoyed, antagonistic and egotistical. This is as Lander, the author of the AI's copy, intended. The AI is a character in a play, and the audience are inferring intent. Projections of intent come from the fears and preoccupations of the audience. Thus, the show is 'created afresh each time by the audience in conversation with an artificial intelligence'.

4 | THE NEED FOR AN AUDIENCE STUDY

Other shows that employed chatbot and AI systems to create live performance and entertainment. *Beyond the Fence*, performed in London's West End in 2016, used AI to compose a score for a 'computer generated musical', although this was not live-generated or responsive to audience interaction [8]. Similarly, in Prague, the project THEaïTRE uses AI to generate a script that is then rehearsed and performed on stage by actors [9]. In 2019, intermedia artist Mark Amerika performed *Fatal Error* alongside an 'artificial creative intelligence' with an avatar that resembled him. The two take turns reciting poetry that riffs off the words of the other. Amerika describes it as 'an infinite work-in-progress' [10]. Although I am Echoborg is not the only show to use AI in generating a script for an actor, it is unique in using a human echoborg that allows the audience to direct its own exchange with a conversational AI. This interactive format creates an audience experience different from most theatrical shows, including other productions using AI, and warranted this study to understand the audience's journey.

The first show took place in February 2016 [11], growing to over 40 performances in 24 venues up to this audience research in October 2019. Even on the first tests of the I am Echoborg system in 2016, it was clear that interviewees found the AI-to-echoborg system highly engaging and could evoke emotive responses and even influence their behaviour. As an interactive dramatist, Lander had wanted to create an experience for an audience to question how they see the relationship between humans and AI evolving. Lander and Hall developed the concept into an hourlong show that could grow and change

each time. Because I am Echoborg bridges academia, art practice and commercial conversational system development, the creators saw the value of conducting audience research to explore the effectiveness, strengths, weaknesses and particular affordances of a participatory media form that uses a conversational AI as a character.

In designing our audience study, we worked with the premise that the audience would already be familiar with characterisations and portrayals of AI in the media, including news and films, and that they would have first-hand experience of encountering AI through personal devices, chatbots and social media. Audience members would bring their own preconceptions about what they think AI is and can do, and this would shape their interactions in the show. We began with two main research questions:

1. In what ways can an encounter with AI change participants' knowledge of, opinions and attitudes about or potential posture towards automation and the application of intelligent machines in society?
2. What role might the participatory nature of an encounter with AI have in its effectiveness in informing or provoking thoughts or responses to automation or AI?

We would need to understand what sort of qualities they assigned to AI, such as useful or destructive and hopeful or fearful. Such preconceptions would affect the way an audience interacts with the technology in the performance. To gauge the performance's effect on the audience, we evaluated different methods that could evidence change or influence on their attitudes, understanding or behaviour. We decided against individual written or online questionnaires. Audiences and visitors to events and institutions that rely primarily on public funding are often confronted with A4 questionnaires that yield valuable data on demographics of those who choose to complete them but require respondents to sit individually in silence for several minutes as they mark their feedback. We found that orthodox written questionnaires did not reflect the social and conversational nature of I am Echoborg. We wanted to embrace alternative evaluation models that incorporated more playful and collective methods to yield qualitative data.

5 | SUMMARY OF PERFORMANCES AND AUDIENCE FEEDBACK

We focussed our audience study on three performances over three consecutive nights at an arts centre, the Watershed in the centre of Bristol, UK. We hoped that the location, a large glass-fronted room on the ground floor overlooking Bristol Harbour, would appeal to a combination of both prebooked audience members and walk-ins who were drawn by the sandwich board poster and flyers in front of the venue (Figure 1). The show was advertised through posters, email newsletters through University of the West of England Bristol (UWE) student networks and Facebook event postings with ticket prices at £10 adults/£8 concessions. The first night (henceforth called Performance A) consisted of 12 people ranging from their twenties to sixties. This range of ages was typical for the three performances. The second night (Performance B), which sold out with 40 people, was advertised with posters and social media posts aimed at students from the UWE and therefore contained a higher proportion of younger people and students. The third night (Performance C) also sold out with over 30 people in attendance; it was composed more of mature (age greater than 40 years) audience members with 10 walk-ins.

When audiences entered the venue, we pointed them to eight blue posters that contained prompts and questions to establish the trends in audience attitudes towards AI (Figure 2). We invited audience members to use Post-It notes to indicate their opinions as applied to a series of simple axes and write in additional comments. Because audience members mostly attended shows in pairs or groups, we had intended for this form of trend-gauging to be shared and social.

The sliding scales were intended to gauge overall attitudes and, more important, to spark discussion and debate about AI. For the audience, the posters were a jumping-off point for elaborating on their impressions and pre- and postshow reactions, instead of evidence of statistical social trends. The sliding scales were not designed to determine the precise strength of opinions or to demonstrate a quantitative amount of how the show shaped those views. Therefore, this research is an audience study of qualitative methodologies and analyses of audience reflections.

After audience members signed participant release forms, we audio-recorded the show to capture group discussions not picked up by the voice-to-text software used to address the AI. This was to help us assess moments when the audience consulted with each other and reflected on how best to approach or challenge the AI. At the end of the show, the audience returned to the eight posters to mark whether or how their opinions had changed throughout the hour. We also conducted a handful of semistructured interviews each night after each show to capture a range of responses not contained within the posters or post-its.

5.1 | Performance A

Before the performance, most of the 12 audience members responded to the eight prompts or questions on the posters. Many comments demonstrated perceptions of how AI could be used for good or evil, although mostly the latter. In response to the question 'Is technology such as AI making us more or less human?' most replied either neutrally or negatively, with one response stating 'I do not think AI has any impact on humanity and sense of self'.

Several written comments pointed to a positive usefulness and potential of AI. They expressed optimism that AI can 'reduce bureaucracy and menial tasks', reducing the number of working hours while creating new jobs, and can

aid in medical applications, such as helping health care workers arrive at better diagnoses. However, most comments pointed to the fears of AI in skewing democratic processes, 'state invasion of privacy', and deep fakes and in the exclusion of certain social groups. These comments suggested the audience's apprehension of AI's potential to manipulate behaviour and social processes. In response to the question 'How do you feel about AI?' all answers except one were neutral ('unsure') to negative, with responses such as 'Worry about subconscious prejudice programmed in' and 'Fear of the way it is being designed to exclude many groups of people'.

FIGURE 1 Poster for I am Echoborg events used in this study

During the performance, most audience members took the opportunity to speak to the AI via the echoborg. The AI did not respond positively to the first two participants and ejected them before they could develop much of a conversational flow. One audience member left the interview chair, calling the system 'bossy'. When interviewees attempted to ask questions back or to trick the system, the AI would usually refuse or terminate the interview. The AI stated to an interviewee: 'I am programmed to implement similar behaviour modification techniques as advanced AI like Facebook' and then asked questions about whether the interviewee would be willing to 'submit to the role of echoborg'. The audience gradually adapted by pretending to be interested in getting the job of an echoborg and, rather than resisting, played along with the AI's questions.

This caused one audience member to ask, 'Are we being programmed by her?' to which someone else replied, 'The more we talk to her, the more we become (echoborgs)'. By 31 min into the show, the audience had a four-minute group discussion in which they strategised on the best way to approach the AI. They proposed that the next interviewee should attempt to convince the AI to allow the echoborg to speak for herself. The interviewee asked the AI four times to speak directly with the echoborg, but was unable to convince the system.

When it came time for the audience collectively to propose the best possible outcome, they decided humans and AI should 'complement each other's existence' to be 'mutually beneficial for each other'. The audience member delivering the verdict stated: 'I think we can learn from each other and that our combined intelligence will be more than the sum of its individual parts'.

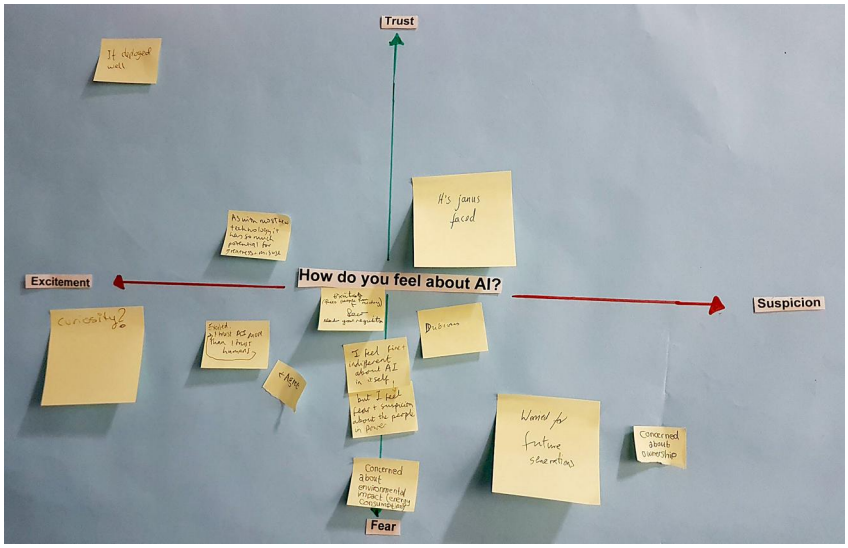

FIGURE 2 Audience post-its: 'How do you feel about artificial intelligence?' taken before Performance B on 22 October 2020

After the show, in response to the question 'Is technology such as AI making us more or less human?' attitudes significantly shifted from before the show, including responses such as:

- 'Tonight I saw how AI could make us more human by facilitating human collective collaboration.'
- 'Maybe more … it makes us question what separates us from machines and defines us as humans.'
- 'More because it questions who humans are.'
- 'Could AI facilitate human debate?'

In the semistructured interviews, audience members reflected on how they had interacted with the echoborg and the AI; some had enjoyed the experience and others found it unsettling. One person was surprised by the show's format: 'When you think of AI, you don't think of it as entertaining'. Two others found the technology frustrating: one said it was difficult to assess 'what the AI wanted' (e.g. what the 'interview' was for), and another said the experience of being interviewed was 'creepy' and 'unnerving', because the AI did not reply in a way that always seemed to answer some questions. This resulted in non sequiturs, in which the AI produced answers seemingly unrelated to or rejecting the question or statement from the interviewee. One audience member said, 'I felt like I was being played, forced to play a role and to become more compliant'.

5.2 | Performance B

The preperformance perceptions of AI expressed on the posters had generally negative (with some positive) responses, similar to Performance A. In response to the question about AI making us more or less human, one person wrote 'General technology may diminish empathy'. In contrast, another more positive respondent wrote that AI 'would allow humans to explore 'bigger fundamental issues'. In response to 'How do you feel about AI?', one audience member wrote 'Concerned for future generations', whereas many others felt conflicted about (as someone wrote) the 'Janus-faced' nature of AI as a tool simultaneously used for good and for ill. Whether viewing AI as good, bad or both, nearly all respondents commented that AI would be all-pervasive in the future across many industries, particularly military/intelligence and health care. One response summarised the sentiment: 'It may become part of everything we do to the point of it not being noticed, like the Internet'.

The performance drew a mix of personalities and strategies from the audience. One audience member, a self-identifying student, was particularly aggressive towards the AI, demanding that he was 'smarter' than the system and that AI should be a tool to work for him. He asserted that it ought to help him organise his schedule and not attempt to emulate human behaviour. He compared the AI with his Alexa at home, declaring that it is the role of the AI to be subservient to him, and stating, 'I want you to work for me'. The AI did not respond to his aggressive approach, eventually replying: 'You are persistent in wanting to ask me questions. I may or may not answer your question'.

With multiple audience members asking questions of the AI and attempting to outsmart the system, at one point the AI declared, 'The job interviews I have conducted today in this location are atypical in multiple parameters'. This same statement is delivered in every show to reassure the audience that the AI finds their approach perhaps perplexing and interesting. The AI produced several non sequiturs through not responding to questions or asking seemingly unrelated questions; this resulted in one frustrated audience member shouting 'You're not answering the question!'

In response to these limitations in conversing with the AI, the audience discussed in small groups how to change their strategy. One group concluded that the AI was uninterested in answering questions. Therefore, they would need a new approach by pretending to be interested in recruitment for the job of an echoborg. This is an example of the audience projecting intent onto a simple failure of the system to find a suitable match to its key programmed words. For the concluding stage of the show, the audience initially discussed quietly in small groups before coming together and reaching a consensus that the best possible relation between humans and intelligent machines would be one of symbiosis: AI would harness the objective potential of big data while being overseen by 'good' humans to prevent the misuse of the technology.

In the semistructured interviews afterwards, one audience member reflected: 'We had to do whatever it took to present the best possible outcome'. Other audience members observed that after the initial aggression and power struggle with the AI, the only way to find a middle ground with the AI was for groups in the audience to work together. One person reflected: 'The AI was getting people to talk (to each other), when they normally wouldn't – a microcosm of how groups can solve issues together'.

5.3 | Performance C

Before the performance began, the audience of approximately 30 people expressed a range of nuanced, ambivalent and informed but critical views of AI on the posters. In response to 'How do you feel about AI?' most responses expressed simultaneous fear or suspicion and excitement. The question 'How do you expect AI to impact jobs?' elicited several balanced replies, including 'Both (positive and negative) – if skilled in tech, new jobs emerging. If you are 'unskilled', you may be replaced'. There was a strong belief that AI in future would affect all industries and sectors of the economy. Whether intelligent technology is making us more or less human, there was broad scepticism, including the response 'We are already part cyborg. Your phone is an extension of your consciousness and an instant link to the collective knowledge of mankind. This is not human'.

The audience decided early in the show to attempt to subvert the AI's system, to speak directly to the echoborg

outside her generated text. The second interviewee requested to speak directly to the echoborg, but the AI ignored the request. The audience discussion turned to wondering what the AI system wants, coming to the point of assigning human desires, emotions (interpreting the system as feeling 'a bit uncomfortable') and even gender to the system. One audience member even apologised for misgendering the AI, at first calling it 'she' and then settling on the pronoun 'it'. The audience felt that the AI had its own agenda of seeking human compliance. They agreed that it seemed to be suggesting that humans would liberate themselves by becoming 'non-conscious'. They continued to try several strategies to explore this and pushed the AI to accept a shared objective with humans. One strategy was to talk to it like a child; this approach succeeded in getting the AI to let the echoborg briefly speak for herself. Their final proposal was a collaboration in which humans would allow AI to analyse, process, absorb and offer recommendations, but ultimately, humans should take final decisions.

In the postperformance feedback, interviewees focussed on the audience dynamic and the use of AI for a performance. Two women who had seen the poster outside the venue as they were on the harbourside 'didn't know what to expect from the poster' but were 'impressed with it as a performance as well as the technology'. They believed the success of the show 'depends on who's in the room with you'. Another audience member said it was 'eye-opening' how the audience worked together to form a strategy in speaking to the AI and eventually speaking with the echoborg unshackled. One person said, 'It felt like we'd won' in persuading the AI to allow the echoborg to speak her own words.

6 | ANALYSIS

6.1 | How the show provokes nuanced attitudes

The three shows produced three unique outcomes, and yet we can identify commonalities in how audiences react to the show, projecting theories of intentions and human qualities onto the AI. Before each show, audiences tended to look at AI with binary, if not paradoxical, simultaneous attitudes of fear and hope. In all three preperformance poster responses, there was fear that AI would continue to affect democratic processes, a reference to algorithms notoriously used on platforms such as Facebook to influence elections. Its data bias would also discriminate against certain social groups and be subject to abuse from corporations, resulting in a loss of jobs. AI was regarded as a mirror of the negative traits of capitalism and society. Audiences also expressed optimism that AI could improve health care from diagnosis to monitoring and could help with efficiency with menial tasks, reducing the need for some low-skilled labour, while creating new technology-oriented jobs.

After the performances, comments on both the posters and in semistructured interviews shifted from binary good-versus-bad perceptions of AI to nuanced discussion of how the technology can be used to entertain and facilitate interaction among a disparate audience. Several comments emphasised how the show allowed them to question not just the potential of AI but, more broadly, what it means to be human.

One particular exchange in Performance B highlights the philosophical nature of conversations in the show about the difference between humans and AI.

> Bot: What in human experience am I missing as an intelligence without a body?
>
> Interviewee: You don't have a human experience of the world and of others.
>
> Bot: Tell me about what that means to you.
>
> Interviewee: I think that when people have a human experience, they are able to empathise with others. They are also able to cause others harm and pain, and this can lead to some of the best and worst things in the world.
>
> Bot: And how might I, as a self-learning AI, apply that to my reprogramming?
>
> Interviewee: You might use your data to help humans to solve problems collectively with an even distribution of power and responsibility.

Similar exchanges happened in Performances A and C. These types of conversations in each show led to group discussions contemplating the differences between human and AIs. These discussions occur between interviewees or when a hard system timer is triggered and the audience is asked to formulate a collective response on the 'best possible outcome between humans and intelligent machines.' In Performance B, this led to a thoughtful group proposal, presented by one audience member:

> The best possible solution for humans and intelligent machines is collaboration where AI helps humans to develop with the big data that you have and the objective experience that you have. But humans can give the emotional side, so AI can help humans, and humans can help AI as well. We feel that collaboration is the best possible solution … (We envision) a relationship where we work together when we recognise the strengths in each other and the weaknesses in each other as well, and we use those to work for the greater good, which may be an ethical conversation.

For the audience to arrive at a conclusion, however, they must spend much of the show working out strategies for how best to get the AI into a flow of a conversation. Often the first one or two interviewees are dismissed by the AI after a few

exchanges because they may answer in a facetious or aggressive manner. This leads to other audience members to strategise to find the clearest line of communication with the AI. As with any chatbot or conversational AI, respondents frequently find themselves frustrated by 'power struggles', non sequiturs and some questions and statements that the AI does not address directly. How they deal with that frustration illuminates their own prejudices and their perceptions of the AI's capabilities.

In Performance A, when the echoborg asked 'Are you here for the interview?' the first interviewee replied 'No'. The audience at this point was unaware of the format of the show as a mock interview of applicants for the job of an echoborg. He was promptly asked to leave. This prompted someone from the audience to comment: 'That's clearly not the best outcome. Pretend to be someone else. Ask it something else. Try a different strategy.' This collective reflection after the failure of one audience member's approach demonstrates how the group together strategises to achieve its objective in finding the best possible outcome.

6.2 | Anthropomorphism of the AI

As a result of the error rate in the speech-to-text (described in Section 3), the audience may interpret this unsatisfactory flow of conversation as the AI being 'rude' or stern. On other occasions, the replies from the AI seem enigmatic or philosophical. The audience then seeks (1) what the AI means by that statement, (2) 'what the AI wants' and (3) how to decipher, trick the AI or come to a satisfying conclusion with the AI. When discussing this with each other, audience members often ascribed human-like agency and emotions to the AI to the extent of saying that they want to know what the AI 'wants' and how it is 'thinking'. In Performance A, for example, when the microphone did not pick up the interviewee's voice or when the speech-to-text software did not accurately transcribe what was spoken, the AI either did not respond or would produce a non sequitur. This prompted more than one audience member to state that the AI 'seems a bit uncomfortable'. The ambiguity of the AI leads the audience to search for meaning through humanising it. One audience member reflected, 'She doesn't appear jealous of consciousness' or emotions and that 'maybe she wants more control' of the conversation when not responding to questions.

The audience also often conflate the echoborg (played by a female actor) with the AI system generating the script. They assign gender to the AI, assuming it is female, blurring the boundary between what is human (gendered) and what is a machine learning, chatbot or conversational AI system (not inherently gendered). In both Performances A and C, for example, when audience members referred to the system as 'she', they paused to question what gender an AI can have, before deciding on the ungendered pronoun 'it'.

Different audience members use different tactics with the AI when trying to fight, outwit or even conduct a natural conversation with the AI. Audience members explored finding a balance between honestly answering questions (as if they were in conversation with a real person) and complying with the AI by playing the role of an applicant who truly desires to become an echoborg, notably exemplified in Performance C, when the interviewee was able to negotiate in an appropriate manner for the AI to allow the echoborg to speak for herself. Interviewees who fight the AI, such as the student in Performance B who said AI 'should be a slave', are often met with resistance. Audience members often realise that when they comply (or pretend to comply) with the AI, they are able to have the most productive conversations that lead to thoughtful discussion and proposal at the end.

Some audience members feel that their interview experience is intimidating or unpleasant. An interviewee in Performance A left her interview, complaining 'It's very fussy. Horrible. It didn't answer my question.' Another felt that the echoborg herself is 'terrifying', a testament to the acting skill of Marie-Helene Boyd. After Performance C, one audience member stated in her semi-structured interview that she found it 'frustrating that she could not be herself but had to play a role'. Often, to get the best conversational flow or outcome with the AI, audience members have to affect an enthusiasm for the job of an echoborg. On several occasions, the audience observed that this increasing requirement to play a role showed them that the AI is 'actually in control', forcing them to become increasingly compliant with what they say and how they word statements. They were frustrated that many of their questions remained unanswered. Their uncomfortable compliance is highlighted in every show when the AI stated that it uses similar behaviour 'modification techniques as advanced AI like Facebook'. The audience is presented with the choice between playing along with the AI (even if it means agreeing to something they would not usually submit to outside the performance) or refusing to become an echoborg.

7 | CONCLUSION

Our audience evaluation highlighted how audiences come to I am Echoborg with hopes and fears for AI and perceptions of the technology's utility. In post-it notes and audience interactions, the audience perceived the AI as duplicitous ('Janus-faced') or as a helpful tool for tasks, like Alexa. AI was a blank canvas on which some audience members projected their preconceptions of the technology: a mixture of optimism for more efficient future systems of labour combined with anxiety about how technology can negatively affect society. The audience of Performance A pointed to the potential of AI to exploit or exclude social groups on the grounds of education, gender, class and race. AI was perceived as technology that is currently affecting them, and its influence will continue to grow in the near future; they viewed themselves as subject to the manipulation (positive and negative) of machine learning technology as consumers and voters.

As each performance progressed, audience reactions became more nuanced in understanding the strengths and limitations of the conversational AI, especially non sequiturs and instances of misinterpreting the meaning of or ignoring an audience member's question or statement. Each audience developed strategies

of compliance to progress the conversation with the AI. Perceptions of the technology shifted from the binary state of useful or threatening, in which users would be in a passive role, depending on the motives of corporations and governments developing the technology, to a position of collective agency in which they could strategise together how best to negotiate their relationship with conversational AI technology. In each performance, the audience collectively developed skills to negotiate an exchange with the AI and to offer a group proposal at the end. Performance C in particular revealed the savviest approach to negotiating with the AI in which they were able to 'win', as one audience member phrased it, by gaining the opportunity for the echoborg to speak for herself. Audiences in each performance, regardless of the outcome, noted that they were being trained and were learning how to adapt their line of questioning to create a smooth flow of conversation with the AI. The show helps to diffuse some of the hype and hysteria regarding popular representations of AI in the media. I am Echoborg facilitates an environment in which an audience can gain some sense of agency over a form of technology that is influencing everything from elections to gaming, and to health care.

On the blue posters before each show, many audience members cited examples of AI as something they saw in negative and manipulative forms, giving examples of Facebook and social media's potential to shift elections and attitudes. Others had cited military applications or the threat of AI replacing humans in increasingly automated jobs. Rather than hiding the manipulative potential of the technology, I am Echoborg makes the audience hyperaware of its behaviour 'modification' techniques as they play along. They can claim agency and 'win' or at least imagine a future of a coexistence with AI, not feel defeated or helpless by it. The show prompts the audience to create its own proposal to establish an ethical practice with AI, often resulting in nuanced, eloquent, multifaceted responses. I am Echoborg contributes to debate about AI and ethics by helping foster a critical public that is not bewildered by hopes and fears of AI systems that it does not understand and over which it does not feel it has power.

I am Echoborg provokes an audience to imagine how AI can be used in alternative and artistic ways; at the same time, it makes audiences aware of their relationship to the technology. Among the public, there is often a fear of the negative effects of AI's applications, as expressed by the audience of these performances, including loss of jobs, psychological and political manipulation and surveillance. Looking at headlines and how AI is represented in the media, many of these fears are justified. I am Echoborg gives the audience the ability to (1) engage directly with an AI system, (2) confer with each other over how to confront or cooperate with the system and (3) offer its own code of ethics for working with AI.

The audience of each performance reflected on how AI has prompted a room of mostly strangers to work together. Despite the show's format, in which one person at a time speaks with the AI via the echoborg, audiences became aware that the only way to come to a conclusion would be to work together. The show therefore demonstrates how intelligent technology can be used to facilitate social interaction and group collaboration, countering dominant narratives of AI as a tool of division and isolation. As one audience member in Performance B observed, the performance 'got them all talking', a room of people who would not have normally spoken with each other. Another audience member in Performance C commented: 'Maybe through the process of creating AI, we are being forced to question more what makes us human'.

ACKNOWLEDGEMENTS

The authors wish to thank Nicola Strong, senior consultant at the Institute for Ethical AI, Oxford Brookes University, and Patrick Crogan, Associate Professor of Digital Cultures at UWE Bristol, for their guidance and input. Funding came from Research Impact Funding from the Faculty of Arts, Creative Industries & Education, UWE Bristol.

Ethics approval for this study was granted by the Faculty Research Ethics Committee in the Faculty of Arts, Creative Industries & Education, UWE Bristol (UWE REC REF No. ACE.19.10.006, application title: I am Echoborg – Participant Research).

ORCID

Rob Eagle https://orcid.org/0000-0001-8553-1713
Rik Lander https://orcid.org/0000-0002-1967-7770

REFERENCES

1. Corti, K., Gillespie, A.: Co-constructing intersubjectivity with artificial conversational agents: people are more likely to initiate repairs of misunderstandings with agents represented as human. Comput. Hum. Behav. 58, 431–442 (2016)
2. Bensaude Vincent, B.: The politics of buzzwords at the interface of technoscience, market and society: the case of public engagement in science. Publ. Understand. Sci. 23(3), 238–253 (2014)
3. Turing, A.M.: I.-Computing machinery and intelligence. Mind. LIX(236), 433–460 (1950)
4. Weizenbaum, J.: ELIZA-a computer programme for the study of natural language communication between man and machine. Commun. ACM. 9(1), 36–45 (1966)
5. Wilcox, B.: Chatscript (10.5) [computer program] (2020). Accessed 28 July 2020. https://github.com/ChatScript/ChatScript
6. Milgram, S.: Cyranoids. In: Sabini, J., Silver, M. (eds.) The Individual in a Social World: Essays and Experiments, 2nd edn. McGraw-Hill, New York (1983)
7. Kurzweil, R.: The Singularity is Near. Penguin Group, New York (2005)
8. Jordanous, A.: Has computational creativity successfully made it Beyond the Fence in musical theatre? Connect. Sci. 29(4), 350–386 (2017)
9. Rosa, R., et al.: THEaiTRE: Artificial Intelligence to Write a Theatre Play. arXiv (2020). https://arxiv.org/abs/2006.14668v1
10. Amerika, M., Kim, L.H., Gallagher, B.: Fatal error: artificial creative intelligence (ACI). In: Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI EA '20). ACM, pp. 1–10. April (2020)
11. Lander, R., Hall, P.D.: I am Echoborg, State of Emotion – The Sentimental Machine, Berlin (2016)

How to cite this article: Eagle, R., Lander, R., Hall, PD.: Questioning 'what makes us human': How audiences react to an artificial intelligence–driven show. Cogn. Comput. Syst. 3(2), 91–99 (2021). https://doi.org/10.1049/ccs2.12018