

# Talking to a Toaster

How Everyday Interactions with Digital Voice Assistants Resist a Return to the Individual

Morana Alač,\* Yelena Gluzman, Tiffany Aflatoun, Adil Bari, Buhang Jing, and German Mozqueda

## Abstract

Discussions of the problematic relationship between AI and society have recently only heightened. These discussions, nevertheless, remain partial until they take into account how we live with AI technologies in the unremarkable circumstances of our everyday affairs. In arguing for the importance of such a “noticing,” this article centers on the internet of things and associated digital voice assistants (DVAs). These commercial *social robots*, designed as conversation-oriented devices, manifest their *incompleteness* in their need for other voices. Paying attention to that relationality at an embodied scale of analysis brings up our involvement in *situated interactional production*, while also manifesting its reciprocal character. This not only puts into question the conviction of DVA designers that these gadgets will generate effects of presence in relation to an intentional mind, but also gives us the resources to resist a parallel return to the individual that more often transpires in the discussions of the problematic relationship between AI and society. We practice this resistance by evoking efforts in *distributed cognition* and the *extended mind hypothesis*, but we also go beyond the instrumentalist reasoning that primarily recognizes the world as carved into *convenient tools* that can *extend our cognition*. To do so we focus on the achieved quality of *bodies* and *environments*—two constitutive elements of DVA technology—thereby pointing out how the *self* in the context of the voiced AI importantly derives from the *openness* between humans and machines in the interactional scenes of which they are a part.

## Keywords

Artificial intelligence, social robots, digital voice assistants, distributed cognition, situated interaction, openness

\* Department of Communication and Science Studies, University of California, San Diego. [alac@ucsd.edu](mailto:alac@ucsd.edu).

## Introduction

This text considers instances of ordinary interaction with AI in the context of the internet of things (IoT) and digital voice assistants (DVAs)—voice-controlled AI software embedded within a variety of commercial products, and colloquially familiar as Amazon’s Alexa, Google Assistant, Apple’s Siri, etc. DVAs are animated by an idea of a hypersocial world where everything is to be social, and we are to live in kitchens where toasters converse with us. That sense of the social is conceived of as unproblematic, and premised on individual selves as its elementary units. We point out that the notion of the individual not only supports the conceptualization of social artificial intelligence (social AI) in its design, but articulates many critiques of AI. We challenge this by turning to relational bodies and their groundedness in lived environments.

In the opening section of this article, we lay out our orientation toward contingencies of everyday situations, which we fully develop in sections three to five. There, we zoom in on instances of interaction with DVAs where these machines are encountered in dormitories of an American university as mundane aspects of everyday life. But before turning to interactional scenes, in section two, we provide an account of DVAs’ relationship to *social robots*, their historical predecessors whose ontological commitments DVAs still carry. Tracing this genealogy allows us to illustrate *how* notions of *body* and *environment* are at play as two constitutive elements of DVAs’ design, where at stake is an idea of the body that *expresses* the mind of the individual behind it while that body is *contained* in an environment assumed to be devoid of affect and inhabitation dynamics. In contrast, our analysis highlights lived, *achieved* features of machinic voices that in turn indicate our *openness-with-the-world*.

Since our thinking on this account stems from a classroom setting, we start there. In the opening section, we introduce the technology of concern, and identify the return to the individual in recent texts on the problematic relationship between AI and society. To resist this return, we repurpose arguments from the *extended mind hypothesis* and *distributed cognition* that those texts oppose.<sup>1</sup> We conclude the first section by indicating how we incline those arguments toward our interest in the everydayness of voiced AI interactions.

## 1. Voiced AI and a Return to the Individual

It is the beginning of 2019, and we are in a university class on *Communication and Social Machines*. An early classroom exercise is to read the *New York Times* article “A Future Where Everything Becomes a Computer Is as Creepy as You Feared,” published on October 10, 2018 by Farhad Manjoo. As suggested by its title, the article talks about the IoT, where objects across our environments will contain chips connected to the internet, making our homes “smart.” The article warns that now (and differently from our past attitudes), we should notice that our relationship with consumer electronics is undergoing a paradigm shift. In tracing a history beginning forty years ago, the article points out that when personal computers (as imagined by Bill Gates and Paul Allen) found their spot on every desk, and subsequently, when social media transferred how we engage others into the digital realm (as predicted by Mark Zuckerberg), these inventions of the tech industry’s “audacious founders . . . started out looking silly,” but they quietly took over our lives. This discursive set-up is to orient the reader toward the present moment when the looming threat of the IoT, if not noticed, may be inviting “a nightmarish set of security and privacy vulnerabilities.” The article’s proposed solution is government intervention, or at least oversight; in reference to Bruce Schneir’s *Click Here to Kill Everyone*,<sup>2</sup> the article argues that “only government can save us from such emerging calamities.”

Students in the class are familiar with the scene of the tech industry’s leading men inventing the technology, the government from where intervention is not currently coming (as indicated by the article stating that “no one is really doing much to stop it”), and fear—the emotion the text employs to depict our condition, and whose strength it further reinforces.<sup>3</sup> With its mention of DVAs, students also recognize the presence of the IoT in their everyday lives. While the article explains that chips with voice software can be easily inserted into objects across our living environments, these voices are presently most frequently heard from “smart speakers” and other smart “home hub devices” (in addition to smartphones).<sup>4</sup> That there is a direct link between the “everything internet” and DVAs is supported by the article’s only concrete example of the technology it discusses:

Amazon last month showed off a microwave powered by Alexa, its voice assistant. Amazon will sell the microwave for \$60, but it is also selling the chip that gives the device its smarts to other manufacturers making Alexa connectivity a just-add-water proposition for a wide variety of home appliances, like fans and toasters and coffee makers. And this week, both

Facebook and Google unveiled their own home “hub” devices that let you watch videos and perform other digital tricks by voice.

In bringing up this interfacing of technology via “voice,” the article stumbles across a critical *incompleteness* that DVAs exhibit once actually engaged. As Amazon’s microwave is “powered by Alexa, its voice assistant,” and Facebook’s and Google’s home hubs perform services by being directed through “voice,” what is at stake is a voice familiar from an *interactional* context—a voice that is toward other voices, is listened to, responded to through voice or other actions, and that talks back. Certainly, in DVA technology, voice is levied as a means of control, allowing us, for example, to turn on a microwave at a physical distance. At the same time, once there is a voice, there is also an orientation toward another that opens possibilities for *conversation*. So, we may say “thank you” to Alexa, but we do not feel compelled to thank our personal computer after doing an Internet search for the same information.<sup>5</sup> And when Alexa says “you’re welcome,” at stake is the interactional character that its designers discover in their efforts to deal with relationality of conversation. Google’s “conversation designer” Cathy Pearl points out: “we found this when we were designing phone systems . . . most people want to say goodbye and they want the system to say goodbye, that’s just what is built into us as politeness, and you need to think about that when designing conversational systems as well.”<sup>6</sup>

During the classroom discussion<sup>7</sup> of the *New York Times* article, this voiced character of the technology came up. While there was strong agreement with the article’s conclusion—that IoT must urgently be regulated by government’s policies—we were also intrigued by the voiced character of these machines. When one student asked, “I personally think, why do you need to talk to your toaster when you can just click a button?” the question turned the class’s attention to where voices are engaged, namely, to the everyday of interaction with digital technologies. For the remainder of the course, we maintained this curiosity, while also staying attentive to market logic, urgently needed regulations, and government oversight. Unearthing details of everyday interaction, and paying attention to lived settings in which talk takes place, allowed us to realize that when “controlling” our digital devices through voice, we also turn to technology in a manner that has the potential to open up a space for participation.<sup>8</sup>

In times when AI’s problematic relationship with society is heightened—in regard to issues such as privacy, possibilities for ideological manipulation, fragmentation into homogeneous social groups, restrictions of critical public discourse, profiling, propagation of cultural stereotypes and other forms of bias<sup>9</sup>—we suggest that these discussions cannot ignore

how AI technologies are ordinarily engaged. While we do not propose an alternative to the other two pillars that the *New York Times* article highlights—tech industry and the government—we argue for attentiveness to mundane engagements with AI as the triangle’s third side. Doing so means unearthing *production*—a production that is lived and interactional, and which shows how talking machines—not complete before their participation in actual events of talk—ask for our involvement, while in turn *articulating us* through those encounters as *open*.<sup>10</sup> In other words, the *interactional production* we witnessed is *bi-directional*: as we describe how the voiced technology is enacted through our engagement with it (rather than being *designed* a priori), we also describe how these enactments condition us as dynamically opened to them (rather than being complete in reference to a psychological individuality).

In foregrounding this involvement in DVAs’ functioning, and by pointing out that we cannot be dismissed as “cultural dopes,”<sup>11</sup> the “market,” or mere “users”<sup>12</sup> of AI technology, this article aligns with *science and technology studies*,<sup>13</sup> and, specifically, its recent discussions of surveillance and algorithms.<sup>14</sup> However, our paying attention to *haecceities*<sup>15</sup> that characterize ordinary events of technological engagement not only reveals some of the logic behind these technologies, but provides opportunities for going beyond the human-only world, privileging instead relations, events, and situations. Associating our efforts with *ethnomethodology*<sup>16</sup> and ethnomethodological work on interaction with technology<sup>17</sup> opens opportunities for questioning the centrality of the individual as the unit of analysis in AI discussions.

DVAs are objects designed to generate effects of self (as we illustrate in the following section), while, at the same time, their form (i.e., the voice)—enrolled to generate those impressions—enmeshes them in *dynamic couplings* (as we point out when we turn to details of everyday interaction). For this reason, the scale of analysis can make a dramatic difference, as tracing how we interact with DVAs puts into question conceptions of *self* designed into these machines, while, in turn, opening possibilities for reflecting upon us, whom these machines are to model. Literature preoccupied with the adverse effects of AI on society has often fallen back on traditional Western conceptions of the individual where *self* is conceived as its attribute. While this may be a consequence of an inattention to the issue (as arguments are directed elsewhere, namely to the level of social structures where extraordinary moments are of main concern), some authors explicitly propose a return to the individual.<sup>18</sup> Shoshana Zuboff, for example, talks about “the ‘collision’ between the centuries-old historical processes of individualization that shape our experience as

self-determining individuals, and the harsh social habitat produced by a decades-old regime of neoliberal market economics in which our sense of self-worth and needs for self-determination are routinely thwarted,”<sup>19</sup> and points out that *surveillance capitalism* threatens “a new society of people born to a sense of psychological individuality, with its double-edged birth-right of liberation and necessity.”<sup>20</sup> Since Michael Patrick Lynch’s discussion of epistemology in the age of “Google-knowing” has a sharply defined philosophical target—and one that exhibits alignments with instances of the everyday interaction we analyze—we will illustrate Lynch’s argument in more detail now.<sup>21</sup>

Lynch imagines (as a thought experiment) a future where we would live with “subcellular computing device” implants—what he calls “neuro-media” (defining them also as “cell-phones that are hooked directly into a person’s brain”)<sup>22</sup>—and points out that this “possibility raises some disquieting questions about society, identity and the mind”: “How is information technology affecting what we know and how we know it? And what happens to society if we not only know more about the world but the world knows more about us?”<sup>23</sup> In asking these questions, Lynch targets literature on the *extended mind*. He cites Andy Clark and David Chalmers,<sup>24</sup> and provides the example of the phone’s “Notes” app functioning as a part of one’s remembering for what to look for in a store by consulting a shopping list on one’s phone. While recognizing that “the cyborging of the human . . . is no longer just fantasy,” and that “the extended mind hypothesis is undoubtedly interesting, and . . . may just be true,”<sup>25</sup> Lynch expresses his hesitation concerning this view, as shown in his hierarchy of modes of knowing and his highlighting of the individual’s centrality.<sup>26</sup> Lynch contrasts knowing through the senses and the role of memory (which is “internal” and where one is “relying on oneself”) to knowing with neuromedia.<sup>27</sup> Correspondingly, as he stresses that “reasonableness matters” and that we must strive to provide possibilities for and sustain critical public discourse,<sup>28</sup> Lynch argues for the importance of the individual:

This brings us to the key point: whether or not a network “knows” something (even in the nonliteral sense) depends on the cognitive capacities (and incapacities) of the nodes on that network—the individual people who make it up . . . So, although networks can embody knowledge, or at least true information, not held by any particular individual, the extent to which they do so depends very much on the cognitive capacities of the individuals that make them up. *You can’t take the individual out of the equation . . .* What we *don’t* want to do is assume that because knowledge is networked, the nodes in the network—the individual knowers—no longer matter.<sup>29</sup>

The purpose of Lynch's example of neuromedia is to exaggerate the claim of the mind's extension into the world, warning us of what may happen if that phone with a grocery list gets implanted into one's brain. What Lynch's example misses, however, is the rest of the body and the world that one would engage with even if equipped with neuromedia. It is true that the extended mind hypothesis<sup>30</sup> may be meager in terms of how it accounts for the world, as pointed out by Edwin Hutchins, the author of "distributed cognition"<sup>31</sup> (a richer predecessor of the extended mind hypothesis<sup>32</sup>); however, that is not a reason to overlook the world we already inhabit. Rather than following Lynch's prescription for "getting up off the couch and plunging into the whirlpool of actual experience" (an alternative that postulates a dichotomy between digital technology and experience, inside and outside, passive and active), we pay attention to how we experience the technologies that already populate the world in which we live.<sup>33</sup>

In their focus on cognition, the extended mind hypothesis and distributed cognition have been preoccupied with processes such as problem-solving, getting things done, and enhancing one's cognitive abilities. While DVAs are designed to be convenient tools that are engaged through talk, once we start paying attention to how a *conjoint involvement in talk*<sup>34</sup> across the human-machine divide takes place, they show themselves to be pieces of the world toward which we are open interactionally (rather than tools that turn humans into more powerful cognitive agents). In that sense, the problem is less about cognition and thinking, and more about being in the world with others (broadly conceived). We accordingly inquire into the individual and its associated self—phenomena postulated on both the human and technological sides—by recalling Erving Goffman's characterization of interaction as an "*unio mystico*, a socialized trance" and by pulling that characterization toward Gilbert Simondon's discussions of *coupling and association between man and machine*, and his conceiving of machines as *open*.<sup>35</sup> Our text mixes and inclines these ideas as we want to know what happens with *openness of technology* when, in unremarkable moments of everyday life, we speak with machines positioned to exhibit a self behind their voice.

In attending to that openness in the context of interaction in ordinary situations, we notice how selves intersect with those who engage with them, while being always also intertwined with features of places we inhabit. This article, thus, describes how my self may, momentarily, be comprised of machinic self and vice versa, and how I may contain pieces of the world that, on that very occasion, involves a DVA. If we were, therefore, to adopt Lynch's vocabulary, we would conceive of his "networks'

nodes” in terms of interactional production, considering them as spaces where one is to notice openness-with-the-world. This means that when we speak about a self, it is an *interactional* self that is both produced and possibly dissolved through situated and embodied engagements across the human–machine divide. In directing our noticing to where we and the world blend, we also reflect upon how an approach—aimed at rendering the dynamicity of singular aspects of these hybrid phenomena—could be articulated. Our text is, thus, written to indicate how it was made, aiming to provide a possibility of engaging in a form of noticing that, to some extent, parallels the noticing we practiced in our class.

## 2. Self in the Design of Social Robots

While the self and the individual to which the self is imputed are at play in the literature on AI and society, they are also the conceptual core of the DVA design. To consider the notion of self in the design context, we think in terms of a *technological genealogy*,<sup>36</sup> which indicates, even in its minimal form, that DVAs descend from *social* or *sociable robots*:<sup>37</sup> instantiations of mechanical and digital life intended to enter into communicative engagements with humans, where these robots are to be positioned as social creatures, eliciting empathy from their human interlocutors. Seeing social robots as DVAs’ predecessors affords the possibility of grasping ontological commitments behind DVAs: while DVAs—constrained by market logic—may obscure the metaphysics they carry, considering their forerunners—whose design more clearly conserves a research component—makes the theoretical envisioning of DVAs more readily available. By outlining this genealogy, the present section indicates how the notions of the *body* and *environment*—two constitutive elements of DVA design—conserve a sense of psychological self that conceptually grounds them. It also lets us point out that the dynamic distribution of the self we catch in DVA interactions (described in the rest of the article) shall not be attributed to the design of these technologies (since the conceptualization behind that design is organized around an intentional understanding of the mind).

Cynthia Breazeal, who is credited with inventing the first sociable robot, Kismet, defines these machines as “able to communicate and interact with us, understand and even relate to us, in a personal way. It is a robot that is socially intelligent in a human-like way. We interact with it as if it were a person, and ultimately as a friend.”<sup>38</sup> In the vision of “social” that these robots embody, the social is not about “social species” (such as the sociality of ants), but about “human-style sociability.”<sup>39</sup> Breazeal explains

that “human-style sociability” regards treating others as having distinct personalities and being treated the same in turn; it is to have a relationship with others on an *individual* basis. According to Breazeal, this is achievable when robots’ observable behaviors can be related to the character’s internal (mental) states—such as intents, desires, beliefs, emotions—governing that behavior.<sup>40</sup> In other words, in the vision of the social that social robotics enacts, interaction is understood in reference to the idea of an intentional self behind it that serves as its organizing principle.

Social roboticists have customarily positioned their robots as testbeds for understanding “human social intelligence,” with a specific interest in atypical social behavior such as autism.<sup>41</sup> But they have also promoted them as “natural, intuitive, familiar to interact with” technologies to “enhance . . . quality of life,”<sup>42</sup> examples of which are museum tour guides and other entertainment applications, as well as household robots and nursemaids to help the elderly. That the engineering aspect of social robots has taken a fully commercial swing in voice-controlled smart hubs is supported by roboticists’ explanation that today, we are witnessing a march of social robots from the laboratory into the world. Breazeal makes this point in a recent interview:

When we started with Kismet, you know, you couldn’t—it was very very difficult to be able to have robust ASR, automatic speech recognition, you know. The machine vision was very primordial. You know, I think one of the most exciting things that’s happened more recently is just the performance and affordability of those core technologies have advanced to the state—although they’re far from perfect—they’re interesting enough that you can start to create services, products, whatever you wanna call it, that actually are interesting to people . . . But just the ability to build these systems in a way that you can actually put them in the world is hugely important, you know, and even from a research standpoint, hugely important. Because until these systems are in real environments with real people interacting with them, you don’t even know what the phenomenon is that you’re trying to develop for and understand. So, that’s been tremendous, and fairly recent, right? And even right now, I would say, whatever, 43 million [Cortana] smart speakers in people’s homes now, people are living with AI in a way they never did before. And who is interacting with AI is fundamentally different. It used to be just kind of like people like us [*gestures towards interviewer and self; he laughs*]. Now it’s young children and seniors and everyone in between, and we know very little about how to design for the diversity of needs and opportunities of such wide demographics. And we still don’t understand what this long-term relationship is gonna be, right. We build these very transactional systems, by and large, but I think the next level is kind of the Star Wars dream, you know? Robot sidekick, helpful companion. Not just a digital assistant but a helpful companion that is with you and works with you towards your goals and really is your collaborator. That’s

still, I think, where this is going. To me, that's a whole other layer to where we are right now.<sup>43</sup>

When Breazeal announces that “these very transactional systems” (such as “smart speakers in people’s homes now”) will, eventually, turn into “helpful companions” beyond “digital assistants,” she implicitly gestures toward her recent attempt to design a commercial social robot, Jibo. But while Jibo, as a more affordable social robot (in comparison to social robots proper that still live in laboratories and are accessible to their designers only), was easily crowdfunded and even featured on the cover of *Time*’s “Best Inventions of 2017” issue, it ended up failing. Regarding its fiasco, commentators noted that its promised functionality quickly appeared in products that were not burdened with the costs incurred by designing expressive movement, relying instead on voice.<sup>44</sup> Jibo was expensive because of its

unique emphasis on physical expression and the potential for a diverse array of capabilities backed by robust hardware. But by the time Jibo started ending up in homes, much of its functionality was duplicated by smartphones and other smart home devices like Siri, Alexa, and Google Home.<sup>45</sup>

An article in *The Robot Report* offered a similar account:

Soon after Breazeal successfully launched the Indiegogo campaign, Amazon unveiled Echo and priced it seventy-five percent less than Jibo. In the end, the robot that eventually shipped for close to \$1,000, was merchandised next to a series of more functional machines priced under \$100.<sup>46</sup>

From Jibo’s physical body (still not conducive to commercialization), DVAs conserve only the voice, while the rest of the robot’s outward perceivable substance is embedded in physical features of the environment. In his discussion of non-verbal vocalizations, Steven Connor pointed out how the voice, in conserving the body, generates effects of subjectivity:

For the mouth is more than just the part of the body specialized for the production of speech. It is the way in which, the place in which, the whole body may be taken up in vocality . . . There is no disembodied voice—no voice that does not have somebody, something of somebody’s body, in it . . . The voice is lived and imagined as the life of its subject.<sup>47</sup>

While for social robots the humanoid body has to do with “the physical and immediate proximate interaction that transpires between humans and robots that share the same social world,”<sup>48</sup> DVAs do not display much of that “emphasis on physical expression.” Instead, the expression intended to generate effects of a self is to be produced by the voice. While the

rest of the physical body is treated as inert, physical matter (including the material form of smart hubs—“that squat slice of black tube”<sup>49</sup>—which, with a full realization of IoT, are to be dissolved into the environment), the DVA’s expressive body is *the body-as-voice*.

In *The Atlantic*’s article “Alexa, Should We Trust You?”<sup>50</sup> Judith Schulevitz points out that as we engage digital technologies via voice—like we engage others that “have minds”—their *voices appear as if proceeding from a human interior*:

Gifted with the once uniquely human power of speech, Alexa, Google Assistant, and Siri have already become greater than the sum of their parts. They’re software, but they’re more than that, just as human consciousness is an effect of neurons and synapses but is more than that. Their speech makes us treat them as if they had a mind. “The spoken word proceeds from the human interior, and manifests human beings to one another as conscious interiors, as persons,” the late Walter Ong wrote in his classic study of oral culture, *Orality and Literacy*. These secretarial companions may be faux-conscious nonpersons, but their words give them personality and social presence . . . And indeed, these devices no longer serve solely as intermediaries, portals to e-commerce or nytimes.com. We communicate with them, not through them.<sup>51</sup>

While roboticists perceive the robot’s mirroring of the human intentional mind as a positive feature of their design, Schulevitz worries that our psychological well-being is threatened by machines who only appear to possess the mind behind their voices:

By now, most of us have grasped the dangers of allowing our most private information to be harvested, stored, and sold. We know how facial-recognition technologies have allowed authoritarian governments to spy on their own citizens; how companies disseminate and monetize our browsing habits, whereabouts, social-media interactions; how hackers can break into our home-security systems and nanny cams and steal their data or reprogram them for nefarious ends. Virtual assistants and ever smarter homes able to understand our physical and emotional states will open up new frontiers for mischief making . . . But there are subtler effects to consider as well . . . I fear other threats to our psychological well-being.<sup>52</sup>

But, while this pessimistic view is different from how technologists see the DVA future, the article embodies the same vision that roboticists evoke when framing social robots: at stake is an account that centers on the individual (with the associated self and psyche) on both sides of the postulated divide: human and nonhuman.

In the present article, however, we take a different stance, as we follow the lead author’s previous work on everyday practices in social robotics.

Morana Alač<sup>53</sup> suggested that the conceptual framing of social robots—which imply traditional Western views of “the social” as propelled from the inner self of a single individual (where the social is related to Cartesian-Kantian notions of personal identity and individual autonomy<sup>54</sup>)—can be reconsidered by turning to what takes place *beyond* robots’ original design. As a part of her noticing that contingencies of unremarkable circumstances confront us with multidirectional dependencies relative to the situated events in which robot encounters take place, Alač<sup>55</sup> has written on how spatial arrangements are components of the social as situationally achieved—that the body extends across those settings, which are, thus, part of its functioning as a social body. But, just as official framings of social robotics do not specifically consider the environment when discussing robots and their interaction with humans, an analogous take appears to be at stake in the DVA case. Also there, the environment is imagined as a stable matter that coincides with the physical layout of the smart home and its appliances, remaining an inert envelope for human–technology interactions. Doing a Google Images search for “smart home voice assistant” returns results that illustrate how designers and marketers portray these environments: there are formulaic-looking kitchens traversed by voices from nowhere, and layouts of houses, abstracted from the mess and other traces of everyday inhabitation, whose rooms are, more often, either people-empty or inhabited by a single dweller.<sup>56</sup>

That the machinic sociality—conceived as entirely attainable via voice—is designed into environments whose eventual and affective aspects are missed *is* precisely why DVAs call for an exploration beyond social robots. Because their expressive body is reduced to voice, these gadgets more apologetically manifest their openness: in contrast to designers’ expectations that they would generate impressions of self behind the voice, DVAs nudge us to notice how their speaking voices entangle them into relationality.<sup>57</sup> The DVA case is also interesting because—differently from social robots—designers, in this context, have to acknowledge the environment as a part of their design domain (even if they treat that environment as an inert, physical matter). And since, with DVAs, social robots “are in real environments with real people interacting with them,” so they provide opportunities for observing social robots in the midst of our ordinary lives; just as for Breazeal having this technology “in real environments” is “tremendous,” so it is for us.<sup>58</sup>

### 3. The Achieved Quality of Bodies and Environments in DVA Interactions

In turning to everyday encounters with DVAs, we follow Natalie Jeremijenko's evocation of "peripheral participation" *in the technological context*, and her explanation that "rather than focusing on the interaction between the device and the 'user,' we pay attention to peripheral participation, the participation between users and surrounding things."<sup>59</sup> We further broaden this orientation by turning to the *situatedness* of DVA interactions, and ask: what is the occasion, what else is going on, and who else participates?<sup>60</sup> In contrast to the IoT imaginaries of the future, we indicate how the involvement of others articulates settings of technological use as thoroughly dynamic. To make this move, we recall Goffman's discussion of *face-work*.<sup>61</sup> Our account, however, orients Goffman's framework toward the phenomenon of talking gadgets, and goes beyond it in that it traces the interactional production of DVAs' expressive bodies, by which we resist anchoring the account in intentionality. The present section, thus, uses face-work to both indicate the dynamic aspects of DVA/loT settings, and—by contrasting it—to foreground the *open* quality of voice.

As part of our class assignments, students were asked to turn in five audio files of their own interactions with DVAs, together with field notes that describe, for each audio file, the situation, the spatial set-up, and the interactional flow. Somewhat surprisingly to the instructor, students' accounts associated the use of DVAs with private settings, where nobody else was present. In her field notes, for example, one student, Amber, expresses a preference for interacting with Alexa (embedded in Amber's Amazon Dot Echo device) in the privacy of her bedroom: "I . . . chose to do the conversation alone in my room because it's a private setting and I didn't want to talk to Alexa around people." On this occasion, Amber (who stores the device under her bed and uses it infrequently) reports on asking Alexa for a weather forecast and what would be an appropriate outfit for the weather:

It was kind of a "professional" feeling conversation with no emotions attached. It felt like I was living in a modern world where AI is always in your house and tells you the weather and helps you get dressed, for example. I could see smart houses in the future being like this.

While we took this expressed need for solitude seriously, we also read it through students' reporting on situational details and their collected interactional material. As students attempted to report on what was going on—describing the activity DVA interactions were part of, and spelling

out *what happened before, then, and after that*<sup>62</sup>—their notes suggest a *dynamism* of the setting, where the seclusion is an *achieved* quality that concerns (even absent) others.

Here is how another student, Flora, describes asking Alexa for a weather forecast in the morning:

I asked Alexa if it would be rainy today. This time I forgot to call her name “Alexa” before asking for questions, so she did not reply me again. I had to repeat the question. My roommate was disturbed by my loud voice and the sound of Alexa and asked me to keep quiet since she was still sleeping. Therefore, I did not talk to Alexa any more this morning. I realized that it is better to talk to smart speaker in a totally private room since talking to a machine will be a little bit embarrassing and disturbing.

Because Flora is a relatively recent adopter of smart speakers, her reported engagement readily manifests a gradually attained quality of the link between the technology and setting, where the setting itself is reflexively oriented toward events taking place within it. Flora has *learned* that she shall use Alexa only when alone; because Flora does not want to “disturb” her roommate, who asked her to “keep quiet,” she “realized” that “it is better to talk to smart speaker in a totally private room.”

That the setting for the voiced technology is affected by the social presence (or absence) of others contrasts with the advertised image of the voice-directed IoT. One frequently featured room in these imaginings is the kitchen. The kitchen, in Western symbolism (and at least since the open-kitchen format of the 1980s), is conceived as a space for communion—for being together with family and friends. That advertised smart home scenarios foreground the kitchen (and a deserted kitchen at that) suggests that social presence in those spaces is conceived as unproblematic. Flora, on the other hand, defines the DVA setting in respect to others—“a totally private room”—where the relationship of the technology to the setting is not only about technological efficiency, but concerns how she relates to others. Highlighting this tension between the tech industry’s portrayals of voiced IoT and the students’ ideas on DVA use is, however, not to ask for the design of DVAs to be directed toward private spaces. It is, instead, to emphasize that intervening in this domain cannot be accomplished in an exclusively top-down manner. Because of the lived quality of these settings, overlooking intricacies of ordinary interactions with talking devices risks obscuring the very character of this technology: its incompleteness.

Amber, who comments that “it is honestly just a little too weird [to interact with a DVA] in a public space,”<sup>63</sup> reports on an occasion in which she

was “getting more and more annoyed and frustrated” when Alexa was not obeying her command to stop playing a song:

The setting was me alone in my room, but one of my roommates walked by and heard me yelling and started laughing out loud at me after I finished the recording. She asked me what I was doing and thought it was funny. I was embarrassed and did not want her to think I was doing this for no reason, so I told her it was for a project. She told me it sounded like I was yelling at a real person and was confused why I had to talk to Alexa for a project . . . A little later, my roommate played music from Alexa while she was putting on makeup and I could actually hear her yelling “Alexa, skip this song!” multiple times also, so I realized it is common for this to happen. I think most people assume they would not get angry or argue with a bot since it is not real, but in the moment it is very easy for you to forget you are talking to a machine and that it does not understand your tone anyway.

As the interaction between Amber and her roommate suggests, being alone with a DVA is still about relating to others. The two roommates interact with their DVAs “on their own,” but that interaction is not immune to what others do when they overhear it (be it as a way to enforce a particular behavior on others or to normalize their own).

Relationships with others and being alone in the context of a DVA’s engagement appear to be about emotions of shame and embarrassment, as indicated by both Flora and Amber’s field notes (where the term “embarrassment” reoccurs). Flora explains that one shall not embarrass oneself by not being considerate toward present others, disturbing and invading their privacy, and Amber suggests that one shall not embarrass oneself by appearing naive or deluded by believing in the sociality of voiced technologies. As both students want to present themselves in a favorable light in front of others, their comments echo Goffman’s writings on how we aim to render ourselves in the public. Goffman defines “face” as “the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact. Face is an *image of self* delineated in terms of approved social attributes.”<sup>64</sup>

Turning to Goffman’s face-work in the context of DVAs allows us to highlight the social aspects of the setting: how students talk about embarrassment indicates that the environment associated with this technology is not only about appliances into which chips are inserted, but also about who else participates (as present or not), and how that relationship is managed. Of interest, however, is to inquire how that framework, accompanied with Goffman’s rather narrow account of the material world, may work in the domain of speaking machines. When thinking about face-work in terms of “staging” situations,<sup>65</sup> Goffman confines the material world

to “equipment” and “tools” that (together with an “audience” and the “team”) are involved in the “maintenance of the front stage.” In our examples, however, students describe a human–machine setup where they are interactionally oriented toward machinic *voices* as well. In other words, students manage their relationship with their roommates in the context of talking with their DVAs. Flora disturbs her roommate by repeating commands to the smart speaker so that the machine can respond back, and Amber is heard by her roommate because of her unsuccessful attempts to direct Alexa to play music. While Amber does this “for a project,” she finds similarities to how her roommate behaves when she overhears her interacting with Alexa. This orientation to DVAs goes beyond equipment and tools that aid humans in their accomplishment of face-work, while its voiced character not only questions the absolute power of the intentional mind behind technology, but also behind its human interlocutors. In other words, we are confronted with a *bi-directional conditioning* that an orientation toward psychology and the individual, as exhibited in Goffman’s face-work<sup>66</sup> and his analysis “devoid of the actual action,”<sup>67</sup> cannot account for.

The discussed example of Flora asking her DVA for a weather forecast in the morning indicates how its interlocutors need to align their bodies with the technology for its voice to speak: Flora repeats her commands, talks in a loud voice, and addresses the machine by its name.<sup>68</sup> In her field notes, Flora includes a further comment from her boyfriend, illustrating a member’s recognition of this interactional production of Alexa’s voice:

My boyfriend thought it was funny when I repeated the question again and again to a machine. He told me that when I talked to Alexa, I spoke in a cool and stiff tone, just like a repeater, and it sounded like two robots were communicating. I asked my boyfriend to ask Alexa some questions about football. He refused since he thought talking to a robot made him awkward. [This is in reference to a clip Flora entitled “Football Match.”]

In saying that Flora’s voice is “robot-like” and that her conversing with the DVA sounds like “two robots were communicating,” the boyfriend highlights the alignment of Flora’s voice to the robot’s voice. In modifying her voice so that the robot can talk back, Flora’s voice projects characteristics of a robot’s voice as it is *articulated by it*. This mutual articulation between Flora’s voice and the voice of the machine—while undoubtedly related to the design and manufacturing process—also goes beyond them, together with the work that users put into training neural nets to support their speaking systems.

As the boyfriend identifies this bi-directional conditioning in Flora’s talk, he also mentions his concern about feeling awkward at the prospect of

engaging with the DVA himself. This feeling of awkwardness evokes, once again, face-work, paralleling the previous examples concerning solitude. When students comment about embarrassment or awkwardness, they are concerned with themselves as individuals, carving a neat demarcation between them and the rest of the world. However, that we are in a presence of an *open* self becomes inevitably apparent in actual courses of interaction, such as the ones to which we turn in the next two sections. There, in seemingly unremarkable moments of everyday encounters with talking gadgets, we witness an articulation of human-machine assemblages, as a part of interactional scenes they derive from. Paying attention to these lived aspects of technological encounter indicates how as the voice of the machine is spread out, so are its interlocutors' images of self.

However, we do not, with this, dismiss the problem of self a priori. We recognize that the idea of the self (with its corresponding voice) is important to designers and AI critics, and we acknowledge that it is evoked by DVAs' interlocutors (when they account for their technological engagements). So we pay it attention, responding to how social AI and its implications are imagined and discursively accounted for. However, we consider this topic of concern in our focus on participants' interactional conduct. While being ethnomethodologically bound to taking the members' perspective seriously,<sup>69</sup> our task is to describe how the self is rendered in human-machine encounters, with all the concreteness that the consideration of the setting calls for.

We do so by engaging two students' examples, Max's and Dave's, where we focus on their audio recordings of DVA interactions and the *conversation analysis* (CA)-style transcripts<sup>70</sup> of those recordings, coordinating these with students' field notes, and considering them in relation to what others reported. In Max's example, below, we trace how the voice of the machine is articulated through the coordinated interaction of two housemates, and how that interaction manifests a dynamic character of the self that the housemates exhibit. When we turn to Dave's example (in section 5), we further inquire into the enactment of humans as DVAs' interlocutors, specifically focusing on how their intentional states are *distributed* across the human-machine divide in singular situations of interaction. As we follow the dynamics of interaction in the reported examples, we move away from face-work, instead engaging Goffman's thinking on *participation*.<sup>71</sup> We conclude the article by speculating (in section 6) on how the *openness* we notice may be aligned with Simondon's ideas on technical objects' integration within a world,<sup>72</sup> and by suggesting (in section 7) that resisting the image of DVAs' "convenience" can function as an engaged act.

#### 4. Producing Alexa through Coordinated Interaction

Max, like his classmates, reports on his concerns about using his DVA when not alone. After asking Alexa for a weather forecast, he comments, “Since we just spoke of the weather, I didn’t mind if my housemates overheard, thinking I’m weird for speaking to a robot.” Max’s comment may be seen as implying that he avoided disapproval (of his housemates) by using the DVA as a tool for problem-solving (i.e., learning about weather conditions). But Max also emphasizes the machine’s employment beyond this tool-like treatment, which he associates with its location—his bedroom—and how he engages with it:

If for say, the Alexa was located in my living room it would give off a completely different vibe. It wouldn’t be seen as a social being and more as a robot to ask only beneficial questions, like “Alexa what is the weather today?” My housemates would not treat the Alexa the same way I do . . . In some way, it seems like they fear the robot, which is why another environment would take away its social characteristics from it. It is only given these social attributes because of how I treat it and ask it questions . . . I treat Alexa as my friend, so when my housemates come into my room, they treat her like that as well.

Max hypothesizes that if used in common spaces (rather than in his bedroom, and when he is present), his housemates would address the DVA only to ask “beneficial questions.” Their using of the DVA in a tool-like manner would thus align with the image of the smart home of the future that Amber envisaged. Reporting, however, on how an actual interaction in his bedroom took place, Max employs descriptors such as “heated,” “surprised,” and “jumping up and down,” rather than portraying the interaction as “professional and emotionless” (the words Amber uses). During the event in which such feelings manifested, Max speaks with Alexa in the presence of his housemate (whom, here, we will call “the housemate”); as is, according to Max, commonly the case, “we usually speak to her together.”<sup>73</sup> In Max’s field notes, he frames the event as follows:

My housemate and I were playing video games together in my room late at night because we were both done with all our work for the week. While we play video games we usually lower the volume on the game and play music through the Soundcloud app on the Xbox to play music in the background. We got in an argument about which artist’s playlist to play. I said Kendrick Lamar, while my housemate said Drake. We decided to ask Alexa and see her response to help us come to a resolution . . . At this time we are both very heated as we had been debating for the past 10 minutes. We were both bringing up random facts and reasons as to why one is better than the other.

Max explains that what happens is not out of the ordinary:

This interaction is a very common one between Alexa, my housemate, and I. Whenever we get into an argument we immediately go to my bedroom to speak Alexa to sort it out, as she is probably the most unbiased source we have around us.<sup>74</sup>

Max also points out that the conversation was exciting because the DVA, when asked who is the best rapper alive, provided an answer that the two did not foresee:

To our surprise she told us “Chance The Rapper.” There are five other rappers who most people would say are better than Chance, so I thought it was quite amazing Alexa was programmed to say him. I think very highly of Chance while other people don’t. We also asked her the best female rapper which she said was Missy Elliot, who is not popular at this moment, but still very good . . . By our reactions you can tell we are overjoyed that she did not give us a basic answer. After our conversation, we decided to play an old Chance The Rapper album in the background while we played our game and our love for him and his music sparked again.

The below excerpt displays the transcript of the interaction between the housemate (H), Max (M), and Alexa (A).<sup>75</sup>

- 1 H: Aye Alexa, (0.1) who do you think is the  
best rapper alive?
- 2 M: Hahaha  
(0.2)
- 3 A: This might answer your question, Best  
Ra[pp]er Aliv[er]e by Lil Wayne=  
4 M: [Kendrick.]
- 5 A: =is December 6th 2005 song whose album is  
the Carter 2 by Lil Wayne and whose album  
is the [Carter]r 2 by Lil Wayne=  
6 M: [Hihhh.]
- 7 H: [Bro, Lil Wayne?]
- 8 A: =Lit was written [ by D Carter
- 9 H: Lil [Wayne is not,]
- 10 A: [Did that answe]er your question?

- 11 H: No, ɾLil Wayne is nɔ:t the best rapper  
alive.
- 12 M: ɫAsk it again. ɫ
- 13 H: Who do you think is the best rapper alive?  
(.)
- 14 A: Hmm, I'm not sure.  
(.)
- 15 M: WHO IS THE BEST RAPPER ALIVE.  
(0.2)
- 16 A: Well, in my opinion, it's Chance The  
Rapper.
- 17 M: Ooɾhhhhhhhhhuuhhuuuɾ=
- 18 H: ɫChance The Rapper?ɫ
- 19 M: =I FEEL ɾyou, Alexa!ɾ
- 20 H: ɫAlexa, whatɫ the hell? huh
- 21 M: No, that's lit!=
- 22 H: ɾHahahaɾ
- 23 M: =ɫThat'sɫ lit!
- 24 H: What the hell?
- 25 M: That's actually hella lit.
- 26 M: Who's the- Alexa, (0.1) who's the best girl  
rapper?  
(0.3)
- 27 A: Hmɾm, I amɾ not sure.
- 28 M: ɫCardi Bɫ
- 28 M: Ah, what?
- 29 H: Who in your opinion is the best girl  
rapper?  
(0.3)
- 30 A: Hmm, I ɾam not sure.ɾ
- 31 M: ɫSay female, ɫ

32 H: Who in your opinion is the best female rapper alive?

(0.3)

33 A: I like Missy Elliot.

34 M: Okay, (.) she's like bom bom ay=

35 H: Yeah.

36 M: =ba ba bom bom.

The housemates engaged the DVA as an arbiter who can resolve their disagreement in an “unbiased” way and by “knowing what would be best.” Other students reported on similar instances where DVAs participate in practices that bring to mind *cleromancy*. As in other forms of cleromancy, those DVA interactions suggest a sense of seeking an impartial, but often playful, outcome, while also treating the received answers as if they were revealing a will behind it. In the case of DVAs, however—rather than by divination (as is usually the case in other forms of cleromancy)—the outcomes delivered through the DVA’s voice indicate the presence of the machine’s self. The details of the interaction that takes place in Max’s room suggest, however, that Alexa’s voice (and through it, the social self of the machine) involves production that actively engages the two housemates. In other words, the machine’s self cannot be separated from the scene where the two housemates work out their disagreement. In that sense, the interaction not only provides the two housemates with a truth (while amusing them), but indicates how the self of the machine, which *speaks* that truth, is open. What is more, as the two participate in making it possible for the machine to tell the truth, their involvement in that production questions their previously taken positions. In the presence of the voiced machine, the housemate—whom Max describes as having “in some way” a “fear of the machine”—is now not only involved, but *coordinates* his action with those of Max, in spite of their divergent agendas. In that sense, tracing the dynamics of the human–machine engagement allows us to spot how the selves of the two housemates—just like the self of the machine—is oriented toward contingencies of the interactional scene.

The transcript displays a series of moves where the housemate—despite being depicted as someone who shuns the machine—indicates his interest in Alexa’s opinion, “interacting with it as if it were a person, and ultimately as a friend.”<sup>76</sup> In Line 1, the housemate stresses “you” when asking, “who do you think is the best rapper alive.” In Line 13, he tries to remedy Alexa’s providing of information about a song (rather than about the rapper) by repeating that same question and expressing an interest in

what Alexa *herself* thinks.<sup>77</sup> The housemate also indicates that curiosity when he inquires about the best female rapper (originally asked by Max more broadly in Line 26) with the phrasing, “Who in your opinion is,” (Lines 29, 32).

Addressing the DVA as an agent is also at play when the housemate and Max contest that Lil Wayne is the best rapper, acting as if that was the answer Alexa provided (Lines 7–11). They know that Alexa reported on the name of a song (giving a “basic factual answer,” similar to what Google does, according to Max’s description). As Max’s field notes state: “When we asked Alexa who the best rapper alive is the first time she told us about the song ‘Best Rapper Alive.’ She perceived the question the wrong way, quite humorously.”<sup>78</sup> At first, they laugh while the DVA gives information on the song (Line 6). Immediately thereafter, however, they realign their actions, treating the answer as if it was in response to what they asked. They do so by showing their surprise and then by contestation, treating the information about the song as if it were a statement on the best rapper: “Bro, Lil Wayne? (Line 7); “Lil Wayne is not,” (Line 9); and “No, Lil Wayne is no:t the best rapper alive.” (Line 11).

As the two address Alexa as an agent, emphasizing that they want to know what *she thinks*, they importantly enact a self behind the machine. As their questions are not only directed at an acquisition of information (but are aimed at learning about an opinion that is to stem from an agent), they participate in an interactional production of the machinic self that is to exist behind its voice. At the same time, that very involvement in the interactional production distributes the machinic self, possibly even dissipating it in interaction. This is available in how the two treat Alexa’s “denotative meaning”<sup>79</sup> as an answer about the musician (while being aware that the machine did not answer their question), and in how they provide the DVA with the status of an interlocutor by designing their own utterances so that the device can respond. Like Flora who modifies her voice in employing prescribed strategies to equip the machine with *its* body,<sup>80</sup> in Line 15, Max raises his voice in his orientation toward the machinic functioning. Similarly, the two housemates repeat and rephrase their questions several times—Lines 13, 15, 29, 32—in addition to (and again, like Flora) repeatedly addressing the DVA by its name in ways that differ from human-to-human conversational flow. Further, they time and repair the delivery of their speech to allow the device to function according to its design. An example is in Line 26, where Max initiates his move with the question about “the best girl rapper,” but then immediately repairs it, opening his new attempt with “Alexa,” followed by a pause, before he asks his question.

One could hypothesize that on this occasion, Max's reputation grows in front of his housemate. When Max, for example, indicates his relationship with the competent machine, he could be seen as not only someone who is acquainted with rap music, but someone who is in the know about worthwhile technologies. When Max approves of Alexa's answers through excited vocalization (e.g., in Line 17), and expresses alignment with the device when he says "I FEEL you, Alexa!" (in Line 19), or when he provides his answers concurrent to the DVA's talk (in Lines 4 and 28), or argues that Alexa's answer is "lit" (Lines 21, 23, and 25), his moves potentially tap into the uneasiness he expressed when commenting about his housemates' stance toward his DVA. In that sense, this making sure that his housemates don't think he is "weird for speaking to a robot" (beyond asking "only beneficial questions") boosts Max's social position (making him gain approved social attributes, while claiming a positive social value for himself through this speaking).

Despite this, it may well be argued that ideas concerning self—together with those of *impression management*, *staging*, and *identity*<sup>81</sup>—do not fully capture what goes on in this interaction. The excerpt, rather, indicates the situated engagement where the actions of the two housemates not only regard their positioning toward each other, but also concern their involvement in producing Alexa's voice. That involvement, in turn, features the two quite differently from how Max originally describes their agendas. This is detectable in how the housemate shows his interest in the robot's opinions (against the original definition of disliking the machine), but becomes most noticeable in the *collaborative* character of the robot's interactional production. While the event continues to be importantly an affair between the two housemates—and an antagonistic one at that (a winner still needs to emerge)—we also witness a shared effort put into making the DVA work.<sup>82</sup> When Max turns the housemate's "who do you think is" (Line 13) into "who is" (Line 15), and suggests (in Line 31) substituting "girl" (the term that appears in the housemate's question in Line 29) with "female," this repairing of the competitor's commands makes the two housemates—who are "in an argument," having been "debating for the past 10 minutes"—collaborators. As the two are involved in making the machine speak,<sup>83</sup> what goes on in Max's room suggests a weakening (if not a full disappearance) of the image of the self as something that is wholly located behind the robot's voice, while it indicates a shifting of that image in what concerns the two housemates.

## 5. Distributed Self in a Conjoint Involvement in Talk across Human–Machine Divide

We now turn to an example of DVA interaction collected by another student, Dave, which we treat as an occasion for further pressing our noticing toward DVAs' interlocutors and their selves in interaction. In the previous section, Max's example indicated how the machinic self is enacted and how it dissipates through interactional production, while also giving us a chance to render how the positioning of the DVA's two interlocutors dynamically shifts through that interaction. As we continue to describe how the DVA's voice is interactionally produced we specifically focus on the *distribution* of the DVA's interlocutor, indicating how *intentional states* of a DVA's interlocutor are extended and distributed. While our argument relies on the *extended mind hypothesis* and *distributed cognition*, we shift attention from an explanation in terms of *tasks* and *tools* to a focus on conjoint involvement in talk across the human–machine divide. As we describe what takes place in the setting dynamically conceived, and pay attention to how the social body of the machine is produced through the course of interaction, we trace the DVA's interlocutor's "wish"—and the social self to which it is attributed—to the openness-with-the-world.<sup>84</sup>

On the occasion of Dave's reporting, the DVA is not involved in music as the topic of conversation (as it was the case in Max's example), but as a smart speaker used for listening to music. The smart speaker, whose owner is Dave's best friend, functions as the "centerpiece of the social space." As we focus on this space, we call the owner—who, in Dave's words, is "a person who likes to be up to date on the newest technology," and is "an experienced conversationalist with Alexa"—"the friend." The porch of his house is where the group of friends listens to music (see figure 1, which Dave submitted with his assignment to illustrate this set-up):

We often convene at this location because we like to sit around and converse while listening to music, which we do through [the friend's] Amazon Echo equipped with Alexa voice control . . .

The voice assistant has become a central piece of social gatherings. Alexa is used as a device for us to discuss and play music, playback information about weather and movie times, and even to settle disputes. It settles disputes by answering questions that are being discussed in conversation and performs the task of fact-checking . . .

Without the need for an extensive design of the human body or other artifact that could resemble a being, through the simple and powerful use of voice Alexa is able to infiltrate the space that we often only designate to other humans.



FIGURE 1. The porch of the friend's house with the Alexa smart speaker.

As Dave highlights the appeal of the machinic voice and its shaping of the social space, he reports on an interaction where the friend (F) asks Alexa (A) to play “hip hop instrumentals.”

- 1 F: Alexa,  
 2 A: ((Sound emitted from Amazon Echo))  
 3 F: Play hip hop instrumental on YouTube.  
 4 A: ((Sound emitted from Echo)) I Couldn't find any enabled video skills that could play that. Go to the Alexa app to manage your video skills and devices.  
 5 F: Alexa,  
 7 A: ((Sound emitted from Echo))  
 8 F: Play sixties music.  
 9 A: ((Sound emitted from Echo)) Here is a station for sixties music. Sixties rock on Amazon Music. Amazon music is streaming on

another device. Would you like to stream from here instead?

10 F: Yes.

11 A: ((Sound emitted from Echo))(5.0)((Music begins to play))

Dave describes the interaction as follows:

In this file [the friend] asks Alexa to play hiphop instrumentals through Youtube for us to listen to while we are outside enjoying the weather. When Alexa responds that her current capabilities do not allow her to complete the action that [the friend] asked her to do, [the friend] quickly changed his mind to ask it to play music from a moment in time. Most likely he knew that Alexa would be able to find a way to play something general from the sixties. [The friend's] demeanor changed because he didn't want there to be empty silence, his brother and I remained quiet in order to allow him to finish talking with Alexa in order to not interrupt his command . . .

Another quick thing to note is the fact that Alexa has the capability of playing the genre of music that he was looking for, but the fact that [the friend] mentioned Youtube caused Alexa to give him directions on how to sort that out. Directions on how to manage apps outside of Amazon's network took precedence over [the friend's] request to listen to a genre of music, which then made him change his mind completely about the type of music he thought Alexa could play through voice commands. This was a very interesting exchange because of the quick change of direction.

What attracts Dave's attention is "the quick change of direction" in the friend's request for music: the friend "quickly changed his mind" and his "demeanor changed." After the friend has asked Alexa to "play hiphop instrumentals on Youtube" in Line 3, in Line 4, Alexa tries to get him to add additional (third-party) apps, possibly broadening the machinic reach into his behavior.<sup>85</sup> However, in Lines 5 and 8, the friend turns away from this possible further information capture, and asks for "sixties music." Dave explains that this orientation has to do with the friend's not wanting "there to be empty silence." The two aspects Dave highlights—the threat of empty silence and the quick change of direction—concern, once again, the setting and the machine's body as organizing notions in DVA technology.

During the interaction, Dave and the friend's brother—whom we call "the brother"—are present. This arrangement suggests an articulated "participation framework,"<sup>86</sup> with two *non-ratified* participants involved.<sup>87</sup> In that sense, the setting, on this occasion, is not only about the furniture, the appliance, and their spatial organization, as Dave's photo in figure 1 may suggest. It also involves Dave and the brother, whose participation exhibits the dynamicity of that setting (in the social situation of listening

to music with friends). Regarding his and the brother's participation, Dave explains:

It is not uncommon for simultaneous conversations to take place in that social space, so it has now become customary for people to hold a pause in their conversations when the voice assistant is addressed by [the friend] . . . We have given precedence to the exchanges between human interactant and voice assistant when they need to take place above our human to human interactions.

That the group members “hold a pause in their conversations when the voice assistant is addressed by the friend” has probably to do with a concern that concurrent conversations would cause the machine to malfunction. Nevertheless, it also provides a possibility for the non-ratified participants to intervene in the friend's interaction with the DVA. In the set-up described by Dave, the friend can perceive Dave and the brother's access to his DVA interaction, and while they are, at this time, participating as *overhearers* (“they can temporarily follow the talk, catch bits and pieces of it, all without much effort”<sup>88</sup>), they could possibly *change their footing*,<sup>89</sup> entering the speaker role. If the friend's conversation with the machine does not maintain a certain flow of talk, and an “empty silence” occurs,<sup>90</sup> opportunities may open up for the two *bystanders* to insert themselves into that interaction, and, by taking the floor, talk across the boundaries of the dominant conversation.<sup>91</sup> This is to say that the problem is primarily an interactional one, not to be explained away with a psychological self and individual's interiority. One can, certainly, suppose that if the friend was to insist on his original request to play hip-hop music, that could (by foregrounding the inadequacy of the technology) make somebody who “enjoys being up to date on the newest technology” lose face. We know from other students' field notes that this is a concern for advanced users of DVAs when their friends are around (particularly if those friends are non-adopters and may be skeptical about the technology). Dave, however, does not write about such concerns, but provides an interactional account, where the machine is positioned as a ratified participant (in the presence of two human overhearers), and where the flow of talk is paramount.

As Dave mentions, the group members suspect that “Alexa has the capability of playing the genre of music that [the friend] is looking for” (Dave says this is a “fact”); the question, however, is how to coordinate with Alexa in a way that is also responsive to what is happening on the porch. The friend needs to enact the robot's competence by occupying that fraction of time in the DVA conversation before the other two can notice the silence or even take the floor (possibly inserting their demands into the

selection of music the group is to listen to). To do that, the friend *changes his mind*. He engages in the production of the machine's voice by now *wishing* for what, in that very moment, the AI is able to execute. We can say that as the friend (by no longer *wanting* to hear a "hip-hop instrumental" but "sixties music") provides the body to the machine (making it capable of responding without detours that could generate the unwanted silence), he *distributes himself* across the world. His intentional state—the *wish* to hear certain music genre—is relative to his linkages with the DVA.

We can thus describe Dave's example as a case of *distributed cognition*:<sup>92</sup> the friend's wish is not simply an internal, brain-bound state; instead, it emerges on the friend's porch, on that afternoon, through a coordination of humans and artifacts. In contrast to Michael Patrick Lynch's proposal (where Lynch calls for "getting up off the couch and plunging into the whirlpool of actual experience" to counter the extended mind hypothesis<sup>93</sup>), the example shows an act of distributed cognition as a very feature of "the whirlpool of experience." By noticing and then indicating it, we resist—from the world we already live in—the return to the individual as a way to deal with AI technologies.

However, in Dave's example, the distribution across the material world does not appear to be about *problem-solving* where the technology is a *tool* to get something accomplished. If it were, the friend would continue acting to realize his original request (as it would also be the case if these devices were primarily about their convenience). One can, certainly, suppose that the friend's DVA functions as a tool for Amazon in that it collects valuable information for the company while training its neural nets,<sup>94</sup> but considered from the perspective of an engagement at its interface—organized around the conversation-like paradigm—it is involved (by the group of friends, and in a playful manner) as a voice. While the request for sixties music does not really make the machine solve the task of playing hip-hop instrumentals (and, in that sense, the device, rather than a convenient tool, is a lacking instrument at best), it makes it converse in front of those who co-participate in the scene.<sup>95</sup> In other words, the distribution, here, concerns this conversation-like relationship with the robot, where the friend is thinking with the *speaking* world. What takes place on Dave's friend's porch is an alignment and co-production of selves, where, as the self of the machine is enacted through talk, so are aspects of the human interlocutor's self. While this does not allow for justifying our engagement with these machines by resorting to utilitarian concepts and schemata of efficacy, it shows our interactional bonds with their dynamic materiality: an openness-with-the-world.

## 6. When the Integration within a World Concerns the Speaking AI

In the last section of his book *The Presentation of Self in Everyday Life*, Goffman summarizes it by explaining:

In this report the performed self was seen as some kind of image, usually creditable, which the individual on stage and in character effectively attempts to induce others to hold in regard to him. While this image is entertained concerning the individual, so that a self is imputed to him, this self itself does not derive from its possessor, but from the whole scene of his action being generated by that attribute of local events which renders them interpretable by witnesses. A correctly staged and performed scene leads the audience to impute a self to a performed character, but this imputation—this self—is a product of a scene that comes off, and is not a cause of it. The self, then, as a performed character, is not an organic thing that has a specific location, whose fundamental fate is to be born, to mature, and to die; it is a dramatic effect arising diffusely from a scene that is presented, and the characteristic issue, the crucial concern, is whether it will be credited or discredited.<sup>96</sup>

This quotation seems to capture our tracing of the self as deriving (to an important extent, at least) from scenes where interactions with DVAs take place. Yet Goffman's account, by relating to self in terms of an "image" or a "dramatic effect," remains disembodied. While this appears linked to the intentionalistic take that characterizes Goffman's text up that point,<sup>97</sup> it is not helped by the exclusion of the material world (beyond treating it as tools and props in staging performances). Turning to the speaking AI meant anchoring our account in that world.

To reflect on the open self that we described, we may recall Simondon's discussion of the genesis of the relationship of man and the world, where Simondon suggests that to understand the technical thought, one shall consider its derivation from the original mode of being in the world—"the magical mode."<sup>98</sup> Simondon explains that when technical objects—considered as *instruments* or *tools*—relate to the world, they do so "only in a single point and at a single moment," unable to "recover the contact with the world in its unity, which was the goal of magical thought."<sup>99</sup> Yet Simondon suggests that *aesthetic reality*—which "appears at the neutral point, between techniques and religion, at the moment of the splitting of the primitive magical thought"—"cannot be said to be either properly object or properly subject":<sup>100</sup>

There is, of course, a relative objectivity to the elements of this reality; but aesthetic reality is not detached from man and from the world like a

technical object; it is neither tool nor instrument; it can stay attached to the world, for instance by being an intentional organization of a natural reality; it can also stay attached to man, by becoming a modulation of the voice, a turn of phrase, a way of dressing; it does not have this necessarily detachable character of the instrument; it can remain integrated, and normally it does stay integrated within human reality or the world; a statue is not placed just anywhere, a tree is not planted just anywhere . . . It is indeed this integration that defines the aesthetic object, and not imitation: a piece of music that imitates noise cannot become integrated into the world, because it replaces certain elements of the universe (for instance the noise of the sea) rather than completing them.<sup>101</sup>

In this context, Simondon talks about technical objects' "integration within a world" in regard to their "aesthetic value," "found within the action":<sup>102</sup>

A tractor, in a garage, is merely a technical object; however, when it is at work plowing, leaning into the furrow while the soil is turned over, it can be perceived as beautiful. Any technical object, mobile or fixed, can have its aesthetic epiphany, insofar as it extends the world and becomes integrated into it. But it is not only the technical object that is beautiful: it is the singular point of the world that the technical object concretizes. It is not only the line of pylons that is beautiful. It is the coupling of the lines, the rocks, and the valley, it is the tension and flexion of the cables: herein resides a mute, silent and ever continued operation of technicity applying itself to the world.<sup>103</sup>

Talking machines suggest a somewhat parallel case, where the integration of the technical object "within a world, whether it be geographical or human"<sup>104</sup> concerns that "*unio mystico*, a socialized trance"<sup>105</sup> that is characteristic of interaction. In such a context, the material world is engaged not only as tools, instruments and objects, but as *selves and individuals achieved in interaction*. Yet while our examples have a certain resonance with Simondon's *reversible conditioning between man and the world*—where "the world influences man just as man influences the world"<sup>106</sup>—we do not see DVAs as having an aesthetic value. Instead, we find a potential for beauty in noticing the *interactional openness* that DVA engagement provides.

In that sense, our examples may parallel some aspects of Yuriko Saito's "everyday aesthetics,"<sup>107</sup> where Saito indicates the importance of going beyond "dramatic, sublime, or picturesque landscapes,"<sup>108</sup> to embrace "unscenic nature" as a part of our green aesthetics.<sup>109</sup> In contrast to Simondon—who describes integration in "primitive magical thought" as structured around "privileged places and moments," while organized in "a network of key-points, of high-places"<sup>110</sup>—our account of interactional

integration moves the discussion closer to “seemingly trivial, innocuous, ordinary, mundane, or even frivolous aspects” that Saito celebrates.<sup>111</sup> We thus traced our *reversible conditioning*—where human and machinic *interlocutors* extend the world as they become integrated in it—through inconspicuous moments of the everyday, in which, rather than identifying how the openness sends us back to privileged spaces of original realms, we pointed out how voices are achieved *here* and *now*, in lived spaces we inhabit with speaking others.

## 7. Attending and Noticing as Intervening

Back in the 1980s, Thomas Disch wrote a children’s book entitled *The Brave Little Toaster*,<sup>112</sup> which was, during that time period, adapted as an animated film of the same title. One of the students in our class reported frequently watching *The Brave Little Toaster* on home video as a child, and described his consequent difficulties in parting from older appliances and other household objects deemed obsolete as these acts of discarding brought to mind the speaking appliances the movie portrays. We wish we could suggest that today, as we live the moment when it is more frequent that voices emanate from things across our environments—and social roboticists acknowledge this move as a march of social robots from the laboratory into the big wide world—the presence of those voices may change our views on obsolescence and what we define as junk. We would like to say that DVAs—while not much more than an embodiment of human narcissism (where everything needs to resemble us)—could, somehow, also serve as a way to curb our unhinged consumerism. But we are afraid we cannot; we have not witnessed anything like that happening.

This said, we do believe that arguing for attentiveness to ordinary encounters with AI can function as an engaged act. While our proposal does not concern an actively intended intervention in the manner that characterizes such phenomena as “emancipatory resistance and activists’ use of technology,”<sup>113</sup> or efforts toward “digital detox and decluttering,”<sup>114</sup> being attentive to what we do and how we do it provides possibilities for decelerating the processes we may not want, as it opens up possibilities for reframing our relationship with AI (whether to engage with technology differently, how to reduce our commitment toward it, and so on).

One example is the rationalization of our massive use of the voiced technologies in terms of “efficiency” and “convenience.” When the question of “why talk to a toaster” was originally brought up in class, students’ first answers were in those terms. In their written comments on the *New York*

*Times* article, they similarly explained away their DVAs' use by bringing up convenience. On that topic, Max explains:

People could suggest it makes you lazy, but I think it is just more time-efficient getting through your day with [Alexa], rather than constantly searching on your phone for simple answers. Because usually when you go on your phone for a simple question, you get distracted and end up using it even longer than intended.

The idea of convenience is also how the *New York Times* article alludes to our employment of DVAs. The article points out that, in addition to there being revenue, surveillance, and marketing opportunities (“Sometimes it will lead to revenue opportunities—Amazon’s microwave will reorder popcorn for you when you’re running low. Sometimes smarts are used for surveillance and marketing, like the crop of smart TVs that track what you watch for serving up ads”), there is the convenience of interacting with the technology through voice: “Sometimes smarts will lead to convenience—you can yell at your microwave to reheat your lunch across the room.” A similar appeal to convenience can be found in Kate Crawford’s and Vladan Joler’s powerful 2018 comment on the materiality of the infrastructure that lies *beneath* each Amazon Echo device, with its “layers of fractal supply chains, and exploitation of human and natural resources, concentrations of corporate and geopolitical power, and continual energy consumption.”<sup>115</sup> That Crawford and Joler contrast these globally dispersed processes of production and exploitation to Amazon Echo’s advertisement (rather than considering the lived aspects of that engagement) may be the reason for their characterization of that engagement in terms of convenience. In its noticing of interacting voices in our kitchens, our account, by marked contrast, attends to paths through which that convenience is achieved, and through which a talking gadget *becomes convenient*.

Accordingly, we claim that noticing how we sustain AI technologies and how we engage in producing effects of convenience or efficiency has a potential to be transformative, as it does not relegate troubles to other spaces and moments, but directs us toward the lived now in which we participate. As this opens possibilities for resisting what we do not like about AI, it concurrently provides an occasion for inquiring into ideas of self, the body, and the fixity of environments we inhabit. In practicing an inquiry into AI that is not exclusively done at the level of social structures, we are confronted with distributions and linkages across human and technological realms, catching openness-with-the-world in unremarkable moments of interaction.

## Acknowledgments

We are grateful to all the students who, over the last decade, took part in the Communication and Social Machines class at the University of California, San Diego. We would also like to thank Antonio Barrios, Maria Giulia Dondero, Allyson Huang, Jamie Lam, Michael Lynch, Maurizio Marchetti, David Pedersen, the *Evental Aesthetics* editors Juljan Krause and James Camien McGuiggan, and two anonymous reviewers; as well as those who hosted the lead author and provided their comments when she presented this work at the University of Southern Denmark, University of Liège, University of California, San Diego (Science Studies Colloquium), and at the 2019 Meeting of the Society for the Study of Symbolic Interaction in New York City.

## Notes

- 1 Even mass-media outlets are taking up these discussions: see Kevin Lincoln, “Where is the Boundary between Your Phone and Your Mind?,” *The Guardian*, December 9, 2018, <https://www.theguardian.com/us-news/2018/dec/09/tech-mind-body-boundary-facebook-google>.
- 2 Bruce Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World* (New York: W. W. Norton & Company, 2018).
- 3 See also, for example, Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York, NY: Public Affairs, 2019).
- 4 Popular media massively reports on the rise of these technologies. For example, *Business Insider* reports that DVA devices and voice-controlled smart home gadgets were the highest selling electronic devices for holiday seasons of 2018 (Dennis Green, “Amazon Reveals Its Best-Selling Gadgets of the Holiday Season as Smart Home Tech Takes Over Gift-Giving,” *Business Insider*, December 27, 2018, <https://www.businessinsider.com/amazon-best-selling-gadgets-holiday-season-2018-12>); see also Maxine Builder, “The Gizmos to Buy Your Techie This Holiday (Before They Sell Out),” *New York Magazine*, October 18, 2018, <http://nymag.com/strategist/article/best-tech-gifts-holiday-trends.html>; Dieter Bohn, “Amazon Says 100 Million Alexa Devices Have Been Sold—What’s Next? Amazon Isn’t Letting Google Own CES without a Fight,” *The Verge*, January 4, 2019, <https://www.theverge.com/2019/1/4/18168565/amazon-alexa-devices-how-many-sold-number-100-million-dave-limp>). “The Smart Audio Report,” conducted by NPR and Edison Research, reports that 21% of U.S. adults, or around 53 million people, own a smart speaker; that the number of smart speakers in U.S. households grew by 78% in one year from December 2017 to December 2018, with 14 million new smart speaker owners in 2018 (“The Smart Audio Report, Winter 2018” National Public Media, Accessed June 26, 2019, <https://www.nationalpublicmedia.com/wp-content/uploads/2019/01/Smart-Audio-Report-Winter-2018.pdf>). An article which heavily relies on the NPR report, explains: “For context, smartphone shipments shrank 6 percent globally in the third quarter of 2018, compared with a year earlier, due to smartphone saturation in the US and other markets. The next great consumer electronics wave is being led by smart speakers and their virtual assistants” (Raini Molla, “The Holidays Ushered in a Smart Speaker Explosion Starring Amazon’s Alexa and Google’s Assistant,” *Recode*, January 8, 2019, <https://www.recode.net/2019/1/8/18173696/amazon-alexa-google-assistant-smart-speaker-sales-npr>).
- 5 This is also different from voices in the context of radio and television, for example, where they do not exhibit the incompleteness of an interactional or conversational voice modeled in DVAs. As Erving Goffman points out, “Often . . . broadcast talk involves a conversational mode of address, but,

- of course, merely a simulated one, the requisite recipients not being there in the flesh to evoke it” (Goffman, “Footing,” in *Forms of Talk* (Philadelphia, PA: University of Pennsylvania Press, 1981), 138).
- 6 Cascade SF, “Designing Voice User Interfaces with Cathy Pearl from Google,” YouTube video, 27:50, December 5, 2018, <https://www.youtube.com/watch?v=RdCmMMwaFRs&t=823s>.
- 7 References to “classroom discussion” refer to a class run by the lead author, in which most of the other authors were students. The exception is the second author, who is the lead author’s PhD student, and was a teaching assistant for a previous version of this course (before the inclusion of DVAs), and who later also ran that version of it. Most uses of the first-person plural (concerning DVAs) in this article refer to this class, most references to “students,” to this class’s students, and all quotations without citations are to work done as part of this class.
- 8 Erving Goffman’s participation framework (Goffman, “Footing,”) highlights how talk as action does not concern speakers alone, but is a product of interaction, where coparticipants (who, as bystanders and non-ratified participants, may even not be speaking) play an active role; see also Marjorie Harness Goodwin, “Processes of Mutual Monitoring Implicated in the Production of Description Sequences,” *Sociological Inquiry* 50 (1980): 303–17; Charles Goodwin, “Audience Diversity, Participation and Interpretation,” *Text* 6, no. 3 (1986): 283–316. To point out that talk in conversation is organized in sequences (and cannot be studied by considering utterances in isolation), see Goffman (“The Neglected Situation,” *American Anthropologist* 66, no. 6, pt. 2 (1964): 133–36) on turn-taking, the phenomenon extensively studied by the tradition of conversation analysis (e.g., Harvey Sacks, Emanuel Schegloff and Gail Jefferson, “A Simplest Systematics for the Organization of Turn-Taking for Conversation,” *Language* 50, no. 4 (1974): 696–735). Goffman points out that “the acts of speaking must always be referred to the state of talk that is sustained through the particular turn at talking, and that this state of talk involves a circle of others ratified as coparticipants” (Goffman, “Neglected Situation,” 135). Another relevant form of organization that conversation analysis highlights is the adjacency pair (on adjacency pairs, see, e.g., Emanuel Schegloff and Harvey Sacks, “Opening up Closings,” *Semiotica* 8, no. 4 (1973) 289–327; Emanuel Schegloff, *Sequence Organization in Interaction: A Primer in Conversation Analysis, Volume 1* (Cambridge: Cambridge University Press, 2007); see also Erving Goffman, *Relations in Public: Microstudies of the Public Order* (New York: Basic Books, 1971)). An adjacency pair (which can be expanded into more complex sequences) consists of utterances where the second part of the pair is selected from some range of possibilities that the first pair part defines. In other words, the first pair part activates a conditional relevance as it calls for the second pair part, which, in turn, exhibits some orientation to that relevance.

- 9 See, for example, John Cheney-Lippold, *We Are Data: Algorithms and the Making of Our Digital Selves* (New York, NY: New York University Press, 2017); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (New York, NY: St. Martin's Press, 2018); Patrick Michael Lynch, *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data* (New York, NY: W. W. Norton & Company, 2016); Safiya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York, NY: New York University Press, 2018); Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York, NY: Random House, 2016); Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, MA: Harvard University Press, 2015); Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (New York, NY: W. W. Norton & Company, 2015); Bruce Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World* (New York, NY: W. W. Norton & Company, 2018); Sara Wachter-Boettcher, *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech* (New York, NY: W. W. Norton & Company, 2017); Shoshana Zuboff, *The Age of Surveillance Capitalism*.
- 10 Our adoption of this expression is also importantly inspired by the discussion of *open text* in Umberto Eco, *The Open Work* (Cambridge, MA: Harvard University Press, 1962).
- 11 Harold Garfinkel, *Studies in Ethnomethodology* (Cambridge: Polity Press, 1984).
- 12 For example, Ronald Kline and Trevor Pinch, "Users as Agents of Technological Change: The Social Construction of the Automobile in the Rural United States," *Technology and Culture* 37, no. 4 (1996): 763–95; Bruno Latour, "Where are the Missing Masses: The Sociology of the Few Mundane Artifacts," in *Shaping Technology-Building Society: Studies in Sociotechnical Change*, eds. Wiebe Bijker and John Law (Cambridge, MA: MIT Press, 1992): 152–80; Marianne De Laet and Annemarie Mol, "The Zimbabwe Bush Pump: Mechanics of a Fluid Technology," *Social Studies of Science* 30, no. 2 (2000): 225–63.
- 13 See, for example, Ulrike Felt, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr, *Handbook of Science and Technology Studies* (Cambridge, MA: MIT Press, 2016).
- 14 For example, Yani Loukissas, whose *All Data Are Local: Thinking Critically in a Data-Driven Society* (Cambridge, MA: MIT Press, 2019) focuses on data as "cultural artifacts, manufactured and presented within contexts that matter" (xv), points out that this field can be defined by its focus on locality (xiii). Paul Dourish draws attention to the "material forms in which digital data are represented and how these forms influence interpretations and lines of actions" (Dourish, *The Stuff of Bits: An Essay on the Materialities of Information* (Cambridge, MA: MIT Press, 2017), 4). See also Kelly Gates, *Our Biometric Future: Facial Recognition Technology and the Culture of*

- Surveillance* (New York, NY: New York University Press, 2011); Kelly Gates, “The Cultural Labor of Surveillance: Video Forensics, Computational Objectivity, and the Production of Visual Evidence,” *Social Semiotics* 23, no. 2, 2013: 242–60; Adrian Mackenzie, *Machine Learners: Archeology of a Data Practice*, (Cambridge, MA: MIT Press, 2015); Samir Passi and Steve Jackson, “Data Vision: Learning to See Through Algorithmic Abstraction,” in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW ’17)* (New York, NY: ACM, 2017), 2436–47; Malte Ziewitz, “Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines,” *Social Studies of Science* 49, no. 5 (2019): 707–31.
- 15 Garfinkel, *Studies in Ethnomethodology*; Gilles Deleuze and Pierre-Félix Guattari, “Chapter 4: November 20, 1923: Postulates of Linguistics,” in *A Thousands Plateaus* (Minneapolis, MN: Minnesota University Press, 1980), 75–110.
- 16 See, for example, Jeff Coulter, “Discourse and Mind” in *Ludwig Wittgenstein: Critical Assessments, Second Series*, ed. Stuart Shanker and David Kilfoyle (London: Routledge, 2001), 143–61; Garfinkel, *Studies in Ethnomethodology*; Michael Lynch, *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory* (London: Routledge & Kegan Paul, 1985); and Michael Lynch, *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science* (Cambridge: Cambridge University Press, 1993).
- 17 For example, Charles Goodwin, “Action and Embodiment within Situated Human Interaction,” *Journal of Pragmatics* 32, no. 10 (2000): 1489–1522; Lucy Suchman, *Plans and Situated Actions: The Problem of Human-Machine Communication* (New York, NY: Cambridge University Press, 1987); and Lucy Suchman, “Embodied Practices of Engineering Work,” *Mind, Culture, and Activity* 7, nos. 1–2 (2000): 4–18; Jack Whalen and Erik Vinkuyzen, “Expert Systems in (Inter)Action: Diagnosing Machine Problems over the Telephone,” in *Workplace Studies: Recovering Work Practice and Information System Design*, eds. Paul Luff, Jon Hindmarsh, and Christian Heath (Cambridge: Cambridge University Press, 2001), 92–140.
- 18 See, for example, Zuboff, *The Age of Surveillance Capitalism*, and Cheney-Lippold, *We Are Data*.
- 19 Zuboff, *The Age of Surveillance Capitalism*, 18.
- 20 *Ibid.*, 35.
- 21 Lynch, *The Internet of Us*.
- 22 *Ibid.*, 3.
- 23 *Ibid.*, 6.
- 24 Andy Clarke and David Chalmers, “The Extended Mind,” *Analysis* 58, no. 1 (1998): 7–19.
- 25 Lynch, *The Internet of Us*, 5.
- 26 See also *ibid.*, 21–40.
- 27 *Ibid.*, 114
- 28 *Ibid.*, 125.

- 29 Ibid., 122, 123, 132. Emphasis in original.
- 30 Clark and Chalmers, “Extended Mind”; Andy Clark, *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence* (Oxford: Oxford University Press, 2003); Andy Clark, *Supersizing the Mind: Embodiment Action, and Cognitive Extension* (Oxford: Oxford University Press, 2008).
- 31 Edwin Hutchins, *Cognition in the Wild* (Cambridge, MA: MIT Press, 1995)
- 32 Hutchins criticizes the extended mind approach for the lack of a fuller appreciation of cultural practices in his “Enculturating the Supersized Mind,” *Philosophical Studies*, 152, no. 3 (2011): 437–46.
- 33 Lynch, *The Internet of Us*, 131.
- 34 Erving Goffman, “Alienation from Interaction,” *Human Relations* 10 (1957): 47–60.
- 35 Goffman, “Alienation,” 47; Gilbert Simondon, *On the Mode of Existence of Technical Objects* (Minneapolis: University of Minnesota Press, 1958), 135, 139, 17.
- 36 Simondon, *Mode of Existence*, 25–51.
- 37 For example, see Cynthia Breazeal, *Designing Sociable Robots* (Cambridge, MA: MIT Press, 2002). Social robots started to be built in the early 2000s at MIT, and their production is still exhibiting an upward swing.
- 38 Breazeal, *Designing*, xi.
- 39 Ibid., 8.
- 40 Ibid., 11.
- 41 Ibid., xii.
- 42 Ibid., 1–2.
- 43 Ashish Kapoor, “Fireside Chat with Cynthia Breazeal,” YouTube video, 30:46, July 20, 2018. <https://www.youtube.com/watch?v=MIGpDyaj0M0>.
- 44 For example, Evan Ackerman, “Jibo is Probably Totally Dead Now,” *Spectrum IEEE*, December 3, 2018, <https://spectrum.ieee.org/automaton/robotics/home-robots/jibo-is-probably-totally-dead-now>; Oliver Mitchell, “Jibo Social Robot: Where Things Went Wrong,” *The Robot Report*, June 28, 2018, <https://www.therobotreport.com/jibo-social-robot-analyzing-what-went-wrong/>.
- 45 Ackerman, “Jibo.”
- 46 Mitchell, “Jibo Social Robot.” The continuity between social robots and DVAs embedded in voice-controlled smart hubs is also noticed by consumers. Following Mitchell's online article, “Jeffrey M Harrison” comments: “Jibo was a great idea, cute, kids loved him, he never did tell a story to the grand kids, since the last update he’s not as fast answering questions, I unplugged him today before I went to work, jibo will go back in his box, be put in the attic, Jibo was Born Nov 21 2017 died Oct 30 2018 . . . so sad, what a waste . . . I have alexia now.”
- 47 Steven Connor, *Beyond Words: Sobs, Hums, Stutters and other Vocalizations* (London: Reaktion Books, 2014), 13–14; 17.
- 48 Breazeal, *Designing*, 5.

- 49 Judith Shulevitz, “Alexa, Should We Trust You?,” *The Atlantic*, November 2018, <https://www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844/>.
- 50 The publication of Shulevitz’s “Alexa, Should We Trust You?” closely followed that of the *New York Times* article.
- 51 See also Raya Jones, “What Makes a Robot ‘Social’?,” *Social Studies of Science* 47 no. 4 (2017): 556–79, for a related argument regarding social robots.
- 52 See also, for example, Zuboff, *The Age of Surveillance Capitalism*, 255.
- 53 For example, Morana Alač, “Moving Android: On Social Robots and Body-in-Interaction,” *Social Studies of Science* 39 no. 4 (2009): 491–528; Morana Alač, Javier Movellan, and Fumihide Tanaka, “When a Robot Is Social: Enacting a Social Robot through Spatial Arrangements and Multimodal Semiotic Engagement in Robotics Practice,” *Social Studies of Science* 41, no. 6 (2011): 893–926.
- 54 Variants of this take may be found in Freud as criticized by Gilles Deleuze and Pierre-Félix Guattari, *Anti-Oedipus: Capitalism and Schizophrenia* (New York: Penguin, 1972), or more recently in neuroscientific approaches to autism (e.g., Nancy Minshew and Diane Williams, “The New Neurobiology of Autism: Cortex, Connectivity, and Neuronal Organization,” *Archives of Neurology* 64, no. 7 (2007): 945–50).
- 55 Alač et al., “When a Robot is Social”; Alač, “Social Robots: Things or Agents?,” *AI & Society* 31, no. 4 (2016): 519–35.
- 56 We specifically refer here to a search we did on May 20, 2019.
- 57 Because of its voiced quality, DVAs manifest more outwardly what would also characterize, while being less apparent, other AI technologies, including their imaginary forms, such as Lynch’s neuromedia.
- 58 When Alač ran this class in previous years, she wanted to turn her students’ attention to the phenomena she has described in her writings: how the social character of a robot is enacted and maintained in specific events of robotics practice (Alač et al., “When a Robot Is Social”; Alač, “Zeigt auf den Roboter und schüttelt dessen Hand: Intimität als situativ gebundene interaktionale Unterstützung von Humanoidtechnologien.” *Zeitschrift für Medienwissenschaft*, 15 (2016): 41–71); how roboticists align their own bodies with robots, not only animating the technology, but enacting and distributing themselves through these interactional linkages (Alač, “Moving Android”); and how spatial arrangements play a role in those enactments (Alač, “Social Robots: Things or Agents?”). The problem was always, however, that social robots exclusively lived in research laboratories, and therefore were largely inaccessible to students. Beyond reading the instructor’s publications, talking with roboticists who visited their classes, or going on lab tours in social robotics, students could not witness how these machines are addressed, looked at, touched, or positioned in space; or what is being said around them in situations in which they are embedded. This has changed with the migration of social robots beyond the laboratory,

- relative to the commercial success of DVAs (notwithstanding their reduced format).
- 59 Natalie Jeremijenko, “If Things Can Talk, What Do They Say? If We Can Talk to Things, What Do We Say?,” *Electronic Book Review*, March 5, 2005, <https://electronicbookreview.com/essay/if-things-can-talk-what-do-they-say-if-we-can-talk-to-things-what-do-we-say/>.
- 60 If students themselves were not frequent users of a DVA, they could observe another frequent user whose use is familiar to them, such as a family member or roommate. The assignment also instructed them to observe DVA interactions for at least a couple of days before starting to record, so as to acquaint themselves with configurations of use, including social settings in which the use takes place. That would also let those they observe acquire familiarity with the observation situation, and give students the ability to organize recordings optimally (to consider, for example, how best to record audio, take photos, and field notes).
- 61 Erving Goffman, *The Presentation of Self in Everyday Life* (New York: Random House, 1959).
- 62 Students were asked to report on *how* interaction took place, who interacted with the DVA, whether anybody else was present (beside the person who directly conversed with the DVA), whether others who were present engaged in the interaction in any way, and whether these others noticed that the interaction took place. In addition, the students were to note emotions that were displayed and write down how frequently they observed related interactions in that very spot, whether the place was different in any way this time, and whether it changed in respect to the interaction in any way.
- 63 These framings may have to do with the newness of the technology, or it may be to do with the fact that the student population is not DVAs’ primary target group; we, however, don’t want to speculate.
- 64 Erving Goffman, “On Face-Work: An Analysis of Ritual Elements in Social Interaction,” *Psychiatry: Journal for the Study of Interpersonal Processes* 18 (1955): 213. Face-work, therefore, concerns “actions taken by a person to make whatever he is doing consistent with face. Face-work serves to counteract ‘incidents’—that is, events whose effective symbolic implications threaten face” (ibid., 216). Goffman points out that when someone, in an encounter, senses being “in wrong face or out of face, he is likely to feel ashamed and inferior because of what has happened to the activity on his account and because of what may happen to his reputation as a participant” (ibid., 214). Embarrassment “occurs in . . . relation to the real or imagined presence of others” (Goffman, “Embarrassment and Social Organization,” *American Journal of Sociology* 62, no. 3 (1956): 264) and “has to do with the figure the individual cuts before others felt to be there at that time” (ibid., 264–65). It takes place when the projected image of self is discredited (ibid., 268).
- 65 Goffman, *Presentation of Self*.

- 66 Scholars who go beyond “real-time observation, clippings and vignettes” (Emanuel Schegloff, “Goffman and the Analysis of Conversation,” in *Erving Goffman: Exploring the Interaction Order*, eds. Paul Drew and Anthony Wootton (Cambridge: Polity Press, 1988), 132) have criticized Goffman for his reliance on psychology and the individual’s inner while overlooking actual interactional systematics. Schegloff has pointed out that Goffman (particularly up to his 1974 publication of *Frame Analysis*) “defended the need for a psychology of the individual” despite recognizing, with an “explicit commitment,” that behavior is a product of the system (ibid., 94). A possible example of this commitment may be Goffman’s stating that embarrassment can “clearly show . . . itself to be located not in the individual but in the social system” in regard to situations where “individuals confront one another with selves incompatible with the ones they extend to each other on other occasions” (Goffman, “Embarrassment and Social Organization,” 269). Schegloff contends, however, that “putting ‘face’ at the centre of interaction drives Goffman’s account toward the individual and the psychological at two levels. On the one hand is the recurrently invoked direct account for, and understanding of, conduct by reference to a concern for preservation or demeaning of face (either own or other’s). On the other hand is the depiction of an organization of interaction which is driven by, whose *raison d’être* is, the individual and his/her interest—namely ‘face’” (Schegloff, “Goffman,” 95).
- 67 In his discussion of shame, Christian Heath (“Embarrassment and Interactional Organization,” in *Erving Goffman*, eds. Drew and Wootton, 155) has highlighted that Goffman’s “analytic model conceals the actual conduct of the participants, it is placed to one side as their behavior is conceptualized in terms of issues of self and identity. The actual interaction itself is lost, and with it our ability to explicate the social organization which underlies the production and recognition of the participants’ conduct in its natural settings.”
- 68 In her field note reproduced above, Flora mentions that she forgot to utter “Alexa” before asking a question, and so did not receive an answer, indexing with this the prescribed elements of voice interaction that users are to follow (such as addressing the DVA or directing a greeting toward it with the “wake word,” such as “Alexa” or “Hey Google”).
- 69 Garfinkel, *Studies in Ethnomethodology*.
- 70 Gail Jefferson, “Glossary of Transcript Symbols with an Introduction,” in *Conversation Analysis: Studies from the First Generation*, ed. Gene Lerner (New York: John Benjamins, New York, 2004), 13–31.
- 71 Garfinkel, *Studies in Ethnomethodology*. In that framework, Flora and Amber’s roommates, for example, would be characterized in terms of *non-ratified participants*, rather than in reference to a self and emotional states, such as the emotion of shame (defined in respect to a threat towards one’s *face*).
- 72 Simondon, *Mode of Existence*.

- 73 Since Max, as well as Dave in section 5, refer to Alexa with the female pronoun, we will, here, follow this convention.
- 74 When reporting on another interaction in which the two defer to the DVA, Max specifies, “We didn’t even discuss speaking to Alexa to ask her opinion on the subject, we just did it, as it is a naturally occurring daily thing that goes on in our apartment.”
- 75 The transcript is rendered according to the style of conversation analysis (see Jefferson, “Glossary of Transcript Symbols”). The transcript adopts the following conventions:
- (.) A dot in parentheses indicates a brief interval.
  - (( )) Double parentheses contain the transcriber’s descriptions.
  - (1.2) Numbers in parentheses indicate pause duration in seconds and tenths of seconds.
  - = The equal sign indicates no interval between the end of a prior and start of a next piece of talk.
  - A hyphen after a word or part of a word indicates a cut-off.
  - [ A left square bracket indicates the onset point at which a current speaker’s talk is overlapped by the talk of another.
  - ] A right square bracket indicates overlap offset
  - : A colon indicates that the prior syllable is prolonged.
  - — — Underscoring indicates stressing.
  - WORD Capital letters indicate that a section of talk is produced in higher volume than the surrounding talk.
  - .,? Punctuation markers are used to indicate ‘the usual’ intonation:
    - . A dot is used for falling intonation;
    - ? A question mark is used for rising intonation;
    - , A comma is used for rising and falling intonation.
- 76 Breazeal, *Designing*, xi.
- 77 Max, who is a more skilled user, will eventually rephrase the question in a more general format, as seen in Line 15 (“who is the best rapper alive”).
- 78 Max also provides an entry that reads, “Wrong wording told us denotative meaning,” which indicates his belief that his housemate’s inapt framing of the question caused Alexa’s trouble.
- 79 These are Max’s words.
- 80 See also Alač, “Moving Android,” for a parallel account in social robotics where not only the voice, but entire bodies are involved in robots’ enactments.
- 81 As in Goffman, *Presentation of Self*.
- 82 There are moments—as in Lines 19–25 and 33–35—where the housemates interact exclusively among themselves, yet even those moments may not entirely be about their positioning toward each other, but about

coordination that produces Alexa’s voice. After Alexa answers with “Chance the Rapper,” the housemates only briefly direct their comments to her (Lines 19 and 20: “I FEEL you, Alexa!” and “Alexa what the hell? Huh”), but, when they then quickly turn to each other, they engage in establishing whether the answer was “lit” or not (Lines 21–25). When Alexa provides “Missy Elliot” (Line 33) as a response to the best female rapper question, the two follow with an interaction among themselves, where Max, in Line 34, orients toward the housemate with a sing-song of Missy Elliot’s verse, presumably to have the housemate confirm it (which the housemate does in Line 35).

- 83 Alexa’s answer—despite its perspectival display (the DVA premises her statements with “in my opinion” in Line 16, and “I like” in Line 33)—is what the two follow. As Max reports, they will continue their gaming session while playing Chance the Rapper, despite originally being interested in a resolution between Kendrick Lamar (Max’s choice) and Drake (the housemate’s choice).
- 84 Consider this point in relation to Colin Jerolmack and Iddo Tavory, “Molds and Totems: Nonhumans and the Constitution of the Social Self,” *Sociological Theory* 32, no. 1 (2014): 64–77. In their discussion of nonhumans in the constitution of the social self, Jerolmack and Tavory criticize *constructivist* discussions of nonhumans for their exclusive preoccupation with how nonhumans reflect cultural symbols of humans who interpret them (65), as well as the “actor-network theory” for its “neglect of sociality” (64). In proposing to turn to how “nonhumans *organize* . . . one’s social position and relations” (65), the two authors discuss Clifford Geertz’s 1973 essay on the Balinese cockfight, to point out that “[p]arts of [Geertz’s] own narrative seem to contradict this constructivist reduction of nonhumans to passive objects onto which people project pregiven social arrangements” (73):

For instance, we catch glimpses of cocks “acting back” by refusing to fight at all of by continually running away. And we see men drawn to the birds as birds (not just status symbols), spending “an enormous amount of time with their favorites, grooming them, feeding them . . . of just gazing at them with a mixture of rapt admiration and dreamy self-absorption” (Geertz 1973: 418–19). For such “cock crazy” men, their birds structured the rhythms of their day. They “spend most of their life” with their roosters and organized their primary groups accordingly. Geertz described cliques of men “squatting idly in the council shed” with a rooster in their hands for hours, holding the birds between their thighs, bouncing them gently, and ruffling their feathers “with abstract sensuality” (419). (74)

While we agree with Jerolmack and Tavory that Geertz’s description provides more than an account of the birds as “passive objects onto which people project pregiven social arrangements” (73), when the two authors say that “it seems that the birds . . . patterned the men’s social relations

and formed the basis for a particular kind of group membership” (74), we find that they skip over what happens in the relationship between the human and nonhuman in terms of that relationship. If we take as an example the interaction in Max’s room from the previous section (where the housemates are asking the robot for opinions while being collaboratively involved in its functioning), it would be limiting to think about the human–nonhuman relationship only in terms of how it articulates the social relationship between Max and the housemate. This becomes even more obvious in what Dave reports. In Geertz’s writing, the relationship with the nonhuman concerns living beings—humans and roosters. In our examples, the relationship is with a robot whose expressive body is condensed to the bare voice, so if there is any of Geertz’s “abstract sensuality” present in the DVA case, it has to concern just that voice. Paying attention to the lived relationship with DVA voice allows us to describe how the self of an DVA’s interlocutor is open. In other words, approaching the problem from moments of actual interactions with nonhumans pushes us beyond the intentionalistic self of social psychology (that Jerolmack and Tavory, in their adoption of Mead’s social psychology (74), are resorting to), as we consider how both humans and nonhumans display a dynamic self (which may also dissipate through interaction).

- 85 See also Zuboff, *The Age of Surveillance Capitalism*, 268. These “skills” often connect to other devices (the IoT), broadening security concerns (e.g., Matthew Pascucci, “Alexa, Are You Keeping My Information Private?” *TripWire*, January 11, 2017, <https://www.tripwire.com/state-of-security/security-awareness/alexa-are-you-keeping-my-information-private/>). That in this case multiple devices are connected is indicated in Line 9 where Alexa says that “Amazon Music is streaming on another device.”
- 86 Goffman, “Footing,” 137. For further discussions of *participation*, see, for example, Marjorie Harness Goodwin, “Mutual Monitoring” and Charles Goodwin, “Audience Diversity.” A related idea can be found in Alfred Schutz, “On Multiple Realities,” *Philosophy and Phenomenological Research* 5, no. 4 (1945): 533–76, as also discussed in Alač, “Social Robots.”
- 87 For the meaning of the term “non-ratified participants,” see Goffman, “Footing,” 131. In pointing out the inadequacies of understanding talk in exclusively dyadic terms, Goffman suggested that “much of talk takes place in the visual and aural range of persons who are not ratified participants and whose access to the encounter, however minimal, is itself perceivable by the official participants” (*ibid.*, 132).
- 88 Goffman, “Footing,” 132.
- 89 *Ibid.*, 128.
- 90 For a discussion of the importance on the temporal flow in conversation in another AI setting (namely, an application of “expert systems” in a work setting), see Whalen and Vinkuyzen, “Expert Systems in (Inter)Action.”
- 91 Goffman, “Footing,” 134.
- 92 Hutchins, *Cognition in the Wild*.

- 93 Lynch, *The Internet of Us*, 131.
- 94 See, for example, Zuboff, *The Age of Surveillance Capitalism*; Kate Crawford and Vladan Joler, “Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources,” *AI Now Institute and Share Lab*, September 7, 2018, <https://anatomyof.ai>.
- 95 For an earlier discussion on how social robots call for a broadening of the idea of human extension into the world beyond tools, see Alač, “Social Robots,” 495–96.
- 96 Goffman, *Presentation of Self*, 252.
- 97 Schegloff, “Goffman,” 94.
- 98 Simondon, *Mode of Existence*, 174–90.
- 99 *Ibid.*, 187.
- 100 *Ibid.*, 174.
- 101 *Ibid.*, 194–95.
- 102 *Ibid.*, 196–98.
- 103 *Ibid.*, 197.
- 104 *Ibid.*, 196.
- 105 Goffman, “Alienation,” 47.
- 106 Simondon, *Mode of Existence*, 179.
- 107 Yuriko Saito, *Everyday Aesthetics* (Oxford: Oxford University Press, 2007).
- 108 *Ibid.*, 62.
- 109 *Ibid.*, 54.
- 110 Simondon, *Mode of Existence*, 178–79. Here he speaks of “the peaks of the mountains or certain, naturally magical, mountain passes, because they govern a land.”
- 111 Saito, *Everyday Aesthetics*, 52.
- 112 Thomas Disch, *The Brave Little Toaster* (New York, NY: Doubleday, 1986).
- 113 For example, see Christina Dunbar-Hester, *Low Power to the People: Pirates, Protest, and Politics in FM Radio Activism* (Cambridge, MA: MIT Press, 2014).
- 114 Cal Newport, *Digital Minimalism: Choosing a Focused Life in a Noisy World* (London: Portfolio, 2019).
- 115 Crawford and Joler, “Anatomy of an AI System.”

## References

- Ackerman, Evan. "Jibo is Probably Totally Dead Now." *Spectrum IEEE*. December 3, 2018. <https://spectrum.ieee.org/automaton/robotics/home-robots/jibo-is-probably-totally-dead-now>.
- Alač, Morana. "Moving Android: On Social Robots and Body-in-Interaction." *Social Studies of Science* 39, no. 4 (2009). 491–528.
- . "Social Robots: Things or Agents?" *AI & Society* 31, no. 4 (2016). 519–35.
- . "Zeigt auf den Roboter und schüttelt dessen Hand: Intimität als situativ gebundene interaktionale Unterstützung von Humanoidtechnologien." *Zeitschrift für Medienwissenschaft*, 15 (2016). 41–71.
- Alač, Morana, Javier Movellan, and Fumihide Tanaka. "When a Robot Is Social: Enacting a Social Robot through Spatial Arrangements and Multimodal Semiotic Engagement in Robotics Practice." *Social Studies of Science* 41, no. 6 (2011). 893–926.
- Bohn, Dieter. "Amazon Says 100 Million Alexa Devices Have Been Sold—What's Next? Amazon Isn't Letting Google Own CES without a Fight." *The Verge*, January 4, 2019. <https://www.theverge.com/2019/1/4/18168565/amazon-alexa-devices-how-many-sold-number-100-million-dave-limp>.
- Breazeal, Cynthia. *Designing Sociable Robots*. Cambridge, MA: MIT Press, 2002.
- Builder, Maxine. "The Gizmos to Buy Your Techie This Holiday (before They Sell Out)." *New York Magazine*, October 18, 2018. <http://nymag.com/strategist/article/best-tech-gifts-holiday-trends.html>.
- Cascade SF. "Designing Voice User Interfaces with Cathy Pearl from Google." YouTube video, 27:50. December 5, 2018. <https://www.youtube.com/watch?v=RdCmMMwaFRs&t=823s>.
- Cheney-Lippold, John. *We Are Data: Algorithms and the Making of Our Digital Selves*. New York: New York University Press, 2017.
- Clark, Andy. *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*. Oxford: Oxford University Press, 2003.
- . *Supersizing the Mind: Embodiment Action, and Cognitive Extension*. Oxford: Oxford University Press, 2008.
- Clark, Andy and David Chalmers. "The Extended Mind." *Analysis* 58, no. 1 (1998). 7–19.
- Connor, Steven. *Beyond Words: Sobs, Hums, Stutters and Other Vocalizations*. London: Reaktion Books, 2014.
- Coulter, Jeff. "Discourse and Mind." In *Ludwig Wittgenstein: Critical Assessments, Second Series*, Vol. 3, edited by Stuart Shanker and David Kilfoyle, 143–61. London: Routledge, 2001.

- Crawford, Kate and Vladan Joler. "Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources." *AI Now Institute and Share Lab*, September 7, 2018. <https://anatomyof.ai>.
- De Laet, Marianne and Annemarie Mol. "The Zimbabwe Bush Pump: Mechanics of a Fluid Technology." *Social Studies of Science* 30, no. 2 (2000). 225–63.
- Deleuze, Gilles and Pierre-Félix Guattari "Chapter 4: November 20, 1923: Postulates of Linguistics." In *A Thousands Plateaus*, 75–110. Minneapolis, MN: Minnesota University Press, 1980.
- . *Anti-Oedipus: Capitalism and Schizophrenia*. New York, NY: Penguin, 1972.
- Disch, Thomas. *The Brave Little Toaster*. New York, NY: Doubleday, 1986.
- Dourish, Paul. *The Stuff of Bits: An Essay on the Materialities of Information*. Cambridge, MA: MIT Press, 2017.
- Dunbar-Hester, Christina. *Low Power to the People: Pirates, Protest, and Politics in FM Radio Activism*. Cambridge, MA: MIT Press, 2014.
- Eco, Umberto. *The Open Work*. Cambridge, MA: Harvard University Press, 1962.
- Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York, NY: St.Martin's Press, 2018.
- Felt, Ulrike, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr. *Handbook of Science and Technology Studies*. Cambridge, MA: MIT Press, 2016.
- Garfinkel, Harold. *Studies in Ethnomethodology*. Cambridge: Polity Press, 1984.
- Gates, Kelly. *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. New York, NY: New York University Press, 2011.
- . "The Cultural Labor of Surveillance: Video Forensics, Computational Objectivity, and the Production of Visual Evidence." *Social Semiotics* 23, no. 2 (2013). 242–60.
- Goffman, Erving. "On Face-Work: An Analysis of Ritual Elements in Social Interaction." *Psychiatry: Journal for the Study of Interpersonal Processes* 18, 1955. 213–31.
- . "Embarrassment and Social Organization." *American Journal of Sociology* 62, no. 3 (1956). 264–71.
- . "Alienation from interaction." *Human Relations*, 10 (1957). 47–60.
- . *The Presentation of Self in Everyday Life*. New York, NY: Random House, 1959.
- . The Neglected Situation. *American Anthropologist* 66, no. 6, pt. 2 (1964). 133–36.

- . *Relations in Public: Microstudies of the Public Order*. New York, NY: Basic Books, 1971.
- . “Footing.” In *Forms of Talk*, 124–57. Philadelphia, PA: University of Pennsylvania Press, 1981.
- Goodwin, Charles. “Audience Diversity, Participation and Interpretation.” *Text* 6, no. 3 (1986). 283–316.
- . “Action and Embodiment within Situated Human Interaction.” *Journal of Pragmatics* 32, no. 10 (2000). 1489–522.
- Goodwin, Marjorie Harness. “Processes of Mutual Monitoring Implicated in the Production of Description Sequences.” *Sociological Inquiry* 50 (1980). 303–17.
- Green, Dennis. “Amazon Reveals Its Best-Selling Gadgets of the Holiday Season as Smart Home Tech Takes Over Gift-Giving.” *Business Insider*, December 27, 2018. <https://www.businessinsider.com/amazon-best-selling-gadgets-holiday-season-2018-12>.
- Heath, Christian. “Embarrassment and Interactional Organization.” In *Erving Goffman: Exploring the Interaction Order*, edited by Paul Drew and Anthony Wootton, 136–60. Cambridge, Polity Press, 1988.
- Hutchins, Edwin. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.
- . “Enculturating the Supersized Mind.” *Philosophical Studies* 152, no. 3 (2011). 437–46.
- Jefferson, Gail. “Glossary of Transcript Symbols with an Introduction.” In *Conversation Analysis: Studies from the First Generation*, edited by Gene Lerner, 13–31. New York: John Benjamins, 2004.
- Jeremijenko, Natalie. “If Things Can Talk, What Do They Say? If We Can Talk to Things, What Do We Say?” *Electronic Book Review*, March 5, 2005. <https://electronicbookreview.com/essay/if-things-can-talk-what-do-they-say-if-we-can-talk-to-things-what-do-we-say/>
- Jones, Raya. “What Makes a Robot ‘Social’?” *Social Studies of Science* 47, no. 4 (2017). 556–79.
- Kapoor, Ashish. “Fireside Chat with Cynthia Breazeal.” YouTube video, 30:46, July 20, 2018. <https://www.youtube.com/watch?v=MIGpDyaj0M0>.
- Kline, Ronald and Trevor Pinch. “Users as Agents of Technological Change: The Social Construction of the Automobile in the Rural United States.” *Technology and Culture* 37, no. 4 (1996). 763–95.
- Latour, Bruno. “Where Are the Missing Masses: The Sociology of the Few Mundane Artifacts.” In *Shaping Technology–Building Society: Studies in Sociotechnical Change*, edited by Wiebe Bijker and John Law, 152–80. Cambridge, MA: MIT Press, 1992.

- Lincoln, Kevin. "Where Is the Boundary between Your Phone and Your Mind?" *The Guardian*, December 9, 2018. [www.theguardian.com/us-news/2018/dec/09/tech-mind-body-boundary-facebook-google](http://www.theguardian.com/us-news/2018/dec/09/tech-mind-body-boundary-facebook-google).
- Loukissas, Yanni. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press, 2019.
- Lynch, Michael. *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. London: Routledge & Kegan Paul, 1985.
- . *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press, 1993.
- Lynch, Patrick Michael. *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York, NY: W. W. Norton & Company, 2016.
- Mackenzie, Adrian. *Machine Learners: Archeology of a Data Practice*. Cambridge, MA: MIT Press, 2015.
- Manjoo, Farhad. "A Future Where Everything Becomes a Computer Is a Creepy as You Feared." *New York Times*, October 10, 2018. <https://www.nytimes.com/2018/10/10/technology/future-internet-of-things.html>
- Minshew, Nancy and Diane Williams. "The New Neurobiology of Autism: Cortex, Connectivity, and Neuronal Organization." *Archives of Neurology* 64, no. 7 (2007). 945–50.
- Mitchell, Oliver. "Jibo Social Robot: Where Things Went Wrong." *The Robot Report*, June 28, 2018. <https://www.therobotreport.com/jibo-social-robot-analyzing-what-went-wrong/>.
- Molla, Rani. "The Holidays Ushered in a Smart Speaker Explosion Starring Amazon's Alexa and Google's Assistant." *Recode*, January 8, 2019. <https://www.recode.net/2019/1/8/18173696/amazon-alexa-google-assistant-smart-speaker-sales-npr>.
- Nass, Clifford and Jonathan Steuer. "Voices, Boxes, and Sources of Messages: Computers and Social Actors." *Human Communication Research* 19, no. 4 (1993). 504–27.
- Newport, Cal. *Digital Minimalism: Choosing a Focused Life in a Noisy World*. London: Portfolio, 2019.
- Noble, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: New York University Press, 2018.
- NPR and Edison Research. *The Smart Audio Report, Winter 2018*. January 2019. Accessed June 26, 2019. <https://www.nationalpublicmedia.com/wp-content/uploads/2019/01/Smart-Audio-Report-Winter-2018.pdf>.
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Random House, 2016.

- Pascucci, Matthew. "Alexa, Are You Keeping My Information Private?" *TripWire*, January 11, 2017. <https://www.tripwire.com/state-of-security/security-awareness/alexa-are-you-keeping-my-information-private/>.
- Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press, 2015.
- Passi, Samir and Steve Jackson. "Data Vision: Learning to See through Algorithmic Abstraction." *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. New York, NY: ACM, 2017. 2436–47.
- Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50, no. 4 (1974). 696–735.
- Saito, Yuriko. *Everyday Aesthetics*. Oxford: Oxford University Press, 2007.
- Schegloff, Emanuel. "Goffman and the Analysis of Conversation." In *Erving Goffman: Exploring the Interaction Order*, edited by Paul Drew and Anthony Wootton, 89–135. Cambridge: Polity Press, 1988.
- . *Sequence Organization in Interaction: A Primer in Conversation Analysis, Volume 1*. Cambridge: Cambridge University Press, 2007.
- Schegloff, Emanuel and Harvey Sacks. "Opening up Closings," *Semiotica* 8, no. 4 (1973). 289–327.
- Schneier, Bruce. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York, NY: W. W. Norton & Company, 2015.
- . *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World*. New York, NY: W. W. Norton & Company, 2018.
- Shulevitz, Judith. "Alexa, Should We Trust You?" *The Atlantic*, November, 2018. <https://www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844/>.
- Schutz, Alfred. "On Multiple Realities." *Philosophy and Phenomenological Research* 5, no. 4 (1945). 533–76.
- Simondon, Gilbert. *On the Mode of Existence of Technical Objects*. Minneapolis, MN: University of Minnesota Press, 1958.
- Suchman, Lucy. *Plans and Situated Actions: The Problem of Human-Machine Communication*. New York: Cambridge University Press, 1987.
- . "Embodied Practices of Engineering Work." *Mind, Culture, and Activity* 7, nos. 1–2 (2000). 4–18.
- Wachter-Boettcher, Sara. *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*, New York, NY: W. W. Norton & Company, 2017.
- Whalen, Jack and Erik Vinkuyzen. "Expert Systems in (Inter)Action: Diagnosing Machine Problems over the Telephone." In *Workplace Studies: Recovering Work Practice and Information System Design*, edited

by Paul Luff, Jon Hindmarsh and Christian Heath, 92–140. Cambridge: Cambridge University Press, 2001.

Ziewitz, Malte. “Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines.” *Social Studies of Science* 49, no. 5 (2019). 707–31.

Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: Public Affairs, 2019.