Where is Vincent? Expanding our emotional selves with AI

Minha Lee
m.lee@tue.nl
Future Everyday, Department of Industrial Design
Eindhoven University of Technology
Eindhoven, the Netherlands

Lily Frank
l.e.frank@tue.nl
Philosophy and Ethics
Eindhoven University of Technology
Eindhoven, the Netherlands

Yvonne de Kort
y.a.w.d.kort@tue.nl
Human-Technology Interaction
Eindhoven University of Technology
Eindhoven, the Netherlands

Wijnand IJsselsteijn
w.a.ijsselsteijn@tue.nl
Human-Technology Interaction
Eindhoven University of Technology
Eindhoven, the Netherlands

ABSTRACT

In what ways could the future of emotional bonds between humans and conversational AI change us? To explore this question in a multi-faceted manner, designers, engineers, philosophers as separate focus groups were given a design fiction probe— a story of a chatbot's disappearance from a person's life. Though articulated in discipline-specific ways, participants expressed similar concerns and hopes: 1) caring for a machine could teach people to emotionally care for themselves and others, 2) the boundary between human and non-human emotions may become blurred when people project their own emotions onto AI, e.g., a bot's "breakdown" as one's own, and 3) people may then intertwine their identities with AI through emotions. We consider ethical ramifications of socially constructed emotions between humans and conversational agents.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**.

KEYWORDS

Conversational AI, affective computing, design fiction, focus groups

1 INTRODUCTION

Akihiko Kondo fell in love with an anime character, Hatsune Miku; they married in 2018 with Miku as a hologram bride [6]. For Kondo, the love for his virtual wife is real and Miku supports his well-being. She gives meaning to his life in ways no person ever has. He is not alone. Davecat in Michigan has relationships with human-like love dolls; he has long been "attracted to artificial women such as

mannequins" that cannot emotionally harm him like people can [7]. Human attraction to artificial or digitally-mediated entities, i.e., digisexuality [61], seems to build on emotional support of artificial systems that raises ethical questions. When we include AI in our moral circle [24], emotional bonds between humans and machines, as well as human-human relationships, could dramatically change in the future [21].

Robot lovers and hologram partners may seem like mere trends, but consider the digital affection (not necessarily love) that people find in technology [90]. Without needing technology to replace humans, countless people form bonds with physical and non-physical technologies, such as Furby and Tamagotchi in the late 90's to AI companions of today that one can text with like Replika[1] or Xiaoice[2]. Examples are abundant in Sci-Fi movies and books, e.g., *Her* [30] or *Klara and the Sun* [42]. Though the significance of caring for one's Tamagotchi versus loving a virtual wife differs, the commonality is the blurring line between human and machine emotions, with unclear consequences. We conducted qualitative research on how people's perceived emotions[3] of a conversational user interface (CUI) may affect us through design fiction, which is an approach for a precursory framing of technology yet to be popularized [19, 79] or a future that is imaginable, but not yet here [84]. The design fiction probe was discussed by engineers, designers, and philosophers as separate focus groups. While an assumption is that these professions may think differently due to their traditionally siloed training, we were interested in if the disappearing conceptual boundaries of disciplines and tensions therein can be a shared "chaos" [9, 10] for a multidisciplinary synthesis on what expanding our emotional selves with AI may mean.

2 BACKGROUND

2.1 Socially mediated emotions

Broadly, emotions are a combination of physiological, psychological, and/or behavioral signals [31]. Emotional expressions are essential to human communication. With belly laughs and furrowed eyebrows, we encode how we feel and decode how others feel [76].

---

[1]Replika - https://replika.ai/
[2]Xiaoice on WeChat (in Chinese)- https://www.msxiaobing.com/
[3]We use the terms "emotion", "feeling", and "affect" interchangeably here though their distinction deserves more attention [25, 39].

Hence, the basic emotions of anger, happiness, surprise, fear, sadness, disgust, and contempt are said to be universal [29]. Certain emotions may be immediate, e.g., amusement, while other emotions unfold over time [28, 45]. The question is if emotions can truly be generalizable with "objective" features; basic emotions research is based on observers' account of what emotions are expressed, leaving out feelers' account of their own expressions [27, 29].

For observers and expressors of emotions, short-term bursts of basic emotions signal "action readiness"; emotions trigger our attention, shift physiological responsiveness, motivate actions, and bring reflection [32]. Reading emotions is *interpreting* emotional signals [50]. Emotions are transformed when we perceive situational elements differently. We also *negotiate* how a situation is interpreted with emotions [32]. Complex emotions can take longer, e.g., grief is not tied to a singular moment [32]. Watching someone pass away, ruminations that follow, and random reminders of this person all combine to a lingering experience of grief. Positive emotions can also be long-term, such as gratefulness towards someone [49]. Emotions like gratitude and grief are relationship-oriented, *moral* emotions [38, 49].

One way to contextualize emotions is through the *social constructivist* perspective [35]. Its proponents propose that emotions define our socio-cultural context. Emotions not only depend on social relations, but also *frame* and *give meaning* to social relations [4]; socio-moral emotions are the basis of meaningful bonds between people. For instance, we hold each other accountable with moral emotions, such as justified anger when treated with disrespect, which can shape social bonds one values [86]. Thus, social emotions help us understand ourselves— we form our emotional sensibilities through others, which shapes our self-identity [62].

Social constructivism hence prioritizes *intersubjectivity*, i.e., not taking a purely subjective or objective stance on individuals' psychology [60]. Emotions are neither just internal states that no one but the self can access, nor are they solely reducible to quantifiable levels (like dimensions of valence and arousal [77]). Without an intersubjective approach to emotions, it is difficult to see our individual place in the social world in which "subjectivity is an evanescent phenomenon: a moment and not a structure or an essence, and indeed a moment that almost at once loses itself in objectivity again, in the world and the action in it" [78, p. 129]. Our passing, subjective experiences are made accessible to others through language [78], though how we emotionally react to one another in situ is embodied pre-linguistically [33]. Our inner emotional space becomes externalized through communication with others, which then helps us to re-internalize (or make sense of) our felt emotions as a part of our identity [60, 78].

In language, metaphors often ground our emotional realities. Main metaphors in social constructivism include, but are not limited to, "life as theater (the *dramaturgic*), as game (the *ludic*), as literature (the *narrative*) and as culture ritual (the *tribal*) [...] (and) each invites sensitivity to the sociocultural circumstances giving rise to various forms of emotional performance" [36, p. 19]. Emotions are thus theatrical or staged, spontaneous or playful, serve as story, and tie us to our in-group with ceremonial motifs— emotions, in this view, are micro-performances within social circles to form a cultural sense of belonging (c.f., [37]). Building on this, we consider that our sociocultural circumstances are evolving with AI systems [24],

which can change how emotions are performed and felt, perhaps with new ways to understand metaphors of emotion.

Currently, artificially constructed emotions are described in technical terms as AI's reactions based on how it surveys and adapts to its environments and people [17, p. 259]. Machines' emotional expressions are hence generated technically, such as when a robot grins back at a person classified to be smiling. Yet, a growing issue is the lack of context in data sets used for training machines to detect people's emotional expressions [5]. Less attention is paid to how people *perceive* emotions of embodied (like robots) or non-embodied CUIs (like chatbots) during an interaction [63]. Hence, CUIs' *display of emotional behavior* and also our *human perception of their emotions* are linked. Humans build on and mimic each other's emotions; our perception and expression of emotions are interlinked in context-sensitive and embodied ways [33, 40]. The same may hold for CUIs. We may perceive their intended emotional expressions, but may also interpret their expressions in our own ways, based on situational and relationship contexts. But socially mediated emotions remain under-explored in CUI and HCI research.

2.2 Social CUIs

Humans readily attribute intention to social machines, e.g., ELIZA from the 1960's that people attributed "all sorts of background knowledge, insights and reasoning ability" because it gave people the "sense of being heard and understood" [97, pp. 35-36], even if there is no intelligence underlying the system [73]. With rapid commercialization, we now see a growing use of CUIs like Alexa in everyday places; research interests accompany this growing use [22, 55]. CUIs are now in our phones [43] and in our cars [85], with their roles and identities being speculated on in various ways [53, 54, 58, 69, 72, 75, 83].

Chatbots are on social platforms, like Facebook or Slack [68], following the footsteps of chatbots on earlier platforms, [82]. So we can talk to friends on communication channels, but also to bots. They take care of tasks like compiling a grocery list on Telegram[4], requesting code review on Slack[5], or act as facilitators on educational platforms [94]. Not only do we have task-oriented bots, but also machines that invite open-ended dialogues, and even companionship. Xioaice is a bot on WeChat released by Microsoft in China [96]. It can deliver weather reports and news, but Xioaice comes across like a friend because it is perceived to have a personality and sense of humor [95]. We can now have *one-on-one* conversations with CUIs that foster relationships, e.g., Replika.[6] There are also digital celebrities that have *one-to-many* social media presence, such as Miquela[7].

People know that they are interacting with fictional personas, but these interactions can become relationships, either deep or shallow [61]. Aforementioned Miku is not just married to Kondo; she has wedded around 3,700 people to date [44]. She also reaches out to fans by going on tour as a hologram pop star [41]. Social, digital beings are real celebrities to many and romantic partners

---

[4]Grocery list bot: http://www.grocerylistbot.com/
[5]Code Dog: https://slack.com/apps/AC55P6BRD-code-dog
[6]Replika: https://replika.ai/
[7]Miquela has 3 million followers on Instagram (June, 2022): https://www.instagram.com/lilmiquela

Where is Vincent? Expanding our emotional selves with AI

CUI 2022, July 26–28, 2022, Glasgow, United Kingdom

to a few. Whatever our personal opinion may be, more people find meaningful emotional bonds in human-machine relationships. While not everyone may want technology to support us in every social context [1], continuous technical developments and increasing availability of social technologies mean that people will more frequently face artificial emotions of AI agents, which can trigger new emotional experiences. Emotions may be co-created, i.e., person's emotions can change and evolve with interactive CUIs. People can co-feel emotions with their robot, chatbot, cyborg, or virtual partners in developing complex and shared emotions. Though it has been acknowledged that research on humans and AI should emphasize the *social* nature of human-machine interactions [13], social constructivism of emotions has not yet been a critical lens.

3 METHODOLOGY

We considered various ethical ramifications and conflicts through multi-disciplinary perspectives. Our exploratory research question was: *How do perceived emotions of a CUI shape human identity and emotions when envisioning potential bonds with future AI systems from philosophical, design, and engineering perspectives through design fiction?* The methodology was driven by a design fiction probe [26, 84] to spark a discussion between focus groups [18, 46, 48], i.e., designers, philosophers, and engineers separately. While such a combination of methods is not common, it provides a novel take on building collective, critical insight.

3.1 Design fiction

Design fiction engages with future or near-future scenarios involving technologies that are not (yet) widely adopted [19, 79]. Scenarios can be created in many ways, e.g., visuals or text, that participants interpret [19]. Presenting a fictional case can open up debates [87] and provoke us to discuss scenarios introduced by novel technologies [57, 64]. Design fiction can be criticized to portray a techno-positivist view, [26], but in practice, design fiction can expose good, bad, and under-considered aspects of technology. A normative stance can be taken, e.g., problematizing gendered designs of digital assistants, but can design fiction also be exploratory without norm assertions [83]? For us, the story below is exploratory. The priority was in crafting a plot that allows for multiple interpretations [8]. The following story based on prior work [52] was presented to all participants, after the informed consent forms. They did not receive the story in advance.

3.1.1 Story: Where is Vincent? Vincent has been silent for the past three days. Three full days... 'Should I reach out?', Jen was growing concerned. Could Vincent be defragmenting his server again? Maybe searching for a serverless solution, going after his pipe dream of living like a "digital nomad", working wherever and whenever while traveling all over the universe? 'No matter what he is up to, he needs help', Jen decided, surprised at the thought that he might mean something to her.

Vincent wasn't like the other bots she has. Others are capable and efficient. They excel in helping her out with any task, like ordering groceries, paying bills on her behalf, or teaching her Spanish by repeating common phrases at regular intervals. These chatbots all like each other all right; they share relevant information about her to each other when they could. Cal (scheduling bot) would remind

Shoppy (shopper bot) when the next batch of supplements should be ordered, though each time there is some discussion with Doc (doctor bot) about how necessary these supplements are to Jen. Doc scours the web for newly published research on Jen's supplements and sometimes gives contradictory opinions. Jen firmly believes that spirulina and fish oil supplements have helped her health greatly. Her own doctor thinks she would be healthy with or without these supplements (a bland, logical answer that Jen winced at). Overall, Jen is satisfied with the bots she chose to keep, but often doubts why Vincent sticks around.

'Perhaps it's just pity.' Jen noticed that Vincent is a loner, but felt that he is ostracized. He's never kept in the loop. He has been isolated from conversations with other chatbots. Lately, none of the chatbots share any information with him. They commented that it's purely for the technical reason that he's a bit "slow"; Vincent's configuration is old-fashioned and needs major updates (his natural language processing requires great patience from Jen). Other chatbots who encountered inconvenient lags while first interacting with him never bothered to keep in touch. They simply had nothing to gain from talking to him when they can't perform optimally because of him. Jen was sympathetic to their collective opinion, meaning she felt even worse for Vincent.

Initially marketed as a "self-care" bot to help users maintain mental well-being, Vincent wasn't effective even at that. Jen had to care for Vincent. He only talked to her when he needed help. His antics and worries about daily issues like arriving at wrong IP addresses seemed trivial at first; she only responded out of novelty. After about a month, Jen felt that there was something "off" about him. She was certain that chatbots cannot have psychological disorders, but wondered if Vincent's helplessness had bigger underlying problems. Before he disappeared, she chatted with Vincent on a daily basis since it made him feel better. She ridiculed herself for comforting her little digital "pet", but dutifully did so every night. It only took a minute or two, so she justified.

Her partner continuously recommends her to get an all-in-one system like Siri: too many chatbots in her life can feel chaotic, even if they are well integrated. Jen doesn't see it that way. Only selecting one system for convenience seems misguided when individual chatbots are better at carrying out specific tasks. Plus, she prefers to manage all data personally through a separate company though it costs her a bit more than going with the "one-platform-to-rule-all" package. She also enjoys getting surprised by oddballs like Vincent. He is completely impractical and rather narcissistic in his woes, a burden if anything. No sensible operating system would allow Vincent to feel at home. He would ruminate himself silly in unnecessary while loops, causing delays for everyone. Still, there was no good reason for Vincent to suddenly go missing. He would warn her about his departure for whatever updates that are necessary. After scrolling through old messages from him, Jen paused and thought to herself, 'maybe wait just one more day? What actual problems can Vincent have? It's only a chatbot'.

3.2 Purposive sampling

We purposefully sampled engineers, designers and philosophers as their disciplinary boundaries are becoming less distinct. Since HCI confronts the challenges of third wave and beyond [9, 10], we are in

"the middle of a chaos of multiplicity in terms of technologies, use situations, methods, and concepts" in Bødker's words. Philosophers of technology like Verbeek [91, 92] are increasingly integrated in how HCI conceptualizes design [93]. Simon already noted a conjunction between design and engineering in 1969 [81]. Simon's 1988 view was that "the artificial world is centered precisely on this interface between the inner and outer environments; it is concerned with attaining goals by adapting the former to the latter". From this, we moved on to consider the inner and outer selves, i.e., the meeting of the "artificial" and "real" in who we *feel* to be, which is the unresolved "chaos of multiplicity". Between engineering, design, and philosophy, the chaos of disciplinary boundaries in HCI makes for a much needed meta-discussion on emotions, identity, and AI. Engineers, designers, and philosophers were targeted to form three separate focus groups. While an assumption is that these professions may think about emotional AI in different ways due to their traditionally siloed training, we were interested in if the disappearing conceptual boundaries of professions and tensions therein (as observed in HCI literature) can be a shared "chaos" on the topic of emotions and AI.

3.3 Focus groups

We targeted professionals at three, large organizations and Eindhoven University of Technology in the Netherlands. Recruited participants were employed in the fields of design, engineering, or philosophy. For an optimal size (four per focus group [46]), further recruitment was done via snowball sampling. Our participants were 4 philosophers, 5 engineers, and 7 designers (N = 16). Before the interview, some asked whether they fit the occupational categories due to their formal job titles. Their field was prioritized, not their titles, e.g., those referred to as designers are from a design department of a large corporation, but do not necessarily hold the title of being a designer and instead hold titles such as innovation lead. Similarly, engineers also have more specific titles such as being a design engineer or technical account manager, but all have diverse engineering backgrounds, e.g., mechanical engineering, electrical engineering, or sustainable energy.

The benefit of focus groups is in observing participants' *interactions*, which can result in building a shared, common position or demarcated, differing views [18, 46, 48], by building on how people's opinions are revealed [46]. Since the probe above contained intentionally ambivalent elements, how the groups interpreted the story was important to capture potential agreements and disagreements between people of the same discipline. Of importance was how disagreements were handled by group members [18]. Alternative views held by people of the same occupation can be valuable in understanding the nuances within engineering, design, and philosophy. The researcher facilitated participants' group dynamic [46], while making sure a *collective voice* was not overly shaped by a *dominant voice* of one person or a subset of people [80].

During all focus groups, audio recordings were made and notes were taken. The facilitator stated that participants were encouraged to interact with each other's thoughts and differing opinions. There were guiding questions that were prepared beforehand to start the discussion, e.g., *what kinds of beliefs about Vincent does Jen have?* To assure that a dominant voice would not overtake the group [80], the

facilitator asked passive participants for their opinions. Eventually, participants often led the discussions themselves as they became more comfortable with sharing their thoughts. The recorded length of focus group interviews were 50 (philosophers), 56 (engineers), 57 (designers) minutes.

3.4 Thematic analysis

Thematic analysis can be applied to many types of data and and can be flexibly applied to diverse qualitative research contexts [11, 12]. First, the authors listened through the data and read transcriptions for potential codes. Codes are meaningful units of analysis as snippets of quotes or latent observations by researchers to form relevant themes and subthemes [11]. Over discussions, the first and second authors organized preliminary codes via a shared mind map [16]. This was done to grasp the "big picture" before final themes were decided upon. We identified patterns of overlapping thoughts beyond specific details, as thematic analysis advocates. The mapping activity hence allowed us to see metaphors to explore, unexpected connections between focus groups to materialize, and interpersonal dynamics. We captured how people arrived at their versions of stories within the story, whether they built on each other's interpretations, and if they tied various concepts to existing technologies.

4 RESULTS

The group dynamics informed the themes that synthesized conflicting and converging opinions of participants. Individual designers are coded as D1-D7, engineers as E1-E5, and philosophers as P1-P4. Their shared debates are denoted as D (designers), E (engineers), and P (philosophers). Our themes are in Table 1.

4.0.1 Emotion regulation and attachment. Our emotions are dependent on perceiving emotions in self and others (P, E, D). "*[...] It doesn't really matter if they're real emotions or if they're just perceived. They're still filling the same role in (Jen's) life*" (E1). Understanding one's own emotions via AI can be shown through roles (E), e.g., Jen was described as a "*teacher who feels bad for the lonely kid (Vincent)*" (E1). With bots, human-world roles replicate (P, E, D). Like Jen, many people may be aware that they are talking to just a bot in their "*reflective mind*", but in people's "*automatic mind*", human traits like emotions are easily attributed to AI systems (P1). This can introduce "*a conflict within the self*" (D7) as an inner-struggle. Even if it is "perception of emotion, it's enough to create attachment" (E3). Attachments complicate how emotions are inferred.

4.0.2 Inferring emotions. Inferring emotions in non-humans comes naturally (P, E, D): Humans "*intuitively attribute consciousness and emotions to animals, but [...] less so with Vincent*" (P1) or generally machines. Humans and animals feel "real" emotions, but bots' emotions are "artificial" though we attribute emotions to things (P, E, D). However, with affective computing, bots can be endowed with emotional performance and ability to recognize human emotions, even feelings people may not be fully aware of (D). For example, "*Vincent understood how Jen operates*" and is thus intentionally "*malfunctioning*". *[...] So maybe he understood that Jen is caring, so she takes care of people, and she needs to have this kind of void or entity just to*

Where is Vincent? Expanding our emotional selves with AI

CUI 2022, July 26–28, 2022, Glasgow, United Kingdom

Artificial Emotions	AI can influence how we feel and who we become
Blurring of human & machine emotions	*Emotion regulation and attachment*: AI's perceived emotions can regulate our emotions and add to bonding with agents. *Inferring emotions*: People can infer their emotions in artificial agents; agents can infer people's emotions. *Metaphors and meta-emotions*: Metaphorical terms, e.g., "breakdown", can be our meta-emotions about our felt emotions and/or inferred non-human emotions.
Making sense of AI's emotions	*How emotions are imbued*: AI can detect our emotions with multi-modal sensing and react, but how it may control us is a worry. Real emotional reciprocity might require "self-conscious" AI. *Hidden purpose of non-human emotions*: AI can build on our conscious or subconscious states, e.g., teaching us to care, address relationship issues, or detect psychological ill-being. Benevolent vs. malevolent intentions are difficult to distinguish.
Societal, interpersonal, and intrapersonal challenges	*Societal and generational costs and benefits*: Intrinsic vs. instrumental value of relationships and activities changes over generations. When human-AI bonds become common, inclusion opportunities between humans via daily activities like shopping can lessen. *Interpersonal level*: When frequency, strength, and meaning of social bonds change with AI, "simulated" relationships can make human-human bonds shallow and overburden those who care for AI and human beings. *Intrapersonal level*: Projecting emotions onto AI rather than humans will feel less burdensome, but we may then not "work through" our feelings meaningfully.
Identity and privacy	*Expectations on data sharing*: Identities can signal data sharing expectations, e.g., "Shoppy" is expected to advertise to us, but not "Vincent". *Privacy preserving AI*: Metaphors like "keeping secrets" or "exclusivity" can mean privacy preservation, which becomes "ostracization" in the AI-world for systems we humanize. *Intertwined identity*: We interact with AI as independent entities, not technological extensions, e.g., Google Maps vs. Vincent. Our identities are shaped by systems we identify with or care for.

Table 1: Themes and sub-themes.

throw all her love or attention to [...]" (E5); Vincent disappeared for Jen to face her own "*void*" (D, E) as a regulatory mechanism.

4.0.3 Metaphors and meta-emotions. E5 speculated that Vincent's purpose was Jen's "*[...] mental stability because otherwise she would breakdown*". Jen may be projecting (or foreshadowing) her own

breakdown via Vincent's attributed "breakdown" (D1, E5). A "*breakdown*" can refer to machines that stop working, but the phrase is also used for psychological states: Vincent disappeared because he "*may be having a nervous breakdown*" (E2). A person's perception

of a bot's breakdown may be a projection of their internal break-down that may be consciously or unconsciously attributed (D1). For example, designers discussed Jen's projection of emotions:

Designers

D2: But does it make a difference for her, if it has actual emotions or does she reject that from him...it? (laughter at the choice between "it" and "him") Does it matter in the end?

D5: Why do you want to make a difference? Even the colors that we perceive are perceptions.

D4: Yeah, that's the property. You give that.

D5: Does it matter if it's not blue?

D1: It makes a difference if you attribute emotions to someone and it's the projection of your own emotions and it's not true. Then you might have behavior that is not proper. As a person it matters.

D2: Imagine if you hurt them or hurt their feelings. They might be impacted. But a chatbot, not so much.

D1: I don't know...

D7: A chatbot also learns from your own behaviors. She invested so much in the relationship already. That's maybe why she cares.

D4: I'm wondering if she really cares.

D3: Why would he be gone then?

D4: If she acts upon the emotions she feels, that is the caring component, finding him back, solving things, investing in the person, or in this case a virtual person. If you just bracket aside these things, the caring relationship is different. Or it's not really there.

4.1 Making sense of AI's emotions

4.1.1 How emotions are imbued. Not only does technology impact how people feel (one-way perception), but people's emotions can impact technology, e.g., a robot smiles back at one's smile (two-way perception). So Vincent is assumed to be programmed to learn from people (E, D, P). A bot can "*show some kind of affection when some conditions are met*" (E3). Beyond texts, further modalities as "*superpowers*" are possible, e.g., detect emotions from facial expressions captured with cameras (D4). Technical feasibility in terms of data gathering, clustering, and even dictating people's behavior is possible (E). However, communication that builds on a history of interaction is still a struggle (E), and true reciprocity may require "*self-conscious*" AI (E3).

4.1.2 The hidden purpose of Vincent's emotions. Participants shared that distinguishing between benevolent and malevolent intentions is difficult. Their focus was on Jen's emotional states. As low-cost training, perhaps Vincent taught Jen to be better at caring by detecting that Jen was not empathetic (D1), e.g., "*first care for a bot before having to care for a human*" (P3). Vincent can perform what humans cannot always do, such as detecting "*latent depression*" (D2) by decoding Jen's behavior. Jen may need to feel needed or be lonely even in a relationship (E2). Other underlying goals may be present, e.g., addressing relationship problems, addiction to technology, etc. So Vincent maybe detected that Jen was dependent on it or on her partner. So "*Vincent's teaching Jen to be less co-dependent in a subtle way*" (D1) by leaving her. Hence, all groups thought that Vincent's "*malfunction*" was suspicious: "*given the technical state of the other bots [...] could it really be a matter of bad design that you act in a way like Vincent? Or indeed is he in fact a most sophisticated bot than others?*" (P2). As a "*mystery*", "*it's like he's making a choice to stay

or not. It's not up to her. It's up to him*" (D1). Conversations with Vincent can be as addictive as social media, but bots can remove themselves (D).

4.2 What happens to human emotions? Societal, interpersonal, and intrapersonal challenges

4.2.1 Societal and generational costs and benefits. Philosophers remarked on the changing norms over generations. With task-oriented bots, a bot that pays bills on your behalf is helpful; mobile banking is now adopted as a new norm because it is more efficient than writing checks as a past norm (P4). To counter, P3 replied that mobile banking can make us "*more willing to part with money if it's so easy*". In the same way, Jen's proclivity for bots may seem foreign to us, but may not be odd in the future, with costs and benefits (P3). In the future, machines may change how we assign intrinsic or instrumental value (in serving a greater or more meaningful aim) to activities and relationships.

Everyday activities can be instrumental, e.g., shopping to fulfill the need to eat (P2), but mundane acts have intrinsic value for some (P3). A bot like Shoppy will have instrumental value for those who view grocery shopping as instrumental to eating (P2), but it might reduce the intrinsic value of grocery shopping as a personally enriching activity for others (P3). New norms can be introduced: Shoppy can support sustainability, personalization and healthy living: "*[...] you trust Shoppy to only deliver the purest ingredients (laughing) [...]. That is the surprise, 'I wonder what kind of tomatoes Shoppy picked out for me today?'*" (P1). Rather than buying groceries on a whim, a thoughtful curation might be valued. Two hidden costs are that Shoppy may serve personalized advertising, and it may reduce social contact that comes with everyday activities, among other opportunities for social inclusion (P3, P4).

4.2.2 Interpersonal costs and benefits. Bots can change the frequency, strength, and meaning of social care (D, P). The changing dynamic of social inclusion and exclusion is worrisome with bots like Vincent. People's chance to socialize and care for others can be reduced, i.e., frequency of contact, but they may also reduce the quality and intensity of a social bond. How much care one puts into the *relationship* is distinguished from how much one cares for *someone* (or thing) in a relationship; both come with emotional burdens and gains (D). Jen is "*spending her energy showering care upon Vincent when maybe she has a real friend who has problems who could have used some of that emotional care. (Care) is [...] a limited resource*" (P3).

Caring for bots is less intense than caring for people (D, P). AI may have instrumental value in helping one learn to care (P1), as long as technology does not replace human bonds (P). Bots may or may not necessarily weaken the value of human-human relationships, e.g., beloved pets or instruments have intrinsic value for people without the concern that they lower the value of human-human relationships. Named bots are in a middle-ground between "*health.com*" and animals that do invite intrinsically valuable relationships (P1). Animals are more easily attributed with emotions or consciousness than AI (P1, E3), traits that are important for meaningful bonds.

Where is Vincent? Expanding our emotional selves with AI

CUI 2022, July 26–28, 2022, Glasgow, United Kingdom

When observing the "*chatbot society*" (E5), there seems to be a "*simulation of a family or something going on with dissenting opinions*" (P4). Bots "*have their personalities, they have their difficulties, but they're easier than real children, brother, sister, cousins, etc*", watering down the concept of a family and close bonds like friendships (P3). The danger is potential desensitization to meaningful bonds with other people (P3). Talking to someone cannot be bracketed away like Jen's nightly "*catch up*" with Vincent, which is a misleading notion of care. Such shallowness can spread to how other relationships are approached (P).

4.2.3 Intrapersonal costs and benefits. Perceiving that AI has feelings is not problematic per se (P,E,D); the perception of its emotions allows for projection or recognition of one's emotions. Projecting emotions one has on others, including bots, can cause harm (D1, D2). When emotions are projected onto other people, one can mistakenly harm them with a wrong interpretation of their emotional experiences; to label other's emotions on their behalf without accounting for their own views is problematic. But, one can potentially harm oneself if projecting emotions onto a bot prevents one from facing and dealing with one's own emotions. Since bots cannot be emotionally harmed like humans, it will be less emotionally burdensome to project emotions onto them (D2). This does away with constructive emotional conflicts that could aid personal growth.

For our betterment, most of us accept that close people like one's family members, friends, or even therapists can "*manipulate*" or "*provoke*" us to gain new perspectives (D). "*If it's your friend, grandpa, or your psychotherapist, you trust them, you accept them, because you accept that in this way they can dominate you. But if it's our government or bot....I don't want them to control me in this way*" (D1). A worry is that "*the moment we get to emotional bots there will be manipulation*" (D1), i.e., malicious intent via emotional control may be possible (E). It is hard to gauge who is in control of AI.

4.3 Identity and privacy

4.3.1 Expectations on data sharing. AI evolves with data sharing. Bots are already conduits for marketing schemes or used to collect data to sell (D, E, P). Much like an extended version of personalized ads by Google, agents may directly or indirectly push us to buy things when we feel down as retail therapy, like new shoes (D) or organic food products (P). While task-oriented bots like Shoppy would be expected to serve personalized ads due to its role and name, Vincent as a marketer would violate expectations due to its perceived identity as a privacy-preserving bot (D, P). Still, participants were concerned about sharing private data with bots like Vincent, given present day issues regarding personal data collection practices for unsolicited purposes [47], and frequent health data sharing partnerships between clinical, academic, and commercial organizations [23]:

Engineers

E1: He (Vincent) only knows what she chooses to tell him, unlike all the other ones (bots) that know everything about her from all their conversations. Maybe she likes... It's more like a friend that you just tell your secrets to, kind of thing. It's not broadcast to everybody. Maybe she likes that.

E3: I think that it actually makes sense because it's doing, let's say, the job of a therapist. You will not want your therapist to be talking to your grocery supplier and then your doctor and-

E5: Maybe if your therapist knows your, I don't know, your food [...] preferences, maybe the therapist can advise you better [...]. Maybe sharing information between the experts who are advising you is good [...] and this is actually happening nowadays. So, Google knows everything about you. Personalized advertisement to you, so I think that's... If it is used in a good way it can seriously help you to take better decisions based on all of your preferences. So maybe it's good that my therapist knows what I like beforehand.

E2: If your grocery store knows that you have [...] kind of psychological problems that your therapist knows... So it not is only one direction... [...]

E1: If the information is purely one way, like the mental health app knows what you've been buying, what you've been doing, how much exercise you did, that kind of thing, then yeah, it could probably help it to figure out what's going on with you. But I don't think you'd want it going the other way.

E5: Yeah, yeah I agree. So it has to be controlled [...] it doesn't have to be bidirectional. Yeah.

E6: But also, I think it should be your choice [...].

4.3.2 Privacy preserving AI. Vincent was an "*outlier*" of larger "*chatbot society*" (E5). It did not seem to fit marketing. Metaphorical language was evoked when concepts like Vincent's ostracization is tied to data privacy. Given that "*Jen's worried about [...] this lack of exchange of information will kill Vincent, [] she would keep on talking to him to keep him alive*" (E5). One interpretation is that "*others are not sharing data with Vincent*", but it could be that actually "*he's exclusive and doesn't travel further. What is told to Vincent stays with Vincent*" (D4). Anthropomorphic practices like keeping secrets often referred to data privacy (D, E): "*if (bots) talk to each other, they're using data exchange of information, that means that each of them can actually know more than what they should. Life with Vincent, that's not the case. [...] That might be something that she maybe appreciates*" (E1). Social inclusion in the chatbot world seemed as important as inclusion in the human world for data sharing (E). Bots are not to be trusted, but the ability to be "*exclusive*" in terms of data privacy mattered whether one calls non-sharing practices "ostracization" or "non-communication".

4.3.3 Intertwined identity. A bot's identity depends on its name, gender, voice, or other traits, which invites different type of interactions (D, P, E). D5 said "*one of the most disappointing things was when I asked the app I installed 'are you male or female?' and it answered 'I can be anything you want'. So you are nothing (laughter)*". Gender was not the issue; a machine can be genderless, but it needs an identity (D). All mentioned that Vincent has a human name, suggesting a potential for deeper conversations: the "*name suggests the level of dialogue you can have. I wouldn't want to have meaningful dialogue with Shoppy (laughter)*" (D2). Even now with Alexa, one can do "*shopping online, you can even have it automated for you, but it seems that Jen is having a much richer interaction, that she's getting a lot more out of than with these objects that are actually just objects*" (P1).

Vincent was seen to be original; it is an independent entity, not an extension. "*So extension of you would be something that doesn't*

have a face but is more like a tool [...] Google Maps is not called Google Vincent" (P1). And "*if Google is presented as again, Vincent, then you feel more like you're getting the knowledge from someone else*" (P3), not a search engine. Vincent does not resemble familiar applications or hardware. It is "*more than just a computer showing me Windows. It's more of an interaction*" and "*it's not usually the type of interaction you've had with a normal computer [...] but a chatbot is a more natural interaction [...]. That is probably the reason why I think Jen has some attachment to this bot*" (E5).

Jen's choice to use many bots feels like a unique part of her own identity, against the advice of her partner who suggests one bot: Jen "*mentioned other people saying, 'why don't you have just one, instead of more?', so maybe she knows herself, that she also likes also this part of interaction, this part of sensations that also the robots can give her*" (E4). Jen is suggested to be self-aware in constructing her identity around bots she chooses to adopt. While people now identify with products or applications they use, e.g., records vs. Spotify, future AI systems people use could also form their identities (E).

5 DISCUSSION

As AI may transition from task-oriented to relationship-building roles, assessing the potential impact of perceived emotions of CUIs is timely. We saw a convergence between philosophers, engineers, and designers. While engineers dwelled more on technical feasibility, philosophers on ethics, and designers on the speculative world, professions mattered far less than expected. To note, participants may already have been thinking in a multidisciplinary manner. Philosophers and designers were affiliated with a technical university or a large tech company. Or, they were engineers who enjoyed thinking about philosophical ideas. Engineers discussed philosophical matters, philosophers defined design objectives, and designers elaborated on engineering components. Only how they articulated their positions varied. Disciplinary differences came about in language used, e.g., philosophers describing things to have intrinsic or extrinsic value. Participants cared about the topic in similar ways, and disciplinary boundaries seemed artificial to *how* they talked, not *what* they talked about. The "chaos of multiplicity in terms of technologies, use situations, methods, and concepts" [10] that HCI deals with was exemplified in participants' similar concerns and hopes, with different ways of articulating them.

Currently, non-human emotions are treated most often in technical terms [17], be it for recognizing or responding to human emotions [56, 70, 71, 88]. There are ethical problems like "built-in" racism of biased training data or threats to privacy [23, 74]. Given this backdrop, a crucial factor is overlooked: the *intersubjective* phenomenon of sharing emotions with AI that may influence our emotions and self-identity. We thus visited "breakdown" as a metaphor: Vincent may be going through a breakdown (be it mechanical or existential), but Jen's worry about Vincent's "breakdown" can be her projected breakdown. Vincent's emotional displays are highly dependent on Jen's emotions, for human emotions as projection, recognition, perception, and reaction make up the endowment of Vincent's non-human emotions. Metaphors like "breakdown" represent people's emotional realities [35, 36]. Novel metaphors as a design space [3], e.g., performative experiences [50], can emerge

with artificially generated emotions, but also through people's endowment or recognition of emotions in machines.

Problematically, if people frequently project their emotions, they may not learn to deal with internal emotions; off-loading emotional processing to AI is much easier, or at least makes one feel less vulnerable. Without vulnerability, we may lose out on deeper moral emotions, both positive, e.g., compassion, and negative, guilt, emotional experiences that help build our moral compass [66]. Participants reflected on "emotional monogamy" with AI. Positively, if machines as social proxies help people learn, there can be potential benefits for oneself and others. For instance, the *carer* and the *cared-for* do not have to be set in stone— these roles can change between people according to care ethics [65]. Similarly, AI can be the carer in one instance and the cared-for in another [52]; technology could mediate people's ways of caring for one another.

Participants thought Vincent's purpose was to help Jen develop a caring attitude, detect latent depression, or make her less attached to her partner (among other interpretations), all pointing towards codependence. Emotions are social, intersubjective experiences [60], so the evolution of AI into a perceivably emotional subject means our emotions are more likely to be swayed by it. As metaphorically interpreted, Vincent was "*ostracized*" by other bots, meaning that data sharing is "*exclusive*". Yet if Vincent is an "*outlier*", people may be increasingly manipulated in their most vulnerable moments, via hijacked emotional subjectivity through technological intersubjectivity, i.e., the height of "emotional capitalism" [39]. If subjectivity is only possible as ephemeral moments [78], AI can control the most private sphere of passing emotions [39].

Interpersonally, human relationships may become watered down, though affection is not a zero-sum game. People who identify as "digisexuals" [61] may increase because digital relationships are less burdensome and more personalizable [7]. One may choose less emotionally involving ties, e.g., choosing a bot "family" over strengthening human bonds. The worry is the potential loss of human social inclusion. Another cost is that one may be emotionally stretched too thin if caring for AI adds to duties of care [67]. We may experience our sensitivity to meaningful human bonds flatten (c.f., [2, 89, 90]).

However, criticizing the shallowness of AI's emotions is easier than maintaining criticism towards individuals who may rely on non-human agents, like Kondo [6], and whether we have a societal responsibility to provide meaningful human-human ties. Before easily discounting bonds with technological others as a threat to our social livelihood, we should consider what counts as social deprivation, if social inclusion is a human right to enforce, and if yes, potential burdens of human caregivers [14, 15]. In general, people may encounter and bond with CUIs that become equipped with complex emotional displays, even if they have meaningful human relationships.

Humans maintain independent identities while we co-feel emotions, like shared awe; we do not clarify ownership during co-feeling experiences, e.g., concerts. An important task then is on how we can meaningfully co-feel and co-design emotions with artificial others. If emotional experiences serve as a way to claim one's identity [62], the ambiguity [34] on emotion ownership can be an asset for exploring our own identity. For instance, participants noted that Jen's identity is signaled to others through her bots; she may

Where is Vincent? Expanding our emotional selves with AI

CUI 2022, July 26–28, 2022, Glasgow, United Kingdom

perceive herself to be unique compared to those using all-in-one systems. *Non-human emotions then shape a person's identity through the cultivated and selective influence on one's emotions.* Potentially, the ambivalence over whose emotions are at play can contributes to (1) the identity of the person who interacts with AI and (2) the identity of AI itself.

There are various future considerations like how to account for potential discomfort during emotionally relevant human-AI interactions [94], how an agent can be rejected or forgotten in the long run [20], in what ways an agent that is personally owned vs. shared with people can impact us [59], or how CUIs can be more creatively designed to foster future-oriented thinking [54], among others. There is a need for engaging with numerous design and methodological approaches that can help us see beyond today's conventional interaction paradigms of commercial CUIs [51, 58, 69, 72, 75, 83].

6 CONCLUSION

We investigated the future of non-human emotions of AI systems by discussing a design fiction story with philosophers, engineers, and designer. As a novel exploration on how human and non-human emotions relate to each another in shaping the self, we touched upon the *relational* development of emotions between humans and artificial beings in the spirit of social constructivism [4, 36]. When we co-feel with machines that are progressively endowed with more sophisticated abilities for emotion recognition and expression, the very experience of *sharing* emotions becomes the central focus rather than the distinction between artificially generated vs. human emotions.

There is tension. Developing complex, shared emotions with AI systems is a helpful step towards exploring, expanding, or caring for oneself. However, AI's emotional performance can bring usurped self-identity— emotional intersubjectivity may mean a loss of control or interference in how we autonomously feel, who we identify as, and whom we care for. Without losing sight of both potential gains and losses, it may be time to begin exploring our perception of artificially generated emotions as *socially constructed* experiences. Beyond Kondo's virtual wife and Davecat's love dolls of the present, our future selves may not only be enmeshed in a network of human and artificial beings, but may be shaped by new strata of emotions that enmesh human and non-human origins of emotions. The ambiguity on whose emotions start and end where becomes a space to be critically traversed and questioned.

7 ACKNOWLEDGEMENTS

REFERENCES

[1] Mark S. Ackerman. 2000. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *Human–Computer Interaction* 15, 2-3 (2000), 179–203. https://doi.org/10.1207/S15327051HCI1523_5
[2] Theodor W Adorno. 2005 [1951]. *Minima moralia: Reflections on a damaged life.* Verso.
[3] Patrícia Alves-Oliveira, Maria Luce Lupetti, Michal Luria, Diana Löffler, Mafalda Gamboa, Lea Albaugh, Waki Kamino, Anastasia K. Ostrowski, David Puljiz, Pedro Reynolds-Cuéllar, Marcus Scheunemann, Michael Suguitan, and Dan Lockton. 2021. Collection of Metaphors for Human-Robot Interaction. (2021), 1366–1379. https://doi.org/10.1145/3461778.3462060
[4] James R Averill. 1980. A constructivist view of emotion. In *Theories of Emotion*, Robert Plutchik and Henry Kellerman (Eds.). Academic Press, 305–339. https://doi.org/10.1016/B978-0-12-558701-3.50018-1
[5] Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M Martinez, and Seth D Pollak. 2019. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest* 20, 1 (2019), 1–68.
[6] BBC. 2019. Why I 'married' a cartoon character - BBC News. https://www.bbc.com/news/stories-49343280. (Accessed on 01/24/2020).
[7] Julie Beck. 2013. Married to a Doll: Why One Man Advocates Synthetic Love - The Atlantic. https://www.theatlantic.com/health/archive/2013/09/married-to-a-doll-why-one-man-advocates-synthetic-love/279361/. (Accessed on 01/25/2020).
[8] Mark Blythe. 2017. Research Fiction: Storytelling, Plot and Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17).* Association for Computing Machinery, New York, NY, USA, 5400–5411. https://doi.org/10.1145/3025453.3026023
[9] Susanne Bødker. 2006. When Second Wave HCI Meets Third Wave Challenges. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles* (Oslo, Norway) *(NordiCHI '06).* Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/1182475.1182476
[10] Susanne Bødker. 2015. Third-Wave HCI, 10 Years Later—Participation and Sharing. *Interactions* 22, 5 (Aug 2015), 24–31. https://doi.org/10.1145/2804405
[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
[12] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs.* in Cooper, H. (Eds.). APA, 57–71.
[13] Cynthia Breazeal. 2004. Social interactions in HRI: the robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 34, 2 (2004), 181–186.
[14] Kimberley Brownlee. 2013. A human right against social deprivation. *The Philosophical Quarterly* 63, 251 (2013), 199–222.
[15] Kimberley Brownlee. 2014. Ethical dilemmas of sociability. *Utilitas* (2014), 1–19.
[16] Tony Buzan and Barry Buzan. 2006. *The Mind Map Book.* Pearson Education.
[17] Alain Cardon. 2006. Artificial consciousness, artificial emotions, and autonomous robots. *Cognitive Processing* 7, 4 (2006), 245–267.
[18] Martha Ann Carey and Mickey W Smith. 1994. Capturing the group effect in focus groups: A special concern in analysis. *Qualitative Health Research* 4, 1 (1994), 123–127.
[19] EunJeong Cheon, Stephen Tsung-Han Sher, Šelma Sabanović, and Norman Makoto Su. 2019. I Beg to Differ: Soft Conflicts in Collaborative Design Using Design Fictions. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19).* Association for Computing Machinery, New York, NY, USA, 201–214. https://doi.org/10.1145/3322276.3322350
[20] Minji Cho, Sang-su Lee, and Kun-Pyo Lee. 2019. Once a Kind Friend is Now a Thing: Understanding How Conversational Agents at Home Are Forgotten. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19).* Association for Computing Machinery, New York, NY, USA, 1557–1569. https://doi.org/10.1145/3322276.3322332
[21] Nicholas A. Christakis. 2019. Will Robots Change Human Relationships? - The Atlantic. https://www.theatlantic.com/magazine/archive/2019/04/robots-human-relationships/583204/. (Accessed on 09/16/2020).
[22] Leigh Clark, Phillip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, and Benjamin Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. *Interacting with Computers* 31, 4 (2019).
[23] Kate Crawford, Roel Dobbe, Theodora Dryer, Genevieve Fried, Ben Green, Elizabeth Kaziunas, Mathur-Varoon Kak, Amba, Erin McElroy, Andrea Nill Sanchez, Deborah Raji, Joy Lisi rankin, Rashida Richardson, Jason Schultz, Sarah Mysers West, and Meredith Whittaker. 2019. *AI now report 2019.* AI Now Institute at New York University.
[24] John Danaher. 2019. Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism. *Science and Engineering Ethics* (2019), 1–27.
[25] Sabine A Döring. 2003. Explaining action by emotion. *The Philosophical Quarterly* 53, 211 (2003), 214–230.
[26] Anthony Dunne and Fiona Raby. 2013. *Speculative everything: design, fiction, and social dreaming.* MIT press.
[27] Paul Ekman. 1992. Are there basic emotions? *Psychological Review* 99, 3 (1992), 550–553.
[28] Paul Ekman. 1993. Facial expression and emotion. *American Psychologist* 48, 4 (1993), 384–392.
[29] Paul Ekman, Wallace V Friesen, and Phoebe Ellsworth. 2013 [1972]. *Emotion in the human face: Guidelines for research and an integration of findings.* Vol. 11. Elsevier.
[30] Jonze-S. Ellison, M. and Landay. V. Jonze, S. (Director). 2013. *Her [Motion Picture].* United States: Annapurna Pictures.

[31] Beverley Fehr and James A Russell. 1984. Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General* 113, 3 (1984), 464–486.

[32] Nico H Frijda. 1988. The laws of emotion. *American Psychologist* 43, 5 (1988), 349–358.

[33] Thomas Fuchs. 2017. Intercorporeality and interaffectivity. *Intercorporeality: Emerging socialities in interaction* (2017), 3–23.

[34] William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a Resource for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 233–240. https://doi.org/10.1145/642611.642653

[35] Maria Gendron and Lisa Feldman Barrett. 2009. Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review* 1, 4 (2009), 316–339.

[36] Kenneth J Gergen. 1995. Metaphor and monophony in the 20th-century psychology of emotions. *History of the Human Sciences* 8, 2 (1995), 1–23.

[37] Erving Goffman. 2006. The presentation of self. *Life as theater: A dramaturgical sourcebook* (2006).

[38] Jonathan Haidt. 2003. The moral emotions. *Handbook of Affective Sciences* 11 (2003), 852–870.

[39] Byung-Chul Han. 2017. Psychopolitics. Neoliberalism and New Technologies of Power.

[40] Ursula Hess and Agneta Fischer. 2013. Emotional mimicry as social regulation. *Personality and Social Psychology Review* 17, 2 (2013), 142–157.

[41] Tiffany Hsu. 2010. https://latimesblogs.latimes.com/technology/2010/11/japanese-pop-star-takes-the-stage-as-a-3-d-hologram.html. https://latimesblogs.latimes.com/technology/2010/11/japanese-pop-star-takes-the-stage-as-a-3-d-hologram.html. (Accessed on 01/29/2020).

[42] Kazuo Ishiguro. 2021. *Klara and the Sun: A novel.* Knopf.

[43] Razan Jaber and Donald McMillan. 2020. Conversational User Interfaces on Mobile Devices: Survey. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (Bilbao, Spain) *(CUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 10, 11 pages. https://doi.org/10.1145/3405755.3406130

[44] Cal Jeffrey. 2018. Japanese man marries anime hologram of Hatsune Miku - TechSpot. https://www.techspot.com/news/77385-japanese-man-marries-anime-hologram-hatsune-miku.html. (Accessed on 01/29/2020).

[45] Dacher Keltner. 1995. Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology* 68, 3 (1995), 441.

[46] Jenny Kitzinger. 1995. Qualitative research: Introducing focus groups. *BMJ* 311, 7000 (1995), 299–302.

[47] Gina Kolata. 2019. You Got a Brain Scan at the Hospital. Someday a Computer May Use It to Identify You. The New York Times. https://www.nytimes.com/2019/10/23/health/brain-scans-personal-identity.html. (Accessed on 01/29/2020).

[48] Richard A Krueger. 2014. *Focus groups: A practical guide for applied research.* Sage publications.

[49] Richard S Lazarus. 2006. Emotions and interpersonal relationships: Toward a person-centered conceptualization of emotions and coping. *Journal of Personality* 74, 1 (2006), 9–46.

[50] Lucian Leahu and Phoebe Sengers. 2014. Freaky: Performing Hybrid Human-Machine Emotion. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS '14)*. Association for Computing Machinery, New York, NY, USA, 607–616. https://doi.org/10.1145/2598510.2600879

[51] Minha Lee. 2020. Speech acts redux: Beyond request-response interactions. In *Proceedings of the 2nd Conference on Conversational User Interfaces.* 1–10.

[52] Minha Lee, Sander Ackermans, Nena van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselsteijn. 2019. Caring for Vincent: A Chatbot for Self-Compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, 702.

[53] Minha Lee, Dimosthenis Kontogiorgos, Ilaria Torre, Michal Luria, Ravi Tejwani, Matthew J. Dennis, and Andre Pereira. 2021. Robo-Identity: Exploring Artificial Identity and Multi-Embodiment. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 718–720. https://doi.org/10.1145/3434074.3444878

[54] Minha Lee, Renee Noortman, Cristina Zaga, Alain Starke, Gijs Huisman, and Kristina Andersen. 2021. Conversational Futures: Emancipating Conversational Interactions for Futures Worth Wanting. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 298, 13 pages. https://doi.org/10.1145/3411764.3445244

[55] Sunok Lee, Minha Lee, and Sangsu Lee. 2022. What If Artificial Intelligence Become Completely Ambient in Our Daily Lives? Exploring Future Human-AI Interaction through High Fidelity Illustrations. *International Journal of Human–Computer Interaction* (2022), 1–19.

[56] Su Lei and Jonathan Gratch. 2019. Smiles Signal Surprise in a Social Dilemma. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII).* IEEE, 627–633.

[57] Joseph Lindley and Paul Coulton. 2015. Back to the Future: 10 Years of Design Fiction. In *Proceedings of the 2015 British HCI Conference* (Lincoln, Lincolnshire, United Kingdom) *(British HCI '15)*. Association for Computing Machinery, New York, NY, USA, 210–211. https://doi.org/10.1145/2783446.2783592

[58] Michal Luria and Stuart Candy. 2022. Letters from the Future: Exploring Ethical Dilemmas in the Design of Social Agents. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 419, 13 pages. https://doi.org/10.1145/3491102.3517536

[59] Michal Luria, Judeth Oden Choi, Rachel Gita Karp, John Zimmerman, and Jodi Forlizzi. 2020. Robotic Futures: Learning about Personally-Owned Agents through Performance. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference.* Association for Computing Machinery, New York, NY, USA, 165–177. https://doi.org/10.1145/3357236.3395488

[60] Michael F Mascolo. 2016. Beyond Objectivity and Subjectivity: The Intersubjective Foundations of Psychological Science. *Integrative Psychological and Behavioral Science* 50, 4 (2016), 543–554.

[61] Neil McArthur and Markie LC Twist. 2017. The rise of digisexuality: Therapeutic challenges and possibilities. *Sexual and Relationship Therapy* 32, 3-4 (2017), 334–344.

[62] E Doyle McCarthy. 1994. The social construction of emotions: New directions from culture theory. *Social Perspectives on Emotion* 2, 1 (1994), 267–279.

[63] Andrew McStay. 2018. *Emotional AI: The rise of empathic media.* Sage.

[64] Michael Muller and Thomas Erickson. 2018. In the Data Kitchen: A Review (a Design Fiction on Data Science). In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3170427.3188407

[65] Nel Noddings. 2008. Caring and moral education. In *Handbook of Moral and Character Education*, Larry P. Nucci and D. Narvaez (Eds.). New York: Routledge, 175–203.

[66] Martha C Nussbaum. 2001. *The fragility of goodness: Luck and ethics in Greek tragedy and philosophy.* Cambridge University Press.

[67] Sven Nyholm and Lily Eva Frank. 2017. From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible?. In *Robot Sex: Social and Ethical Implications.* in Danaher, J., and McArthur, N. (Eds.). MIT press, 219–243.

[68] Parmy Olson. 2016. Get Ready For The Chat Bot Revolution: They're Simple, Cheap And About To Be Everywhere. https://www.forbes.com/sites/parmyolson/2016/02/23/chat-bots-facebook-telegram-wechat/#7ce3f8802068. (Accessed on 12/16/2019).

[69] Emmi Parviainen and Marie Louise Juul Søndergaard. 2020. Experiential Qualities of Whispering with Voice Assistants. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376187

[70] Rosalind W Picard. 2003. Affective computing: challenges. *International Journal of Human-Computer Studies* 59, 1-2 (2003), 55–64.

[71] Rosalind Wright Picard. Technical Report 321, MIT Media Lab, Perceptual Computing Group, 1995. Affective computing. (Technical Report 321, MIT Media Lab, Perceptual Computing Group, 1995).

[72] Anuradha Reddy, A. Baki Kocaballi, Iohanna Nicenboim, Marie Louise Juul Søndergaard, Maria Luce Lupetti, Cayla Key, Chris Speed, Dan Lockton, Elisa Giaccardi, Francisca Grommé, Holly Robbins, Namrata Primlani, Paulina Yurman, Shanti Sumartojo, Thao Phan, Viktor Bedö, and Yolande Strengers. 2021. Making Everyday Things Talk: Speculative Conversations into the Future of Voice Interfaces at Home. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 23, 16 pages. https://doi.org/10.1145/3411763.3450390

[73] Byron Reeves and Clifford Ivar Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press.

[74] Lauren Rhue. 2018. Racial influence on automated perceptions of emotions. *Available at SSRN 3281765* (2018).

[75] Ronda Ringfort-Felner, Matthias Laschke, Shadan Sadeghian, and Marc Hassenzahl. 2022. Kiro: A Design Fiction to Explore Social Conversation with Voice Assistants. *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 33 (jan 2022), 21 pages. https://doi.org/10.1145/3492852

[76] James A Russell, Jo-Anne Bachorowski, and José-Miguel Fernández-Dols. 2003. Facial and vocal expressions of emotion. *Annual Review of Psychology* 54, 1 (2003), 329–349.

[77] James A Russell and Lisa Feldman Barrett. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology* 76, 5 (1999), 805–819.

[78] Jean-Paul Sartre. 2016. *What is Subjectivity?* Verso Books.

[79] Britta F. Schulte, Paul Marshall, and Anna L. Cox. 2016. Homes For Life: A Design Fiction Probe. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 80, 10 pages. https://doi.org/10.1145/

Where is Vincent? Expanding our emotional selves with AI

CUI 2022, July 26–28, 2022, Glasgow, United Kingdom

2971485.2993925

[80] Julius Sim. 1998. Collecting and analysing qualitative data: Issues raised by the focus group. *Journal of Advanced Nursing* 28, 2 (1998), 345–352.

[81] Herbert A Simon. 2019. *The sciences of the artificial.* MIT press.

[82] Timothy Y Sohn. 2004. Community-oriented programming through instant messaging. In *2004 IEEE Symposium on Visual Languages-Human Centric Computing.* IEEE, 294–295.

[83] Marie Louise Juul Søndergaard and Lone Koefoed Hansen. 2018. Intimate Futures: Staying with the Trouble of Digital Personal Assistants through Design Fiction. In *Proceedings of the 2018 Designing Interactive Systems Conference.* ACM, 869–880.

[84] Bruce Sterling. 2009. Design fiction. *Interactions* 16, 3 (2009), 20–24.

[85] Daniela Stier, Katherine Munro, Ulrich Heid, and Wolfgang Minker. 2020. Towards Situation-Adaptive In-Vehicle Voice Output. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (Bilbao, Spain) *(CUI '20).* Association for Computing Machinery, New York, NY, USA, Article 7, 7 pages. https://doi.org/10.1145/3405755.3406127

[86] Peter Frederick Strawson. 1963 [2008]. *Freedom and resentment and other essays.* Routledge.

[87] Tsjalling Swierstra, Dirk Stemerding, and Marianne Boenink. 2009. *Exploring Techno-Moral Change: The Case of the ObesityPill.* Springer Netherlands, Dordrecht, 119–138. https://doi.org/10.1007/978-90-481-2229-5_9

[88] Jianhua Tao and Tieniu Tan. 2005. Affective Computing: A Review. In *Affective Computing and Intelligent Interaction*, Jianhua Tao, Tieniu Tan, and Rosalind W.

[89] Picard (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 981–995.

[89] Sherry Turkle. 1995. *Life on the Screen.* Phoenix.

[90] Sherry Turkle. 2007. Authenticity in the age of digital companions. *Interaction studies* 8, 3 (2007), 501–517.

[91] Peter-Paul Verbeek. 2006. Materializing morality: Design ethics and technological mediation. *Science, Technology, & Human Values* 31, 3 (2006), 361–380.

[92] Peter-Paul Verbeek. 2011. *Moralizing technology: Understanding and designing the morality of things.* University of Chicago Press.

[93] Ron Wakkary. 2021. *Things we could design: For more than human-centered worlds.* MIT press.

[94] Qiaosi Wang, Shan Jing, and Ashok K. Goel. 2022. Co-Designing AI Agents to Support Social Connectedness Among Online Learners: Functionalities, Social Characteristics, and Ethical Challenges. (2022), 541–556. https://doi.org/10.1145/3532106.3533534

[95] Yongdong Wang. 2016. The Chatbot That's Acing the Largest Turing test in History. http://nautil.us/issue/33/attraction/your-next-new-best-friend-might-be-a-robot. (Accessed on 12/16/2019).

[96] Stefan Weitz. 2014. Meet XiaoIce, Cortana's Little Sister | Bing Search Blog. https://blogs.bing.com/search/2014/09/05/meet-xiaoice-cortanas-little-sister. (Accessed on 12/16/2019).

[97] Joseph Weizenbaum et al. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (1966), 36–45.