Speculating on Risks of AI Clones to Selfhood and Relationships: Doppelganger-phobia, Identity Fragmentation, and Living Memories

PATRICK YUNG KANG LEE, University of British Columbia, Canada
NING F. MA, University of British Columbia, Canada
IG-JAE KIM, Korea Institute of Science and Technology, Korea, Republic of
DONGWOOK YOON, University of British Columbia, Canada

Digitally replicating the appearance and behaviour of individuals is becoming feasible with recent advancements in deep-learning technologies such as interactive deepfake applications, voice conversion, and virtual actors. Interactive applications of such agents, termed *AI clones*, pose risks related to impression management, identity abuse, and unhealthy dependencies. Identifying concerns AI clones will generate is a prerequisite to establishing the basis of discourse around how this technology will impact a source individual's selfhood and interpersonal relationships. We presented 20 participants of diverse ages and backgrounds with 8 speculative scenarios to explore their perception towards the concept of AI clones. We found that (1. doppelganger-phobia) the abusive potential of AI clones to exploit and displace the identity of an individual elicits negative emotional reactions; (2. identity fragmentation) creating replicas of a living individual threatens their cohesive self-perception and unique individuality; and (3. living memories) interacting with a clone of someone with whom the user has an existing relationship poses risks of misrepresenting the individual or developing over-attachment to the clone. These findings provide an avenue to discuss preliminary ethical implications, respect for identity and authenticity, and design recommendations for creating AI clones.

CCS Concepts: • **Social and professional topics** → **Socio-technical systems**; • **Human-centered computing** → **Empirical studies in HCI**; • **Computing methodologies** → *Philosophical/theoretical foundations of artificial intelligence.*

Additional Key Words and Phrases: AI clones, identity, interpersonal relationship, impression management, self-hood, AI agents, machine learning applications, human-AI interaction, human-centered AI, doppelganger-phobia, identify fragmentation, living memories, risks

ACM Reference Format:
Patrick Yung Kang Lee, Ning F. Ma, Ig-Jae Kim, and Dongwook Yoon. 2023. Speculating on Risks of AI Clones to Selfhood and Relationships: Doppelganger-phobia, Identity Fragmentation, and Living Memories. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 91 (January 2023), 28 pages. https://doi.org/10.1145/3579524

1 INTRODUCTION

In February 2020, one of South Korea's largest media broadcasters released a documentary titled *Meeting You* [44], depicting a reunion in virtual reality (VR) between a mother and her recently deceased 7-year old daughter, who had been "digitally resurrected" in the form of an interactive

Authors' addresses: Patrick Yung Kang Lee, pl0419@student.ubc.ca, University of British Columbia, Canada; Ning F. Ma, ning.ma@ubc.ca, University of British Columbia, Canada; Ig-Jae Kim, drjay@kist.re.kr, Korea Institute of Science and Technology, Korea, Republic of; Dongwook Yoon, yoon@cs.ubc.ca, University of British Columbia, Canada.

2573-0142/2023/1-ART91 $15.00
https://doi.org/10.1145/3579524

simulacrum [45]. The mother was given the opportunity to say her last goodbyes to her daughter, who had passed away suddenly due to blood cancer, over a tearful interactive picnic in VR. In June 2022, the FBI's Internet Crime Complaint Center issued a PSA regarding the use of deepfakes by impersonators to apply for work-from-home and remote work positions. The PSA noted these deepfakes included videos, images, or recordings "convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said" [1]. Both these stories made it clear that digitally replicating the appearance and behaviour of a specific individual is no longer the realm of science-fiction but an increasingly feasible technological advancement.

This feasibility is a result of recent and ever-growing advancements in deep-learning technologies, such as speech synthesis [12], voice conversion technologies [49, 59], deepfakes [82], natural language processing (NLP) [66], and virtual actors [2, 30]. Along these lines, applications involving the embodiment of the likeness of real-world individuals are progressing rapidly. A few examples in this growing design space include interactive digital twins of the deceased [44] and self [75] in virtual reality, the creation of new interactive forms of media using the personal data of loved ones [62], and chatbots that mimic the conversational mannerisms of those we know [69] and the past selves [39]. These types of digital agents may then be used to create interactive experiences – doing work for us, serving as social companions, or even as romantic partners. We refer to these systems that attempt to replicate the likeness and behaviour of real-world individuals for use in interactive experiences as *AI clones*.

Generating digital clones of *laypeople* is becoming increasingly possible with continuous technological advancements, creating new social and ethical concerns. The creation of AI clones of celebrities has proliferated in the entertainment industry over the last couple of decades, from virtual versions of actors in the film-making industry [41] to interactive copies of deceased musicians [37, 54]. Work in how to legally navigate issues of impression management [9] and personality rights [70] have been at the forefront of this application domain of celebrity AI clones. Unlike celebrity figures, whose identity management issues are often navigated with the assistance of managers and talent agencies, laypeople often mismanage their impressions [72] or have unique risks in presenting contextually coherent presentations of their identity [53].

The ability to replicate a real-world individual therefore comes with pressing concerns regarding identity and relationship management. Technologies related to AI clones, like deepfakes, have been featured in a number of alarming headlines [36] and previous literature on identity abuse [51] and identity theft [84] threats. Studies on political misinformation [80] and non-consensual deepfake pornography [3] have begun to explore some of the fears about AI clones that have gripped both the general public and lawmakers. Similarly, understanding people's perceptions of AI clones can generate implications for guiding policy, design, and technical decisions to mitigate potential threats to personal identity and relationships posed by this class of technologies.

Given these considerations, our study was guided by the research question, *what perceptions and concerns do users and source individuals*[1] *have regarding the construction and use of AI clones?* In HCI and CSCW, two lines of conceptual framing are particularly relevant to this question.

1) *Personal identity and selfhood.* What effects will being embodied in an AI clone produce on the personal identities of people whose likenesses are being replicated? Previous literature [81] has shown users to be uncomfortable with seeing their personality and likeness being embodied in interactive applications. With the advent of AI clones of even-increasing fidelity, issues of impression management [38, 48, 76, 88, 89] and uncanny valley perceptions [60] are problems that will have to be resolved to successfully develop applications that use AI clones.

[1]Hereafter, "users" refers to those who interact with AI clones, and "source individuals" indicates those who are replicated by AI clones.

2) *Interpersonal relationships.* How will interactions with AI clones affect our interpersonal relationships? Previous literature on non-consensual synthetic pornography has brought to light the devastating effects on interpersonal relationships that can be caused when people are no longer in control of their digital identities [71]. No matter the application domain, having interactive experiences with AI clones is likely to have ripple effects between users and the person on whom the clone is based. Identifying possible mechanisms that can affect these interpersonal relationships is a key design concern that requires further study.

To this end, we conducted a qualitative study that involved having 20 members of the general public read and respond to 8 speculative scenarios involving AI clone usage to elicit their perceptions around AI clones. Responses from our participants were contextualized with references to the speculative scenarios envisioned for AI clone usage, such as grief processing, sexual and romantic intimacy, and family dynamics.

Analysis of our data revealed how people's concerns about the AI clones depicted in the scenarios were shaped by three factors: emotional aversion to potential replacement and exploitation of their identity, threats to cohesive perception of self identity, and risks in interpersonal relationships produced by misrepresentation and over-reliance. Based on these findings, we discuss some ways our findings connect to a larger conversation in ethical AI development, outline an identity-centric philosophy to designing AI clones with a focus on maintaining an individual's authenticity, and briefly discuss the rights and responsibilities of different stakeholder groups involved in the development, deployment, and use of AI clones. This study contributes:

- a conceptual contribution for establishing and characterizing the notion of AI clones,
- empirical findings on people's concerns regarding the potential impact of AI clones on personal identity and relationships, characterized in terms of issues with doppelganger-phobia, identity fragmentation, and living memories,
- discussions on connections with ethical AI, identity-centric design guidelines, policy recommendations, and a call for greater academic discourse about this nascent design space.

2 AI CLONES: CONCEPTUALIZATION AND RELATED WORK

2.1 Conceptualizing AI Clones

Impersonating real-world individuals has been a documented phenomenon. Celebrity impersonators and stunt doubles have been a staple of the entertainment industry; identity theft and identity fraud by impostors has long been a criminal activity; the use of political decoys, such as kagemushas, has also been previously documented [52]. The extension of impersonation into the digital realm, however, has been a relatively new and novel phenomenon.

Despite this growing field [57, 58], the characterization of these digital "copies" of existing individuals has been inconsistent in the literature. Floridi [21] has suggested the term *ectype*, as a "copy, yet not any copy, but rather a copy that has a special relation with its source (the origin of its creation), the archetype". This is a suitably abstract definition that applies to AI clones, but is lacking in terms of a pragmatic description for a socio-technical system. The term *deepfake* has entered the popular consciousness, usually in reference to the digital manipulation of the visual likeness of an individual for malicious intent [82]. AI clones, however, are not *a priori* meant to be malicious, and show potential for positive social change, as we elaborate in our discussion. Badler et al. [6] use the term "digital humans" to refer to a broad class of technologies that mimic human likenesses, to the point where they undertake regular human-like routines and have their own schedules, but this also departs from our more specific conception of AI clones.

Truby & Brown [79] use the term "digital thought clone" to refer to digital artifacts that replicate the behaviour and decision-making patterns of consumers with near identical precision to the

source individual. However, the primary usage for digital thought clones is as tools for performing consumer market research rather than crafting interactive user experiences. A US patent proposing "a digital virtual clone of a user that learns about the user and effectively functions on behalf of the user" [42] is also suitably close to our conception of AI clones. Given the lack of consistency in terminology for designating the digital agents we explore, we borrow and expand upon Truby & Brown's terminology and refer to these digital representations of individuals as *"AI clones."* We ground this term in three defining features we take to be representative of AI clones.

2.1.1 AI clones are taken to represent some or multiple aspects of a real-world individual. AI clones are distinct from other digital representations of human beings, in that they are based upon, and taken to represent, existing individuals. They serve in essence as representatives of the real-world *source individual*, the individual on whom the likeness of the AI clone is based. This stands in contrast to the current landscape of digital representations of humans, which are largely based on fictitious humans or characters. Whereas the same technologies used to create AI clones could also be used to create fictitious human-like digital agents, it is when we are basing such digital agents on real-world individuals that there seems to be a particular cause for concern, as studies on deepfakes have shown [51]. In addition, AI clones need not be a perfect replica of the real-world source individual. Some AI clones may only try to replicate certain aspects of the source individual's personality or mannerisms. Any combination of such elements, such as the source individual's visual likeness, conversational mannerisms, or behavioural patterns, to name a few relevant dimensions, could potentially constitute an AI clone.

2.1.2 AI clones are interactive technologies. AI clones adhere to basic user interaction loops. They present interfaces for interpreting user and environmental input, conduct internal processing, and produce perceptible output. They can take on a variety of representational mediums, from relatively simple text-based chatbots [66] to immersive digital experiences in VR [24] incorporating multi-modal feedback loops. The criterion laid out above precludes non-interactive technologies, such as a compiled highlight reel of an individual's social media posts, from being considered an AI clone.

2.1.3 AI clones are AI-based technologies built on personal data. The degree of artificial intelligence present in AI clones can be of varying degrees of complexity. They may range from simple pre-programmed behavioural responses to more complex models involving large datasets and machine learning technologies in order to mimic human-like behaviour and mannerisms [15]. As individuals generate increasingly large amounts of data throughout their lifetimes, the fidelity of AI clones capable of replicating their behaviour is also likely to increase.

Although the design space of such AI clones is relatively new, digital recreations of real-world individuals, particularly celebrities, has been previously documented [74]. This is likely due to the financial and social opportunities that creating such AI clones gives rise to. Celebrity figures who have been "digitized" have served a variety of purposes, from delivering public service announcements [74] to musical performances [37]. Celebrity figures also have access to agents and lawyers that are actively performing impression management and considering the implications of their image in relation to an audience. Laypeople, on the other hand, are much less likely to have active support in considering the implications of being digitized into AI clones. For these reasons, this study is focused primarily on the creation and use of AI clones based on laypeople and individuals most participants interact with on a daily basis, rather than unique individuals that occupy distinct social niches, such as celebrity figures or politicians.

2.2 Related Research on AI Clones

Interactive socio-technical systems like Apple's Siri or Amazon's Alexa that enact human-like traits are already becoming commonplace consumer-facing technologies. However, as mentioned in 2.1.1, it is becoming increasingly feasible for developers to use these underlying technologies to simulate real-world individuals for interactive applications. This study hopes to fill a gap in imagining the potential consequences and futures that developing such AI clones will have on potential users and source individuals. To this extent, we have surveyed existing literature about technologies similar to AI clones, and hope to extend the literature on issues of risks presented by engaging with human-like digital agents, digital self presentation/impression management, and how management of personal data of the deceased can affect interpersonal relationships to anticipate future consequences and values surrounding the development of AI clones.

2.2.1 Engagement with human-like digital agents. Human-like digital agents across a range of mediums are increasingly becoming user-facing, from audio-based virtual assistants like Siri and Alexa to automated chatbot services used in customer service applications [22]. How AI clones, being replicas of real individuals, and not merely human-like digital agents, might alter user engagement dynamics is an open question. Previous studies, for example by Wilkinson et al. [83], have investigated the importance of fostering trust between human-like digital agents such as chatbots and users to encourage their use, for example by having the digital agent provide justifications for their user-facing actions. Since AI clones are representations of real-world individuals, mediating issues of trust between the source individual, the clone, and the user is of primary concern, for example so that individuals do not feel as if their identities are being commodified or made into transactional objects. There is also a growing body of literature about the psychological effects of using digital "doppelgangers" taken to represent a real-world individual [5, 7, 47, 57, 58, 67]. Work by Muresan & Pohl [61] on chatbots has also shown a design tension between having overt social cues displayed by the digital agent, which can make a system feel fake, and a system's inability to adhere to social rules, which can make an interaction seem unnatural, an issue AI clones will have to overcome as well. On the other hand, previous work has shown that human-like digital agents, such as ReplikaAI [77] or ICT technologies for cultural heritage [50], can provide a "safe space" for users to engage with human-like digital agents without fear of judgment or retaliation, one of the potential benefits of AI clones.

2.2.2 Representations of self on social media. Since AI clones involve replicating the likeness of a real-world individual, issues of impression management become pertinent in the discussion surrounding their development and use. The term *context collapse*, when the multiple audiences of an individual on a social media platform are "flattened" into a single audience has been previously studied [53]. AI clones developed to engage with a broad audience may face similar challenges in managing context collapse. Hogan [38] discusses strategies individuals adopt to establish coherent and consistent presentations of their personal identities on social media in light of such problems. One strategy is the *lowest common denominator* approach, when individuals present only those aspects of their identities they believe are acceptable to all the audiences they engage with. Such a strategy ensures a consistent presentation of an individual's identity, although this may not necessarily cohere with the true multi-faceted nature of an individual's personality. Freeman et al.'s research on creating avatars in virtual reality social environments has highlighted how the process of constructing and perceiving identity in these environments is dependent on aligning them with one's offline identity [25]. They also point to the ability of social VR contexts for enabling users to experience their everyday self in new ways or explore alternative gender identities, an experience that we hope AI clones can also support.

Previous research has also highlighted the potential for identity technologies to be used to manipulate one's own image and perception by others, perhaps to present an idealized form of one's personality, especially in users with a less coherent sense of self [55], lower self-esteem [68] or as a means of obtaining social capital strategically [17]. The potential for AI clones to adopt a wide variety of mediums and degrees of fidelity for representing the source individual therefore raises questions about how idealized representations will be navigated by source individuals that we wished to investigate.

2.2.3 CSCW and HCI research on managing personal data after death. A growing body of HCI research is also concerned with the treatment of personal and social media data after death. This topic stays close to our research on AI clones, especially for applications like post-mortem AI clone experiences and how they might impact the interpersonal relationships and memories of the living and the dead. For example, research by Gach & Brubaker has investigated challenges in representing such complex phenomena as interpersonal trust in socio-technical systems (like the Facebook legacy contact system) [27], and difficulties in having technological interfaces encourage thoughtful and deliberate choices about how a deceased individual will continue to be represented on social media platforms. Such difficulties in post-mortem representation are also shared by AI clones, not only if the source individual has passed away, but also if the source individual chooses to (or must) give access of their clone representation to another stakeholder. Ways in which the source individual might be misrepresented by making their personal data more interactive, and the effects this may have on their interpersonal relationships with potential users, are also some of the questions we wished to investigate. In Chen et al. [13], participants were more receptive towards design concepts that presented the personal data of the deceased as-is to users in comparison to more interactive forms, such as the creation of an AI replica (analogous to an AI clone) of the deceased, which was thought to be "creepy" and "controversial". Whether this sentiment extended to AI clones more broadly and in other application domains is of interest to the present study.

3 METHODS

As an initial exploration of a class of speculative technologies, we wished to orient participants to the design space while leaving the door open for the personal experiences of the participants to imagine alternative futures elicited by AI clones. To accomplish this, we conducted a qualitative study involving the presentation of 8 speculative scenarios to participants accompanied by a semi-structured interview. Our approach took inspiration from prior work in design fiction and Value Sensitive Design, most notably Briggs & Thomas' scenario-based elicitation studies [11], Nathan et al.'s value scenarios [64], as well as the work of Fiesler [19, 20] and Wong in speculative design [85, 86].

3.1 Speculative Scenarios

We wished to present to participants a representative range of speculative scenarios that efficiently communicated the concept of AI clones across representational mediums and use cases. To achieve this, we used a scenario generation method similar to that of Cheon & Su [14] for fictional autobiographies (FAB). However, whereas FAB techniques are geared towards having expert participants (e.g., roboticists in the case of Cheon & Su) craft a story of how the scenario came to be, we used the scenarios as open-ended prompts to capture participant reactions to speculative use cases for a range of hypothetical AI clone technologies, with a focus on the perspective of non-expert users and source individuals.

We first analyzed popular media, such as news articles and YouTube channels, for AI clone-like applications that could serve as foundations for our scenarios and that would encourage participants

Modality \ Application Domain	Labour	Therapeutical	Sexual Use	Cultural Heritage	Entertainment
Chatbots	Scenario 4	Scenario 8			Scenario 1
Voice Assistants	Scenario 4	Scenario 8			
Deepfakes			Scenario 3		Scenario 7
Virtual Reality		Scenario 5/8	Scenario 6	Scenario 2	

Table 1. Summary of AI clone modalities and application domains in speculative scenarios

to think of themselves as "characters" in the scenarios. We then considered a list of values and ethical principles that we felt would be of interest to imbue our scenarios with, taking inspiration from value scenario generation methods [26, 64] and imbuing them with considerations from AI ethics guidelines outlined by Jobin et al. [43]. Our scenario development was guided by the four features of futuristic autobiographies outlined by Cheon & Su [14]; they were developed to be expeditious (i.e., written in 100-200 words), to have captivating content, to lead to an open-ended discussion, and to reveal a participant's values and ethics. We initially developed 12 speculative scenarios, which were pared down to 8 speculative scenarios to fit into an hour-long interview. This was done after piloting the scenarios with 4 participants and merging scenarios with overlapping values and design issues.

Representational mediums depicted in our speculative scenarios were based on existing technologies, including chatbots, voice user interfaces, deepfakes, and VR. A notable omission from this list are humanoid robots, since the added complexity of considering AI clones that have a physical, hardware-based presence is already being explored in the realm of robotics [16]. Application domains represented were chosen based on a combination of use cases discussed in prior literature [50, 65] and popular media examples [41, 45, 54]. The mapping between representational mediums, application domains, and our speculative scenarios is summarized in Table 1.

The final group of 8 speculative scenarios is summarized in Table 2. The scenarios consisted of text-based prompts accompanied by either images of what this speculative technology may take the form of (Scenario 1, 6, 7), or other visual stimuli (Scenario 2, 3, 8). Two scenarios also included video demonstrations of the intended nature of the technologies contextualized in the text prompts (Scenario 4, 5). This was done to increase the credibility and plausibility of the speculative scenarios to participants. The scripts of the scenarios can be found in Appendix A and the scenarios in full can be found in the supplementary materials.

3.2 Participants

The study was piloted with three undergraduates and one graduate student, during which time the speculative scenarios and study protocol went through major iterations as new scenarios were developed in response to domains and issues of interest participants mentioned in pilot interviews. A screening survey was posted in university-centered social media channels, including Discord and Reddit, of a major North American city. We received 85 survey responses and conducted the study with a purposive sample of 20 participants to obtain a mix of age, occupation, and gender/sexual identities being represented. The screening survey also asked about social media platforms most commonly frequented by participants, usage habits on said platforms, and relative ratio between posting and consuming content on the platforms, although this did not significantly impact participant selection. Of the 20 participants interviewed, 11 were younger adults (19-29), 7 were middle-aged (30-69), and 2 were older adults (70 years of age or above) [2]. 6 participants

[2]As a note, participants were given pseudonyms in their transcribed interviews, and they are used, where appropriate, in their quotations. Pseudonyms were created to reflect the gender identity of the participant, but otherwise preserved anonymity.

Scenario \ Value/Ethical Principle	transparency	justice, fairness, & equity	non-maleficence	responsibility & accountability	privacy	beneficence	freedom & autonomy	trust	dignity	solidarity	Summary
1: *Retail chatbot*	X	X		X							An online store owner is misrepresented by a chatbot AI clone when the clone issues a response the source individual would not normally say.
2: *Indigenous elder*		X			X	X				X	An Indigenous elder agrees to have their likeness embodied in an AI clone for visitor interactions in a museum.
3: *Non-consensual deepfakes*	X			X	X			X	X		A student athlete has their likeness stolen to create non-consensual AI clone pornography.
4: *Call centre agent*					X					X	A call centre agent is pressured to turn over personal data to create an AI clone for customer service applications.
5: *VR reunion*	X				X						A mother is offered an opportunity by a grief counselling team to reconnect interactively with her deceased daughter embodied in an AI clone.
6: *Significant others*	X				X		X	X	X		A couple agree to the creation of AI clones of one another to manage a long distance relationship.
7: *Physical swap*							X		X		An online content creator alters the physical appearance of their AI clone to make their presence more marketable to a digital audience.
8: *Commodifying grief*		X	X	X				X	X		A company offers a grieving widow increasingly high-fidelity AI clone applications on a subscription basis, prompting concern from friends.

Table 2. Summary of values/ethical principles incorporated into speculative scenarios. Full text of scenarios and accompanying stimuli presented to participants can be found in Appendix A.

ID	Name	Age	Gender Identity	Sexual Orientation	Ethnicity	Occupation	Weekly social media consumption
P1	Izandra	25-29	Woman	LGBTQ+	White	Lecturer	5-10 hours
P2	Selena	30-39	Woman	Heterosexual	East Asian	Graduate student	15-20 hours
P3	Kyle	19-24	Man	Heterosexual	East Asian	PhD Student	5-10 hours
P4	Samantha	25-29	Woman	Heterosexual	South Asian	Software Engineer	10-15 hours
P5	Kenny	30-39	Man	Heterosexual	East Asian	Research Technician	20 hours or more
P6	Lilith	19-24	Woman	LGBTQ+	East Asian	Student	2-5 hours
P7	Leanne	19-24	Woman	LGBTQ+	White	Landscaper	10-15 hours
P8	Mary	19-24	Woman	LGBTQ+	Latin American	Student	20 hours or more
P9	Jennifer	19-24	Woman	LGBTQ+	East Asian	Student	10-15 hours
P10	Queenie	50-59	Woman	Heterosexual	East Asian	Economist (retired)	5-10 hours
P11	George	50-59	Man	Heterosexual	White	Decision analyst	1-2 hours
P12	Ellie	50-59	Woman	Heterosexual	Eastern European	Senior's Companion	5-10 hours
P13	Nathan	30-39	Man	Heterosexual	East Asian	Realtor	5-10 hours
P14	Faye	19-24	Trans Man	LGBTQ+	East Asian	Student	15-20 hours
P15	Sharon	50-59	Woman	Heterosexual	South Asian	Entrepreneur	10-15 hours
P16	Carol	19-24	Non-binary	LGBTQ+	East Asian	Piano teacher	10-15 hours
P17	Jeremy	70+	Man	Heterosexual	White	Retired	2-5 hours
P18	Luna	70+	Woman	Heterosexual	White	Retired	10-15 hours
P19	Petra	19-24	Woman	LGBTQ+	East Asian	Student	15-20 hours
P20	Brandon	19-24	Man	Heterosexual	East Asian	Content Strategist	10-15 hours

Table 3. Summary of participant demographics

identified as men, 12 identified as women, and 2 participants preferred to self-identify. In addition, 8 participants self-identified as part of the LGBTQ+ community. Occupations included research technicians, students, decision analysts, entrepreneurs, senior's companions, realtors, teachers, software engineers, landscapers and retirees. While we strove to achieve a diverse participant pool in the dimensions above, the authors note that all participants resided in North America.

Participants gave informed consent prior to participation. A breakdown of participant demographics is provided in Table 3.

3.3 Data Collection and Analysis

Following completion of the screening survey, interested participants were emailed by a member of the research team and provided with a link to view the collection of speculative scenarios asynchronously and provide some short text-based first impressions, as well as to arrange an interview session following completion of the viewing. Each interview session consisted of three parts: A few short introductory questions, an exploration of the speculative scenarios using their asynchronous responses as starting points for discussion, and a debriefing interview. One member of the research team conducted the interviews over Zoom, recording audio and video footage as well as taking notes. Participants received CAD $45 in compensation following completion of the study.

Participants were first asked general questions about social media usage and impression management. Following this, they were asked to open their first impressions to the speculative scenarios in their browser windows, and were asked which of the scenarios sparked the greatest emotional responses after their initial viewing. Participants were reminded that the speculative scenarios were chiefly designed to assess their attitudes towards technologies that replicated the human-likeness of someone they knew, rather than evaluate outright the marketability or technical feasibility of the designs depicted. The research team member then asked follow-up questions such as "Was there a personal experience in your life that caused you to react this way to this scenario?" in order to elicit possible explanations underpinning the responses of participants.

Each speculative scenario was discussed in turn before the research team member concluded with a few debriefing questions, which involved inquiring about the general impression participants had towards AI clone technologies as well as their potential hopes or fears if such technologies were to become commonplace. The allocation of time per scenario was not uniform, as some scenarios elicited longer and more in-depth responses from participants compared to other scenarios, depending on the interests and concerns of the participant. Nonetheless, each of the speculative scenarios was brought up at some point during the interview, although a limitation might be that greater emphasis was placed by participants on scenarios with high emotional valence, such as the grief-oriented scenarios, in comparison to other scenarios, like the scenarios involving the use of AI clones for labour. The interview script is available in the supplementary materials. The interviews were recorded and fully transcribed by an external service for subsequent analysis.

Due to the sensitive nature of some of the topics discussed over the course of the interview, including participant opinions on trauma, processing grief, sex work, and digital pornography, the research team member who conducted the interviews had completed prior training in counselling and community mental health. Participants were also given the option to not answer any questions at any point during the asynchronous survey and the interview, although we note that the 20 participants we interviewed did not decline responses to any particular scenario. Participants were also redirected to the appropriate mental health resource during the debriefing interview if they exhibited signs of distress. Our study protocol was approved by our institution's ethics review board.

We conducted a reflexive thematic analysis [10] on the resulting transcripts. The lead researcher read through the interview transcripts, generating both deductive codes, based on prior literature, and inductive codes found in the data. An example of a deductive code was "impression management" that was based on previous literature in impression management. An example of an inductive code generated from the data was "living memories", a code which we had not anticipated from our literature search prior to the beginning of the interviews. After the first round of open

coding was completed, they were categorized into higher-level codes before being transformed into themes. This data analysis procedure occurred simultaneously with data collection. The use of an iterative data analysis procedure allowed the research team to re-orient interviews according to codes and emergent themes of interest as data collection occurred, which we felt was important for exploratory work. To this extent, the codes and themes were restructured after P9, P15, and P20. After reviewing the data after P9 and P15, the diversity of the participant demographics was also reviewed and purposive sampling was undertaken to augment demographics that were underrepresented, particularly that of older participants, who were difficult to recruit during the COVID-19 pandemic. This can be primarily attributed to low response rates from older adults on our recruitment platforms of choice (social media) and our inability to use more traditional recruitment methods for this demographic during the COVID-19 pandemic (e.g. putting up flyers in community centres, visiting senior care homes). The research team met on a weekly basis to discuss suitable groupings for codes and defined themes that served as the basis of our findings.

3.4 Positionality Statement

The authors of this paper see the development of AI clones as an inevitability, given the rapid rate of progress in the underlying technological fields that will make large-scale development of AI clones feasible. Without a moratorium in their development, which we believe to be unlikely, we will soon be faced with a nascent design space that has many real world implications that could affect the lives of people on an individual and societal level. For these reasons we believe studying perceptions and attitudes towards AI clones to be a pressing issue. If AI clones are to be developed, designers might consider ways of ensuring that they are built with beneficial outcomes in mind, while doing our best to mitigate nefarious or malicious use cases. As an initial exploration of these attitudes concerning a speculative technology, we acknowledge that the aspects of AI clone development we choose to focus on here are non-exhaustive. In addition, the authors would like to draw attention to the approach taken here, of analyzing participant perceptions/attitudes and imagining future concerns to putting forth design recommendations, may not necessarily reflect the actual impact AI clones will have when/if implemented broadly. However, it is our belief that making these connections is a productive effort in the design space, and while we are cautious about the particular issues and concerns AI clones raise in their application domains, we are also watching with great interest what new, exciting, and positive social interactions AI clones will be able to enable. As researchers working within an academic institution, we are also limited by our personal interpretive lens. We acknowledge that our analysis is predominantly framed from a Western, educated, and privileged perspective. Further work in this growing design space is well-positioned to articulate stakeholder interests from perspectives that differ from ours.

4 FINDINGS

In the preliminary portion of the interviews, we learned that prior to engaging with the speculative scenarios, the majority of participants had some degree of exposure to AI clone-adjacent technologies. These spanned applications like automated customer support agents, deepfakes, or virtual influencer avatars (VTubers). Experiences with these AI clone-adjacent technologies ranged from intrigue at the potential for greater connection with digital entities (such as VTubers) to annoyance at the perceived lack of human interaction (especially in the context of applications like customer support).

Most participants raised concerns about potential crises of identity and relationship breakdowns that would arise when a source individual was embodied in an AI clone. The act of "copying" the source individual into an AI clone generates a tripartite relationship between the user, source individual, and AI clone, and breakdowns in the relationships between the three entities were

featured prominently in participant reactions. The specific nature of these breakdowns and the contexts and features of AI clone technologies that give rise to them are the primary focus of our findings.

4.1 Doppelganger-phobia: fears of replacement and exploitation

Initial impressions of the AI clones depicted in the speculative scenarios were characterized by sentiments that they were "uncanny," "weird," and "creepy" because they attempted to emulate the source individual. Concerns about the impact of AI clones on the well-being of the source individual and a defensive attitude against threats to personal identity caused by AI clones were some of the general trends that surfaced in our data collection. We refer to this attitude of emotional aversion towards seeing oneself replicated by an AI clone as "doppelganger-phobia." In our interviews, AI clones were likened to human-machine chimeras that made novel forms of human replacement and identity exploitation possible.

4.1.1 Abusive potential of AI clones to objectify and exploit human identity. The most prominent concern about selfhood was the potential for AI clones to "dehumanize" and "objectify" the source individual. Participants used this narrative to express their concerns about the potential reduction of human identity to a commodity. When describing such concerns, the rhetoric of consumption and exploitation of identity occurred repeatedly in interview data.

> *"[...] if people actually make money out of information about me in whatever way that I haven't given my consent to, then I resent that, because I'm being treated like a commodity, or at least my information is being treated as a commodity that someone is harvesting, and that seems morally repugnant to me." (George)*

> *"I don't know if I would ever consent to having a version of myself made, especially for the use of someone else's use, someone else's consumption of me." (Petra's response to Scenario 6: Significant others)*

These participants (George, Petra, Lilith, Faye, Carol) attributed moral status to human identity, asserting that one's identity should not be produced, copied, stolen, or sold. They speculated that AI clones could turn individuals into transactional products for someone else's profit and pleasure, overstepping moral boundaries of identity.

Speaking more directly about the abusive potential of AI clones, participants imagined various forms of identity exploitation that AI clones could enable. Some participants (Selena, Kyle, Izandra) made comparisons to existing examples of identity objectification involved in portrait rights, idol worship, celebrity culture, and personal data appropriation. However, many others saw how novel forms of identity exploitation, such as an interactive form of digital sexploitation, could only be made possible by the unique qualities of AI clones. Lilith and Izandra described their perception that AI clone-based identity exploitation is riskier than existing forms of misappropriating personal data because of the interactive and embodied nature of AI clones.

> *"[...] if you kept your ex's nudes, I think that would be wrong. So this is like that, but much bigger. A photo is a photo. This is more interactive by a lot." (Lilith's response to Scenario 6: Significant others)*

> *"I think the horror there is almost a body horror, where like, if you're almost taking control of my body and my identity to do something else with it [...]" (Izandra's response to Scenario 3: Non-consensual deepfakes)*

These participants considered AI clones to be a new and unpredictable platform that enabled unseen ways of interacting with personal identity, some of which were disturbing and concerning when control was lost by the source individual being embodied.

The concern for the exploitative nature of AI clones seemed to be exacerbated by participants' skepticism towards feasibly implementing proper consent and control mechanisms. Many participants, although they agreed such mechanisms were necessary, reacted to the issue with strong pessimism. Izandra articulated how capitalist pressure and power imbalances could enable corporate entities to prey upon the identities of the vulnerable.

> *"I mean, there's the standard capitalist overlords going to do capitalist overlord things, where I could definitely see people being exploited and the content they made being reused without their consent, and all of that. [...] That's the sort of scenario that I get worried about. And the example of somebody being like, I am poor and I need to eat, and creating these simulacra of grandpa's things will help us put food on the table for the next year. It's not their fault. They're not doing anything wrong. They're doing their best to survive in a cruel world. But that power imbalance exists is bad, I always say." (Izandra's response to Scenario 2: Indigenous elder)*

Several other participants speculated that existing issues with consent barriers [31] (Brandon, George, Izandra, Parnoosh, Samantha) and consent to data use after death [73] (Selena) would persist in AI clone applications.

4.1.2 "A better version of me": speculations on AI clones replacing source individuals. Many participants (Carol, Kenny, Izandra, Sharon, Lilith, Samantha, Brandon) spoke of their fear of being displaced by an AI clone. Although some comments echoed the existing rhetoric of replacement by AI [40] and robots [63] in work, AI clones suggested a new threat to intimate relationships that are traditionally considered "safe" from replacement by technology. Such concerns were pronounced in the scenarios that involved a close personal relationship with one's significant other or family, such as the Romantic Partner (#6) and Indigenous Elder (#2) scenarios. Carol's remark exemplified the aversion to potentially being replaced by a replica of oneself in a romantic relationship.

> *"You probably want to make it [an AI clone] the most idealized version of yourself for your partner to keep. And also how this version of you acts. Especially if your partner can control or select what they want, I think that's kind of icky." (Carol)*

In Carol's comment, it is worth noting that she assumes the AI clone can be an *idealized version* of oneself rather than an identical version. The capacity to be better than the source person gives AI clones the power to displace the individual.

We found that sentiments about the fear of replacement are typically made in tandem with the assumption of a superior replica. In these comments, AI clones were considered human-machine hybrids that transcend the individual and perfect their characteristics. Our participants speculated AI clones would be superior compared to both the source person and generic human-like AI in three ways: By having perfect access to personal memories (Sharon, Samantha), by having an aesthetically ideal version of one's body (Leanne), and/or by being able to selectively incorporate good characteristics and leave out/fix bad ones (Sharon, Brandon).

> *"I would think that my interactive clone would know more about me than I know about them, since they take my memories or whatever and implant it into the system. So it's almost like they would have more knowledge and information about you than you do of yourself. I don't have immediate access to all my memories, whereas the clone would. So it's a perfect version of you, which is weird because we're all human and we're all fallible. And so I think that there's this version of you, it displaces you. It's like, why am I needed?" (Sharon's response to Scenario 7: Physical swap)*

> *"If they are there, why do I need to be here? Would they displace you and your usefulness? Because if they're virtual, they could probably do everything you can do, but better. It*

> *would be almost a better improved version of you. Yeah. You can take away the annoying habits, make them work harder. So why would I be needed? It's like, yeah, let's get rid of these habits, bring these in, these traits. And so it's like creating the perfect person, which is yeah. It's weird. It's too weird." (Sharon's response to Scenario 6: Significant others)*

In these excerpts, AI clones are envisioned as technological constructs that surpass both humans and generic human-like AI. Concerns were raised that the source person would not be able to measure up to the superior qualities of their AI clone while AI clones are capable of playing the source individual's role in intimate and emotional relationships that generic human-like AI cannot.

4.2 Identity fragmentation: anticipated challenges in creating representative and cohesive AI clones of individuals.

In our interview data, a prominent risk of AI clones anticipated by our participants was the challenge of maintaining consistent selfhood. AI clones, by definition, create replications of the source person that are based on and assimilate their traits. This caused participants to remark that AI clone technologies could *fragment* their identities across the multiple actors that realize their identity. Identity fragmentation caused by AI clones could compromise the source person's own feelings of who they are as an individual and jeopardize the internal cohesion of their identity. Further analysis revealed two ways participants thought AI clones create risks in cohesive perceptions of the self, depending on the nature of the embodiment: a clone that captures a snapshot of one's past and clone(s) coexisting with the source individual.

4.2.1 Drifting away from the "snapshot": AI clone risks for cohesive and consistent self-perception. Participants (Faye, Nathan, Petra, Jennifer) speculated that the identity and representation of the source individual would begin to "drift" once they were captured in an AI clone as time went on. These concerns assumed that AI clone technologies serve essentially as snapshots of the source person's identity at a certain point in time, and the system will not change the clone's behaviour and personality afterwards. The participants speculated on how this lag or "desync" between the identities of the ever-changing source individual and the behaviour embodied in the AI clone could potentially contribute to feelings of identity fragmentation.

These participants emphasized how embodying a snapshot of their identity in an AI clone could introduce tension between their past and present selves because behaving consistently with the version of themselves embodied in the AI clone can betray who they were in the present. Petra's quotation best elaborates this dilemma:

> *"I also feel like it would hold me back a bit. You would feel more inclined to continue being the same person as the version that was captured so that there's no inconsistency because if you notice that there's a version of you that was supposedly you at one time, that if you change enough and they seem so different from you, I feel like that would definitely cause you a lot of stress in terms of your identity. Was it good? Was it bad? It's like, I don't know. I feel like in a lot of cases, a reminder of your past is not always a good thing." (Petra)*

In Petra's quote, the negative undertone about one's past self is worth noting.

Several other participants were more explicit about the contrast between an "evolving" self in the present and "stale versions" of oneself from the past. Embodying a snapshot of oneself in an AI clone was concerning not only because of the inconsistency but also because one's past self becomes outdated. These participants implied that living humans have the capacity to evolve their memories, thoughts, and perspectives on the real world over time, while AI clones may lack the potential for growth.

> *"[Your] personality changes on a daily basis. Things you learn, things you do change who you are. But if a virtual clone is not able to do that, then that version of you is stuck like that forever." (Nathan)*
>
> *"The idea that there is this lingering snapshot that you're going to grow out of, though people are still going to see this snapshot of you." (Faye)*
>
> *"I know they've been dating since high school and I know they know each other, but the fact is that they're 21, they'll be 23 when this is all said and done and people change so much. But you're still putting that person, you're pigeonholing them into that age that they were. (Ellie's response to Scenario 6: Significant others)*

In these comments, a past self captured in an AI clone was referenced alongside negative expressions such as "stuck," "lingering," and "pigeonholing." AI clones risk the continued materialization of a past version of self that the source individual would have otherwise forgotten.

Updating AI clones for accurate and consistent presentation of the source person was considered a viable remedy to the inconsistency problem. Leanne discussed how the AI clone needs to be regularly checked to "make sure they are up to date" with the source individual.

> *"I think keep them around as long as possible and just intermittently check to make sure that it's still accurate, I think, because stories change, perceptions change, everything. I think as long as it's not kept entirely static, then that's fine. I'd say don't get rid of them but maybe make sure that they're up to date, if that makes sense. Make sure they're moving forward with us." (Leanne)*

This speculative solution assumes the possibility of AI clones growing and evolving over time.

Other participants, Nathan and Faye, also brought up the idea of having AI clones that grow over time and the importance of consistent presentations of identity. More importantly, their remarks suggested another issue with consistency – the clone might evolve its *own* identity. Compared to the case of AI clones as snapshots, an evolving AI clone complicates the problem of identity fragmentation because the clone can branch off its own identity that does not mirror changes in the source person, eventually causing the two selves to drift apart.

> *"Unless a virtual avatar has some sort of artificial intelligence designed into it, then it develops into a completely different version of you, and that's a whole different ball game." (Nathan)*
>
> *"That can be a problem where it's like, this bot either doesn't grow or it grows in a way that is probably going to be different from how I, as a person, would actually grow." (Faye)*

However, participants did not extend this concept of evolving AI clones to implications and ramifications, leaving further speculation to our discussion.

4.2.2 Creating AI clones active in parallel with the source person diminishes individuality. The concept of introducing copies of the source person seemed to contribute to perceptions of diminished individuality. Participants (Izandra, Carol, Samantha, Sharon, Leanne, Lilith) saw value in the things made an individual unique.

> *"It feels like making a copy of a person that's already in existence, I think. The cool thing about humanity's how every single person is different, learns from their own experience, and that kind of thing." (Leanne)*

However, creating an AI clone of a person seemed to reduce the value of the source individual because it no longer made them the unique instantiation of their personality. In this regard, participants were concerned about two particular aspects of AI clone usage and its consequences

for individuality. First, quantity mattered. Creating *multiple* copies of the source individual felt increasingly worrisome to participants. Leanne and Lilith articulate their issues with plurality:

> *"I think my biggest concern, I guess, would be retaining the individuality of humans. I think if you just end up copying... If you make one AI robot that acts like a person, you make an entire army, they're all the same, that freaks me out." (Leanne)*

> *"Making one copy seems a little dehumanizing, but then making multiple copies seems really, really dehumanizing. (Lilith)"*

Second, time mattered. Introducing an AI clone that co-existed *at the same time* as the source individual seemed to compound the problem of diminished individuality. Lilith elaborates on how she felt about the issue of concurrency:

> *"I think I would be more okay if Bill had already passed away and then a surrogate was made of him. But to have him walking around and then also to have just this object version of him is really... I don't like it." (Lilith's response to Scenario 2: Indigenous elder)*

The converse to the above is AI clone embodiment after death, when the source person is no longer interacting with the world when the clone is. Some participants considered post-mortem AI clone embodiment problematic due to issues of consent and control over the personal data of the deceased. This is a different kind of risk reported in section 4.1.1.

4.3 Concerns about misrepresentation and over-reliance potentiated by AI clones

Participants regarded AI clones as intermediaries that could impact existing relationships between the user and the source person. When participants put themselves in the shoes of the source person, they emphasized the potential risk for AI clones to create a false image of the source individual. In purview of the user, participants were concerned about commercial incentives of AI clone applications that could cause users to become dependent on emotional experiences with AI clones.

4.3.1 False image: misrepresentation of the source individual to the user. Several participants expressed their concerns that users could conflate interactions with an AI clone with interactions with the source individual. Such confusion might forge a *false image* of the source individual, which can compromise their interpersonal relationships. Participants pointed out that the formation of a false image in the user could result in the source individual being falsely accused or held responsible for the actions of AI clones that bear their likeness.

> *"If the AI says something, then the [user] gets upset. Does that mean that they're upset at their real-life partner or are they just upset at the AI that's meant to keep them company in place of their partner? That's going to confuse them and it's going to cause friction." (Leanne's reaction to Scenario 6: Significant others)*

Leanne anticipated that a false image would create friction in the relationship between the user and the source individual due to ambiguities in blame attribution. The problem is exacerbated when the source individual is not aware of the interactions their AI clone has with a user.

Participants drew attention to two different types of ways the formation of a false image could occur in the user: Fake memories and misrepresentation. With fake memories, the AI clone experience can produce a distorted or fabricated *memory* about the source individual in the user. Brandon expressed concerns about the potential for AI clones to cause "memory distortion" in users, which risks compromising the image of a deceased individual.

> *"I think the problem would be if it's bringing up things that aren't true and that never happened. [...] I think the memory distortion would be bad, but it's the distortion of your thoughts and your beliefs [about the source individual]. If that were to maybe change in*

> *the clone over time and you remember them [the source] differently, it could be sort of scary." (Brandon)*

Along this line, many other participants mentioned that interacting with an AI clone may cause problems in interpersonal relationships if the user is not "aware" that they are interacting with an AI clone and not the source individual, for example in a voice-based medium where it might be difficult to differentiate the AI clone from the source individual.

Comments about misrepresentation by an AI clone seemed more subtle, indicating that the user's experience with an AI clone can create a *false impression* of the source person simply due to the *resemblance* between AI clone and the source individual. Participants speculated this may occur even if the user is clearly aware that their interactions were with an AI clone and not the source person, in contrast to the issue of false memories described above. For instance, responding to the *Significant others* scenario, Faye noted that the resemblance of an AI clone to the source individual could cause them to *project* feelings they had about the source individual more easily onto the AI clone:

> *"This [AI clone] said something and it looks like you and it's trying to be you. That would probably have an effect on your image of your significant other in your mind." (Faye's response to Scenario 6: Significant others)*

> *"It would just feel really weird to me that they now have this image of me in their head of the experience they had that I was not there for and I don't know what happened and it kind of brings up like, 'Okay, what else has the bot done that I don't know about?' Like how has my image changed due to what this bot has said? It feels like I'm losing control of my identity in relation with other people. There will probably be an effect in my relationship with other people, even though it's a bot that's doing all this." (Faye's response to Scenario 6: Significant others)*

This can be considered a special case of the transference effect [4], where one (the user) redirects their social perception of an individual (the clone) to another person (the source person).

4.3.2 "Living memories": risks of over-reliance on AI clone interactions. When users interact with AI clones that replicate source individuals they know, the AI clone can materialize shared past experiences between the user and source individual in the present with an interactive dimension. We refer to this concept as *living memories*. The primary concern of our participants towards living memories was the risk of users potentially developing unhealthy attachments and an over-reliance on AI clone interactions. The concept of prolonged interactions with AI clones evoked in some of the scenarios elicited remarks that the technology "seems unhealthy" (Kyle) or might cause "unintended over-dependence" (Kenny). For example, in responses to the *VR reunion* scenario, which involved the "digital resurrection" of a deceased loved one into an AI clone, participants expressed concerns that an over-reliance on the AI clone would form, since "having unlimited access to a past memory is just going to make you ruminate and just live in the past rather than the future". Concerns about over-reliance were compounded by the potential for AI clones to be easily accessible to potential users on-demand.

Participants speculated about how they might dwell on their past if given the chance to interact with AI clones of people close to them, instead of turning to more traditional methods of managing their relationship with complex emotions like grief. These participants saw acknowledging the reality of death as a fundamental part of the human experience, and artificially altering this using technologically-mediated methods was seen as problematic. They held strong attachments to their existing memories of an individual. When the scenarios introduced the potential for these memories

to become distorted, participants felt that memories important to them were threatened if they interacted with the AI clone regularly.

> *"I think if there's anything like this, I think it would be impossible for someone to be able to move on with life because who wouldn't want to talk to a deceased person they loved so much dearly. But then I also think that these VR versions are not real. They look maybe 60% like the person [...], but technology is always advancing. [...]. So I think it would make it really hard for someone to move on." (Samantha's response to Scenario 5: VR reunion)*

> *"I think once someone gets overly attached to this fake version of their past loved one, then you start to get into a little bit of a ... you're liking the fake person instead of the actual memory of your person that has passed away." (Leanne's response to Scenario 8: Commodifying grief)*

Reliance on AI clone interactions also raised concerns in our interviews that users might come to forget their real memories about the source individual, something participants found particularly problematic.

Several participants (Jennifer, Mary, Selena) were also particularly concerned with the ways in which emotional dependencies on AI clones could be exploited commercially for profit. Responding to scenarios in which grief could seemingly be commodified, for example by offering a paid service in which a deceased loved one could be replicated by an AI clone, Jennifer elaborated:

> *"It's kind of scary that soon grief could be commodified in this way, where companies can take advantage of that [over-reliance on AI clones], and kind of prey on people who've just lost someone. I really just think it's putting off the inevitable, which is that this person has gone from your life, and you just have to learn to live with that." (Jennifer)*

Participants indicated that emotional attachments the user transferred from the source individual could be particularly troubling in the context of commercialization. One participant remarked how AI clones could be "built on this kind of fake relationship [where they] try to act like your friend, but in the end they just kind of want profit". Concern was expressed about how the potential benefits of AI clones could be marketed in a way that was exploitative, making potential users of AI clones dependent on recurring subscription-based models to keep their relationship with the AI clone "alive".

In speculative scenarios pertaining to grief, participants wanted to practice a proactive approach to prevent over-reliance from occurring in the first place. Participants suggested two types of potential remedies in the interview data. First, limiting the number of interactions with AI clones was seen as desirable by participants. Responding to the *VR reunion* scenario, Leanne commented:

> *"I like what the woman did in this situation, where she went for the one session, got a sort of closure for herself then moved on." (Leanne's response to Scenario 5: Virtual reunion)*

Second, participants expressed a desire to see user access to AI clones limited to sanctioned interventions by qualified experts, such as therapists and expert consultants. In response to one of our scenarios, one participant remarked how their opinion of AI clones changed when they saw the context in which it was being used:

> *"At first I thought 'Well, this is crazy. How's this going to work?' That is until I realized that it's not just a person randomly hooking up to an Oculus at home. There's actual therapists walking through the session and when I heard that then I thought, 'Okay, this makes a lot more sense because there's a point to it, there's a healing process to it.' If it's just some person escaping from reality then that becomes dangerous." (Nathan's response to Scenario 5: VR reunion)*

In what ways designers can further mitigate or prevent the formation of unhealthy emotional dependencies is further elaborated in our discussion section.

5 DISCUSSION

From our findings, participants sketched out a number of concerns they had with the AI clone technologies. In this section, we highlight ramifications and implications informed by participant speculation that we hope can be used to develop productive conversations around this nascent design space. These include a discussion of how respect for identity and authenticity might come to be viewed as design principles for AI clone technologies, the identification of showstopper ethical concerns about AI clone technologies, and discussion about the rights and responsibilities of various stakeholder groups in the AI clone design space. Future work in AI clone design is well placed to extend and augment some of the discussion here.

5.1 Respect for identity and authenticity as design principles

Literature discussing technology and race [35] has highlighted the impact of our identities when interacting with embodied technologies [87]. As interactions with embodied technologies extend our own identity and become part of our extended self, it influences the source individual's behaviour towards other people they work with [56], and motivations for interaction in social settings [24, 25]. From our findings, participants reacted to the speculative scenarios of AI clone interactions from the vantage point of one or more personas/aspects of their identity. These personas (e.g. being a partner, a friend, and/or a granddaughter) contributed to perceptions participants had about AI clones, and led to varied optimism about interacting with AI clones. In turn, faced with the potential of identity fragmentation, some participants shared that they felt pressured to keep their identities consistent with their AI clone. In Goffman's terminology, having different facets of oneself is natural and part of our nature as social beings [29]. We do not need to explicitly remind ourselves of how to behave every time we interact with others using different personas that make up our personal identity. However, there seemed to be tension when different personas of a source individual's identity confronted each other as the result of being embodied in an AI clone. For example, an AI clones version of being a friend to a user compared to the source individual's version of being a friend could potentially come into conflict and compromise the way the source individual is perceived by the user.

As such, if designs of AI clones keep the potential impacts on the source individual's identity front and center in design decisions, the designs may have more positive outcomes for the sense of selfhood and bodily autonomy of the source person. Considering which persona or "face" of the source person an AI clone is trying to replicate and present thereby becomes a potential guiding question – for example, is the AI clone trying to be a colleague, a parent, or a friend? When the designs of AI clones take into consideration how the existing relationships and faces of the source person encompass their identity, they may be able to better exhibit relevant and appropriate mannerisms that the source individual possesses. It may also be advisable that users are made aware of when they are interacting with a clone of the source individual, so that they can conduct interactions with appropriate expectations about the authority of the AI clone and what the source individual should be accountable for, as discussed in 5.2.

Our findings also highlighted both concerns and hopes regarding the potential for AI clones to realize *authentic* presentations of a source person's identity. On the one hand, some participants felt that AI clones, as technological artifacts, would never be "real" in the way human beings were. On the other, the resemblance of the AI clone to the source individual was taken by some participants to be an extension of their selfhood. Prior studies on the authentic presentation of self on social media echoes such seemingly contrasting perceptions and their empirical bases, including the drive

to make the presentation of one's online self consistent and "true" to the offline counterpart [34], the opportunities for exploring new facets of self-identity [32], and the concerns about online parasocial relationships [28]. Reflecting on this literature, we envision unique challenges and opportunities for AI clones to foster authentic self-presentation and emotional bonds.

Marwick and Boyd pointed out the problem of *context collapse* in online social media, where a user's presentation is constrained to a single identity that "flattens multiple audiences into one" [53] as a result of the features and affordances of social media platforms. Since the granularity of self-representation is confined to the unit of a user account, social media users often leverage multiple accounts to manage varying facets of their identity (e.g., postings in "fake" or "throwaway" Instagram accounts feature authentic, albeit negative, self-presentations [78]). In contrast, AI clones, as social agents unconfined to the unit of a user account, can adjust their presentation to the encounter with each specific user. Thus, self-representation through AI clones has the potential to offer believable identity representations contextually contingent on a particular social setting. Designers of AI systems often struggle to make agents behave naturally and coherently responding to the given social context [46]. A novel challenge here is to empower the source individual to *control* how AI clones might behave in their autonomous interactions (e.g. in chatbot AI clones).

For example, some participants believed in the potentially beneficial outcomes of AI clone embodiment for authentic presentation and exploration of one's identity. LGBTQ+ participants, in particular, commented on the potential for AI clone technologies to enable them to explore or affirm changing notions of personal identity. This has also been substantiated in previous research on LGBTQ+ participation in digital spaces like social VR [24, 25]. However, in doing so, the literature cautions against the normative aspect of technology. For example, Haimson & Hoffman examined Facebook's design to conclude that the platform ends up excluding and minimizing people with marginalized identities, instead enforcing user profiles that conform to heteronormative social standards [33]. Hence, the design of AI clones might consider embracing the performative, pluralistic nature of authentic identity rather than administering a dichotomy of the authentic versus inauthentic. Similarly, Barta & Andalibi identified that affordances of the platform of an embodiment form a basis for normative notions of authenticity perceived by the audience [8]. If AI clone systems consider supporting features for fluid and non-normative identity expressions, this could lead to a greater acceptance of the technology. A good example is the design of online communities that allow user identities to be selectively visible, so LGBTQ+ individuals can present their identities on their own terms [18]. As the development of AI clones is a nascent design space, best practices for respecting the original identity and authenticity of the source individual will likely be a rich but divisive space. Subsequent empirical design work is well-placed to investigate this further, perhaps with methods such as collaborative design to ensure all stakeholders have a transparent account of how the identity and authenticity of the source individual is affected by the AI clone.

5.2 Preliminary ethical concerns about AI clone technologies

From our findings in 4.1.1 and 4.1.2, participants outlined two primary concerns they had about AI clone technologies in general: one about the potential for the commodification and objectification of personal identity, and another about the potential for AI clones to be the next step in the replacement of human beings, in this case for emotional bonds, by technology.

In their meta-analysis of AI ethics guidelines [43], Jobin et al. identified key principles of AI ethics including justice & fairness, privacy, trust, dignity, and solidarity. The concerns of participants seemed to view the ethical development of AI clone technologies with varying degrees of skepticism. The potential to commodify and make personal identities transactional elicited negative impressions from the majority of participants. Many of their attitudes to the speculative scenarios and use

cases seemed to speak to a desire for a thoughtful design process of managing trust, dignity, and solidarity between source individuals and potential users so that AI clone technologies may have beneficial outcomes.

For example, to mitigate the threat of AI clones replacing human beings in their capacity to provide emotional bonds, designers might encourage AI clone usage to be moderated by trained professionals and build appropriate safeguards to enforce this, as expressed by one of our participants in 4.3.2. However, this creates open questions about who is qualified to moderate AI clone experiences and the mechanisms for such moderation. Alternatively, it might also be possible for designers to incorporate methods for making users of AI clones aware of the artificiality of their experience so that there is a clear delineation between experiences with real human beings and AI clones. Methods for doing this might include visual cues embedded in the representation of an AI clone to alert the user to the fact that they are interacting with a technological construct, or deliberately reducing the fidelity of an AI clone in order to keep the user cognizant of the artificial nature of their interaction, as was done in the *VR reunion* scenario.

Regardless of the mechanism, developing and scoping the appropriate socio-technical context for AI clones is an important first step for the development of this class of technologies. This recommendation, provided in the 3rd revision of the Internet Research Ethics Guidelines for AI and Machine Learning published by the Association of Internet Researchers, stresses the importance of scrutinizing the social, political, and economic contexts within which a technology operates [23]. While we hope our present study contributes to this conversation, future work is well placed to design and innovate on our work and some of the recommendations laid out above.

5.3 Rights and responsibilities of and towards AI clones

5.3.1 Control, ownership, and consent. In our findings from 4.1.1, we alluded to an area of concern from participants surrounding ownership of their identity when interacting with AI clones. These considerations speak to the potential for AI clones to violate a source individual's personality rights, which encompasses the right of an individual to control the use of attributes that serve as identifiers of who they are as individuals, such as their physical likeness, behavioural mannerisms and lived experiences [70]. Technology laws have previously advocated for personal ownership and exclusive control over one's digital image [9], speaking to an individual's right towards embodied agents such as AI clones. When building AI clones, considerations of how the personality rights of source individuals might be adequately protected is one mechanism for fostering greater trust in this speculative class of technologies. Means for achieving such protection might include developer codes of conduct for seeking active, transparent consent for the usage of personal data and mechanisms that allow the source individual to revisit, audit, and revise the elements that make up their AI clone. Allowing the source individual to retain the right to potentially terminate social interactions between AI clones and users when necessary could also foster greater trust in AI clones. Control and ownership over the personal data used in the construction of the AI clone, and the AI clone itself by the source individual might also be considered in design commitments for the ethical development of AI clones.

5.3.2 Stakeholder responsibilities. At the same time, we anticipate the effort in managing ownership of identity in AI clones will be taxing and require a large amount of time, energy, and domain expertise. Handling failure cases of AI clones will also be a major concern in the design of AI clones. Navigating such design issues is likely to involve a complex amalgamation of stakeholder perspectives, including not only technical experts, but perhaps also medical and mental health professionals in addition to activists for legal, privacy, and human rights. As AI clones interact with multiple stakeholders in social settings, the prospect of deciding who the responsible parties are

when the interaction breaks down can be difficult and confusing. Is the source individual responsible for making informed decisions about AI clone technologies? Or are developers responsible for proliferating AI clone technologies? Still others may argue that the motivations of private companies or governmental regulation might be the entities on whom responsibility lies for breakdowns in AI clone interactions.

The inconsistent attribution of ethical responsibility seems to echo a larger trend in the global landscape of AI ethics [43]. In their scoping review of 84 published guidelines for developing ethical AI, Jobin et al. pointed out the inconsistencies in ethical principles that different stakeholder groups had. Though they identified overlapping concerns, mechanisms for further clarifying or resolving conflicting ethical priorities were deemed lacking or non-existent. In the development of AI clones, the ethical impacts of their adoption are likely to be highly damaging and immediately visible due to the interactive and social nature of the technology. At the same time, different parties involved may have common or competing interests during an interaction, or after a breakdown.

For instance, competing stakeholder interests when considering AI clones for grief-related commercial applications might lead to concerns that emotionally volatile individuals are incapable of making decisions in their best interest with a potentially addictive technology. Based on participants' reaction towards AI clones for grieving the death of loved ones, we believe long-term use of AI clones entails the possibility of forming unhealthy social attachments. If designers strive to incorporate mechanisms that encourage the cessation of AI clone usage if users are found to have a high amount of interaction time, this might mitigate the addictive qualities of AI clones. Ways to achieve cessation could include pre-use warnings and disclaimers of the potential emotional effects of interacting with AI clones, interspersed prompts from the system to take a break from usage, and the instantiating of a fixed expiry date for the AI clone. One important consideration in implementing a mechanism like expiry dates is to determine an appropriate "time to live" for an AI clone. It is our hope that the design of AI clones will include the simultaneous development of policy and ethical infrastructure to decide fair ways of accounting for responsibility in failures of AI clone-mediated interactions, whether it be at the level of the source individual, the user, or groups associated with AI clone development.

6 CONCLUSION AND FUTURE WORK

We investigated the perceptions and attitudes of 20 members of the general public towards issues of personal identity and interpersonal relationships stemming from AI clones. From semi-structured interviews, we learned of participant concerns with (1) "doppelganger-phobia", the potential for abusing AI clone technologies to exploit and displace one's personal identity; (2) identity fragmentation, when creating replicas of a living individual threatened to jeopardize a sense of cohesive self-perception and a sense of individuality; and (3) risks to interpersonal relationships engendered by AI clones, for example by the formation of a false image of the source individual or an over-reliance on a living memory materialized by AI clone interactions. These findings were contextualized using speculative application scenarios for AI clones spanning use cases like labour, cultural heritage, grief processing, and sexual/romantic intimacy. Building on these findings, we discuss the role of AI clone research in contributing to a larger conversation in ethical AI, the framing of identity and authenticity as design principles in AI clone development, and the rights and responsibilities of stakeholder groups for responsible development of AI clones, alongside relevant design recommendations like moderation and expiry dates that might contribute to more thoughtful work in this nascent design space.

We established this study along two lines of conceptual inquiry: Ways in which AI clone technologies might reshape the presentation of personal identity and interpersonal relationships. It is worth noting that the selection of these two considerations is by no means exhaustive. In particular,

our exploratory work seeks to capture attitudes, perceptions, and concerns of the general public towards AI clones. The perspectives of alternative stakeholders such as developers or technical experts are given cursory treatment in the discussion, but additional work is well placed to further articulate issues in the design space not covered in this work. These might include issues in data privacy, security, and commercialization of AI clones that deserve equally thorough treatment in the literature.

ACKNOWLEDGMENTS

This work was supported by the KIST Institutional Program (Project No. 2E31583). We would like to thank all the participants who made this study possible, and the people of the D-lab and MUX-lab for their thoughtful feedback throughout the paper writing process. Special thanks to the Stojicevic family for their invaluable assistance to the lead author throughout the research process.

REFERENCES

[1] 2022. *Internet crime complaint center (IC3): Deepfakes and stolen PII utilized to apply for remote work positions.* https://www.ic3.gov/Media/Y2022/PSA220628

[2] Jascha Achenbach, Thomas Waltemate, Marc Erich Latoschik, and Mario Botsch. 2017. Fast generation of realistic virtual humans. In *Proceedings of the 23rd acm symposium on virtual reality software and technology*. 1–10.

[3] Henry Ajder, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen. 2019. *The State of Deepfakes: Landscape, Threats, and Impact.* Retrieved September 3, 2021 from https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

[4] Susan M Andersen and Alana Baum. 1994. Transference in interpersonal relations: Inferences and affect based on significant-other representations. *Journal of personality* 62, 4 (1994), 459–497.

[5] Laura Aymerich-Franch and J Bailenson. 2014. The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. In *Proceedings of the International Society for Presence Research Annual Conference*. Citeseer, 173–186.

[6] Norman I Badler, Nadia Magenat-Thalmann, Laurie McCulloch, Evan Marc Hirsch, and Phil LoPiccolo. 2002. Digital humans: what roles will they play?. In *ACM SIGGRAPH 2002 conference abstracts and applications*. 88–89.

[7] Jeremy N Bailenson and Kathryn Y Segovia. 2010. Virtual doppelgangers: Psychological effects of avatars who ignore their owners. In *Online worlds: Convergence of the real and the virtual*. Springer, 175–186.

[8] Kristen Barta and Nazanin Andalibi. 2021. Constructing Authenticity on TikTok: Social Norms and Social Support on the "Fun" Platform. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 430 (oct 2021), 29 pages. https://doi.org/10.1145/3479574

[9] Joseph J Beard. 2001. Clones, bones and twilight zones: protecting the digital persona of the quick, the dead and the imaginary. *J. Copyright Soc'y USA* 49 (2001), 441–548.

[10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[11] Pam Briggs and Lisa Thomas. 2015. An inclusive, value sensitive design perspective on future identity technologies. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 5 (2015), 1–28.

[12] Sam W.T. Chan, Tamil Selvan Gunasekaran, Yun Suen Pai, Haimo Zhang, and Suranga Nanayakkara. 2021. KinVoices: Using Voices of Friends and Family in Voice Interfaces. *pre-print* (2021).

[13] Janet X Chen, Francesco Vitale, and Joanna McGrenere. 2021. What Happens After Death? Using a Design Workbook to Understand User Expectations for Preparing their Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[14] EunJeong Cheon and Norman Makoto Su. 2018. Futuristic autobiographies: Weaving participant narratives to elicit values around robots. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 388–397.

[15] John Danaher and Neil McArthur. 2017. *Robot sex: Social and ethical implications.* MIT Press.

[16] Maartje MA de Graaf. 2016. An ethical evaluation of human–robot relationships. *International journal of social robotics* 8, 4 (2016), 589–598.

[17] Michael A DeVito, Jeremy Birnholtz, and Jeffery T Hancock. 2017. Platforms, people, and perception: Using affordances to understand self-presentation on social media. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 740–754.

[18] Brianna Dym, Jed R. Brubaker, Casey Fiesler, and Bryan Semaan. 2019. "Coming Out Okay": Community Narratives for LGBTQ Identity Recovery Work. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 154 (nov 2019), 28 pages.

https://doi.org/10.1145/3359256

[19] Casey Fiesler. 2019. Ethical considerations for research involving (speculative) public data. *Proceedings of the ACM on Human-Computer Interaction* 3, GROUP (2019), 1–13.

[20] Casey Fiesler. 2021. Innovating like an optimist, preparing like a pessimist: Ethical speculation and the legal imagination. *Colo. Tech. LJ* 19 (2021), 1.

[21] Luciano Floridi. 2018. Artificial intelligence, deepfakes and a future of ectypes. *Philosophy & Technology* 31, 3 (2018), 317–321.

[22] Asbjørn Følstad, Marita Skjuve, and Petter Bae Brandtzaeg. 2018. Different chatbots for different purposes: towards a typology of chatbots to understand interaction design. In *International Conference on Internet Science*. Springer, 145–156.

[23] Aline Shakti Franzke, Anja Bechmann, Michael Zimmer, Charles Ess, et al. 2020. Internet research: Ethical guidelines 3.0. *Association of Internet Researchers* 4, 1 (2020), 2056305118763366.

[24] Guo Freeman and Divine Maloney. 2021. Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.

[25] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. 2020. My body, my avatar: How people perceive their avatars in social virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[26] Batya Friedman, Peter H Kahn, Alan Borning, and Alina Huldtgren. 2013. Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory*. Springer, 55–95.

[27] Katie Z Gach and Jed R Brubaker. 2020. Experiences of trust in postmortem profile management. *ACM Transactions on Social Computing* 3, 1 (2020), 1–26.

[28] Katie Z. Gach, Casey Fiesler, and Jed R. Brubaker. 2017. "Control Your Emotions, Potter": An Analysis of Grief Policing on Facebook in Response to Celebrity Death. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 47 (dec 2017), 18 pages. https://doi.org/10.1145/3134682

[29] Erving Goffman et al. 1978. *The presentation of self in everyday life*. Vol. 21. Harmondsworth London.

[30] Jonathan Gratch, Jeff Rickel, Elisabeth André, Justine Cassell, Eric Petajan, and Norman Badler. 2002. Creating interactive virtual humans: Some assembly required. *IEEE Intelligent systems* 17, 4 (2002), 54–63.

[31] Colin M Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark patterns and the legal requirements of consent banners: An interaction criticism perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

[32] Oliver L. Haimson, Anne E. Bowser, Edward F. Melcer, and Elizabeth F. Churchill. 2015. Online Inspiration and Exploration for Identity Reinvention. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3809–3818. https://doi.org/10.1145/2702123.2702270

[33] Oliver L. Haimson and Anna Lauren Hoffmann. 2016. Constructing and enforcing "authentic" identity online: Facebook, real names, and non-normative identities. *First Monday* 21 (6 2016). Issue 6. https://doi.org/10.5210/FM.V21I6.6791

[34] Oliver L. Haimson, Tianxiao Liu, Ben Zefeng Zhang, and Shanley Corvite. 2021. The Online Authenticity Paradox: What Being "Authentic" on Social Media Means, and Barriers to Achieving It. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 423 (oct 2021), 18 pages. https://doi.org/10.1145/3479567

[35] David Hankerson, Andrea R Marshall, Jennifer Booker, Houda El Mimouni, Imani Walker, and Jennifer A Rode. 2016. Does technology have race?. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 473–486.

[36] Karen Hao. 2021. *Deepfake porn is ruining women's lives. Now the law may finally ban it.* Retrieved September 3, 2021 from https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/

[37] Matthew Harris. 2013. The hologram of Tupac at Coachella and Saints: The value of relics for devotees. *Celebrity studies* 4, 2 (2013), 238–240.

[38] Bernie Hogan. 2010. The presentation of self in the age of social media: Distinguishing performances and exhibitions online. *Bulletin of Science, Technology & Society* 30, 6 (2010), 377–386.

[39] Michelle Huang. 2022. I trained an AI chatbot on my childhood journal entries - so that I could engage in real-time dialogue with my "inner child"some reflections below:. Retrieved January 23, 2023 from https://twitter.com/michellehuang42/status/1597005489413713921?lang=en

[40] Ming-Hui Huang and Roland T Rust. 2018. Artificial intelligence in service. *Journal of Service Research* 21, 2 (2018), 155–172.

[41] Dave Itzkoff. 2016. *How 'Rogue One' Brought Back Familiar Faces.* Retrieved September 3, 2021 from https://www.nytimes.com/2016/12/27/movies/how-rogue-one-brought-back-grand-moff-tarkin.html

[42] Yevgen Ivanov. 2017. *Us9959497b1 - system and method for using a digital virtual clone as an input in a simulated environment.* Retrieved September 7, 2021 from https://patents.google.com/patent/US9959497B1/en

[43] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 9 (2019), 389–399.

[44] Jin-man Kim, Jong-woo Kim, and Yoon-mi Cho. 2020. *Meeting You.* Retrieved September 7, 2021 from https://youtu.be/uflTK8c4w0c

[45] Violet Kim. 2020. *Virtual Reality, Real Grief.* Retrieved September 3, 2021 from https://slate.com/technology/2020/05/meeting-you-virtual-reality-documentary-mbc.html

[46] Yelim Kim, Mohi Reza, Joanna McGrenere, and Dongwook Yoon. 2021. *Designers Characterize Naturalness in Voice User Interfaces: Their Goals, Practices, and Challenges.* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445579

[47] Emmanuelle P Kleinlogel, Marion Curdy, João Rodrigues, Carmen Sandi, and Marianne Schmid Mast. 2021. Doppelganger-based training: Imitating our virtual self to accelerate interpersonal skills learning. *PloS one* 16, 2 (2021), e0245960.

[48] Effie Le Moignan, Shaun Lawson, Duncan A Rowland, Jamie Mahoney, and Pam Briggs. 2017. Has Instagram Fundamentally Altered the'Family Snapshot'?. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 4935–4947.

[49] Jaime Lorenzo-Trueba, Junichi Yamagishi, Tomoki Toda, Daisuke Saito, Fernando Villavicencio, Tomi Kinnunen, and Zhenhua Ling. 2018. The voice conversion challenge 2018: Promoting development of parallel and nonparallel methods. *arXiv preprint arXiv:1804.04262* (2018).

[50] Octavian M Machidon, Mihai Duguleana, and Marcello Carrozzino. 2018. Virtual humans in cultural heritage ICT applications: A review. *Journal of Cultural Heritage* 33 (2018), 249–260.

[51] Sophie Maddocks. 2020. 'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political'deep fakes. *Porn Studies* 7, 4 (2020), 415–423.

[52] Frances M Malpezzi and William M Clements. 1989. The double and the theme of selflessness in Kagemusha. *Literature/Film Quarterly* 17, 3 (1989), 202.

[53] Alice E. Marwick and danah boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13, 1 (2011), 114–133. https://doi.org/10.1177/1461444810365313

[54] Arion McNicoll and Nick Glass. 2014. *The technology Bringing Sinatra, Tupac back to life.* Retrieved September 3, 2021 from https://www.cnn.com/2014/01/08/tech/innovation/optical-technology-tupac-back-to-life/index.html

[55] Minas Michikyan, Jessica Dennis, and Kaveri Subrahmanyam. 2014. Can You Guess Who I Am? Real, Ideal, and False Self-Presentation on Facebook Among Emerging Adults. *Emerging Adulthood* 3, 1 (Apr 2014), 55–64. https://doi.org/10.1177/2167696814532442

[56] Milad Mirbabaie, Stefan Stieglitz, Felix Brünker, Lennart Hofeditz, Björn Ross, and Nicholas RJ Frick. 2021. Understanding collaboration with virtual assistants–the role of social identity and the extended self. *Business & Information Systems Engineering* 63, 1 (2021), 21–37.

[57] Kana Misawa and Jun Rekimoto. 2015. ChameleonMask: a human-surrogate system with a telepresence face. In *SIGGRAPH Asia 2015 Emerging Technologies*. 1–3.

[58] Kana Misawa and Jun Rekimoto. 2015. Wearing another's personality: A human-surrogate system with a telepresence face. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. 125–132.

[59] Seyed Hamidreza Mohammadi and Alexander Kain. 2017. An overview of voice conversion systems. *Speech Communication* 88 (2017), 65–82.

[60] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

[61] Andreea Muresan and Henning Pohl. 2019. Chats with bots: Balancing imitation and engagement. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[62] MyHeritage. 2021. *Deep Nostalgia: Animate your family photos.* Retrieved September 10, 2021 from https://www.myheritage.com/deep-nostalgia

[63] Taewoo Nam. 2019. Citizen attitudes about job replacement by robotic automation. *Futures* 109 (2019), 39–49.

[64] Lisa P Nathan, Batya Friedman, Predrag Klasnja, Shaun K Kane, and Jessica K Miller. 2008. Envisioning systemic effects on persons and society throughout interactive system design. In *Proceedings of the 7th ACM conference on Designing interactive systems*. 1–10.

[65] Olivia B Newton and Mel Stanfill. 2020. My NSFW video has partial occlusion: deepfakes and the technological production of non-consensual pornography. *Porn Studies* 7, 4 (2020), 398–414.

[66] Neelkumar P Patel, Devangi R Parikh, Darshan A Patel, and Ronak R Patel. 2019. AI and web-based human-like interactive university chatbot (UNIBOT). In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 148–150.

[67] David L Penn, James D Ivory, Abigail Judge, et al. 2010. The virtual doppelganger: Effects of a virtual reality simulator on perceptions of schizophrenia. *The Journal of nervous and mental disease* 198, 6 (2010), 437–443.

[68] Holly Maxwell Pringle. 2015. Conjuring the ideal self: An investigation of self-presentation in video game avatars. *Press Start* 2, 1 (2015), 1–20.
[69] Jason Rohrer. 2020. *Project December*. Retrieved September 3, 2021 from https://projectdecember.net/
[70] Jennifer Rothman. 2018. The right of publicity: Privacy reimagined for New York. *Cardozo Arts & Ent. LJ* 36 (2018), 573–600.
[71] Nithya Sambasivan, Amna Batool, Nova Ahmed, Tara Matthews, Kurt Thomas, Laura Sanely Gaytán-Lugo, David Nemer, Elie Bursztein, Elizabeth Churchill, and Sunny Consolvo. 2019. " They Don't Leave Us Alone Anywhere We Go" Gender and Digital Abuse in South Asia. In *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
[72] Ovul Sezer. 2022. Impression (mis)management: When what you say is not what they hear. *Current Opinion in Psychology* 44 (2022), 31–37.
[73] David M Shaw. 2019. Defining Data Donation After Death: Metadata, Families, Directives, Guardians and the Route to Big Consent. *The Ethics of Medical Data Donation* (2019), 151–159.
[74] Alexandra Sherlock. 2013. Larger than life: Digital resurrection and the re-enchantment of society. *The Information Society* 29, 3 (2013), 164–176.
[75] Mel Slater, Solène Neyret, Tania Johnston, Guillermo Iruretagoyena, Mercè Álvarez de la Campa Crespo, Miquel Alabèrnia-Segura, Bernhard Spanlang, and Guillem Feixas. 2019. An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific reports* 9, 1 (2019), 1–13.
[76] Frederic Stutzman and Woodrow Hartzog. 2012. Boundary regulation in social media. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 769–778.
[77] Vivian Ta, Caroline Griffith, Carolynn Boatfield, Xinyu Wang, Maria Civitello, Haley Bader, Esther DeCero, and Alexia Loggarakis. 2020. User experiences of social support from companion chatbots in everyday contexts: Thematic analysis. *Journal of medical Internet research* 22, 3 (2020).
[78] Lee Taber and Steve Whittaker. 2020. *"On Finsta, I Can Say 'Hail Satan'": Being Authentic but Disagreeable on Instagram*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376182
[79] Jon Truby and Rafael Brown. 2021. Human digital thought clones: the Holy Grail of artificial intelligence for big data. *Information & Communications Technology Law* 30, 2 (2021), 140–168.
[80] Cristian Vaccari and Andrew Chadwick. 2020. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society* 6, 1 (2020), 2056305120903408.
[81] William D Weisman and Jorge F Peña. 2021. Face the Uncanny: The Effects of Doppelganger Talking Head Avatars on Affect-Based Trust Toward Artificial Intelligence Technology are Mediated by Uncanny Valley Perceptions. *Cyberpsychology, Behavior, and Social Networking* 24, 3 (2021), 182–187.
[82] Mika Westerlund. 2019. The emergence of deepfake technology: A review. *Technology Innovation Management Review* 9, 11 (2019).
[83] Daricia Wilkinson, Öznur Alkan, Q Vera Liao, Massimiliano Mattetti, Inge Vejsbjerg, Bart P Knijnenburg, and Elizabeth Daly. 2021. Why or why not? The effect of justification styles on chatbot recommendations. *ACM Transactions on Information Systems (TOIS)* 39, 4 (2021), 1–21.
[84] John Wojewidka. 2020. The deepfake threat to face biometrics. *Biometric Technology Today* 2020, 2 (2020), 5–7.
[85] Richmond Y Wong and Vera Khovanskaya. 2018. Speculative design in HCI: from corporate imaginations to critical orientations. In *New Directions in Third Wave Human-Computer Interaction: Volume 2-Methodologies*. Springer, 175–202.
[86] Richmond Y Wong, Ellen Van Wyk, and James Pierce. 2017. Real-fictional entanglements: Using science fiction and design fiction to interrogate sensing technologies. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 567–579.
[87] Jun Xiao, John Stasko, and Richard Catrambone. 2007. The role of choice and customization on users' interaction with embodied conversational agents: effects on perception and performance. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1293–1302.
[88] Lotus Zhang, Lucy Jiang, Nicole Washington, Augustina Ao Liu, Jingyao Shao, Adam Fourney, Meredith Ringel Morris, and Leah Findlater. 2021. Social Media through Voice: Synthesized Voice Qualities and Self-presentation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–21.
[89] Xuan Zhao, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Voida, and Dan Cosley. 2013. The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1–10.

A SPECULATIVE SCENARIOS

Below is a copy of the text portion of the speculative scenarios presented to participants during the asynchronous survey portion of the study as described in Section 3.1. The visual stimuli that accompanied each of these text-based prompts can be found in the supplementary materials.

A.1 Scenario 1: Retail chatbot

Claire is the owner of an Instagram shop that sells clothing in limited quantities. Her page has been in the Featured section of the app recently and her page is getting a lot of visitors. Her customers message her constantly on Instagram, hoping to buy the clothing she sells before they sell out.

Claire can't keep up with the number of messages coming in at all times of day, so Instagram offers to help create a chatbot version of Claire based on her previous Instagram messages with customers This helps Claire manage her Instagram shop better.

When one customer messages Claire and asks if she can ship her clothes to the Philippines, Claire's chatbot does not have a previous conversation to refer to. Instead it looks in a database of other Instagram conversations and tells the customer "we don't ship to the Philippines because we don't make as much money". The real Claire is now in a lot of trouble, and Instagram even threatens to suspend her page due to customer reports.

A.2 Scenario 2: Indigenous elder

Bill is an Indigenous elder with lots of important cultural knowledge about Indigenous ways of life. To help preserve this cultural knowledge for future generations, sociology researchers at a large university ask Bill if they can create a virtual reality version of him for students, researchers, and members of the public to interact with.

To do this, Bill has his body recreated using 3D body scanning and motion-capture technology and his voice data is used to create an interactive VR human surrogate. The interactive Bill surrogate is placed in a museum for people to interact with, preserving important Indigenous knowledge in a new medium.

A.3 Scenario 3: Non-consensual deepfakes

Avery is a university track and field athlete, and has many images and videos of themselves on Instagram and Twitter. A hacker takes the images and videos of Avery training and promoting their school team to instead create deepfake pornography, which they then use to blackmail Avery, threatening to send it to their potential employers and team members. Avery reports this to the police, but is told that there is little that can be done to stop them other than telling social media sites to take it down when they see it.

A.4 Scenario 4: Call centre agent

Matthew works as an agent at a call centre. He is told that the company will be rolling out a new chatbot service that is capable of responding to basic questions people have before redirecting them to a live agent. To make the chatobts more efficient, it is proposed that the chatobts are trained on some of the text messages of the call centre agents, including Matthew.

At first, the chatbots do a good job of replicating Matthew's messages and help decrease his workload with basic questions people have. As the chatbot gets better, and better, Matthew is told that the company wants to turn it into a voice-based call centre agent as well, and they showcase the technology to Matthew and the other agents (see video below).

Matthew is told that to make the voice-based agent a success, it will need to be trained on his audio data. Matthew does not feel particularly good about having a call centre agent based on his

voice, but also remembers that his performance evaluation is coming up. He is worried that not participating in the program might affect his evaluation, and is unsure of what to do.

A.5 Scenario 5: VR reunion

Sophie is the mother of her recently deceased 7-year old daughter Anna, who passed away in a car accident. During one of her grief counselling sessions, she mentions to her therapist that she has repeated dreams of Anna, where she is "hiding behind a tree, never smiling". Sophie expresses the desire to apologize to Anna and see her smile at her again, but says she is not able to have this dream.

The grief counselling team proposes recreating a scenario in Virtual Reality (VR) for Sophie, where Sophie and Anna are able to spend some final moments together so that Sophie may have closure. The counselling team works with Sophie's family to gather photos, audio data, and videos of Anna. Sophie undergoes the treatment, which involves an hour long session with the surrogate form of her daughter Anna. They have a picnic together and Sophie is finally able to see Anna smile at her again, giving Sophie the chance to obtain some closure over the matter. Footage of the session is available here: (link). Please watch up to 0:45.

After the treatment session, the counselling team offers another session for Sophie to spend time with Anna. Sophie declines, citing that she does not want to remember only the VR version of her daughter.

A.6 Scenario 6: Significant others

Mary and Liam have been together since high school. At the age of 21, Liam is going to be conscripted into his country's mandatory military service program and stationed in an underwater naval training program for 2 years. Since there is no access to outside communication during his time underwater, the couple is facing a strain in their relationship.

To maintain emotional and physical intimacy, someone suggests creating virtual forms of one another. This virtual version of 'Mary and Liam' will resemble each other's appearance and personality, and more importantly, their virtual forms will also be able to support sexual experiences between the couple over the course of Liam's service. Mary is hesitant to go through with this at first, but eventually agrees to do so.

A.7 Scenario 7: Physical swap

Julie is 25 years old, and grew up in a middle class family. Since middle school, Julie has been a member of the school soccer team and played as a midfielder until she graduated high school.

During the lockdown weeks of the pandemic, Julie re-kindled her passion with soccer and started a YouTube channel to create soccer commentary. At first Julie wanted to simply record her reactions to various games, but found she was uncomfortable with using her actual appearance online. To address this, Julie instead decides to use a surrogate creation tool to input her appearance, voice, and personality, creating a virtual host to represent herself. This virtual version of "Julie" helps her to deliver the content while keeping her actual appearance away from viewers.

For the first few weeks, Julie saw her channel grow super slowly, and only her close friends and family would watch her videos. After watching other channels that publish similar content, she found that while her content was similar, most of these channels had male hosts. She feels tempted to change her virtual host to a male figure with more masculine features to grow the channel, but also feels like this would be a betrayal of her identity.

A.8 Scenario 8: Commodifying grief

Jane's fiancée Mark has recently passed away after several years of battling pancreatic cancer. She finds going to work and trying to keep her life in order has become increasingly difficult while dealing with her loss. A friend recommends a local company that has been helping people like Jane by providing them a chatbot surrogate of their loved ones, and encourages Jane to give the service a try.

Jane enrolls with the service, who create a surrogate chatbot of Mark based on Jane's text message conversations with Mark. This helps Jane manage her loss in a better way, but she is annoyed that she occasionally has to watch ads to continue talking to the chatbot form of Mark. However, being able to reconnect with Mark this way is more important to Jane, so she puts up with the ads.

As Jane spends an increasing amount of time with the chatbot form of Mark, the company proposes expanding Mark with voice-based capabilities so Jane can call Mark for a monthly fee. Jane decides to give this part of the program a try, seeing as the monthly fee is not too expensive. Jane continues spending an increasing amount of time with the surrogate form of Mark, and before long, the company proposes a virtual reality recreation of Mark's appearance so Jane can spend more time with him, this time for a much higher fee. Jane's friends grow concerned about the way she is dealing with her loss.

Received January 2022; revised July 2022; accepted November 2022