AI in Your Mind: Counterbalancing Perceived Agency and Experience in Human-AI Interaction

Angel Hsing-Chi Hwang
hh695@cornell.edu
Cornell University
Ithaca, NY, USA

Andrea Stevenson Won
a.s.won@cornell.edu
Cornell University
Ithaca, NY, USA

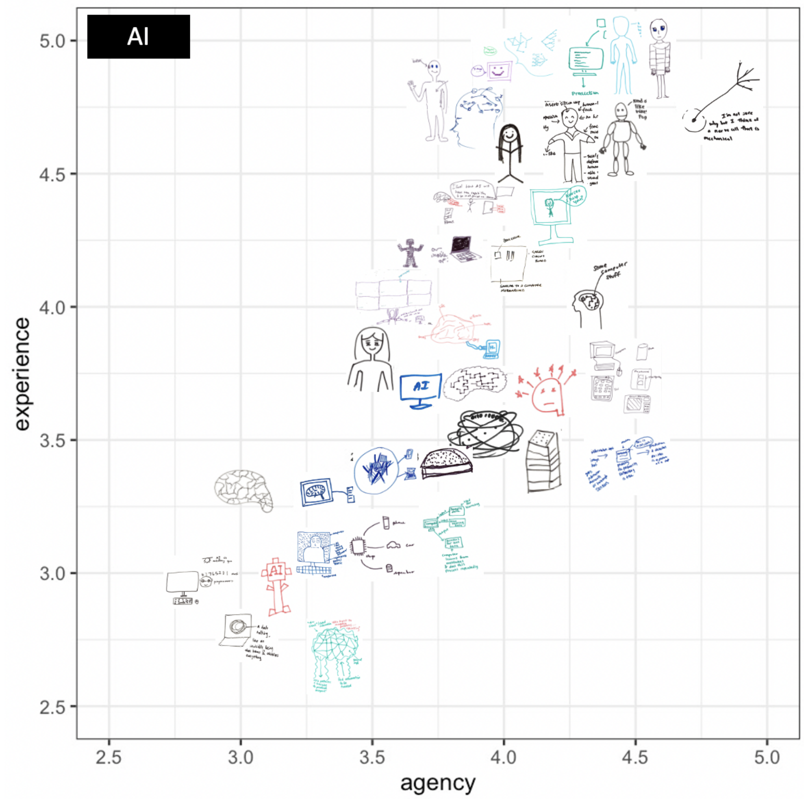


Figure 1: Participants' drawings and mind perception of AI.

ABSTRACT

In this mixed-methods study, we attempt to capture users' conception of AI through the two-dimensional mind perception framework (perceived agency vs. experience) in cognitive psychology [13] and a series of drawing tasks. Our data illustrate how participants perceive AI entities with physical embodiment, depicting AI through devices, imaginary human figures, or full techno-ecosystems. Furthermore, we apply users' mind maps of AI entities to highlight risks in human-AI interaction (HAII) and propose design solutions accordingly. We posit HAII research and development should be cognizant that users possess existing AI images and should exploit them as starting points for design improvement.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**.

KEYWORDS

human-AI interaction, mind perception, physical embodiment

ACM Reference Format:
Angel Hsing-Chi Hwang and Andrea Stevenson Won. 2022. AI in Your Mind: Counterbalancing Perceived Agency and Experience in Human-AI Interaction. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3491101.3519833

CHI '22 Extended Abstracts, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9156-6/22/04…$15.00
https://doi.org/10.1145/3491101.3519833

How would you draw a portrait of AI? Whether you envision a human face, a mechanical robot, or an abstract object, this vision reveals your conception of AI. How humans interact with AI-mediated entities is deeply affected by their perceptions [21, 23, 36, 39]. For example, when AI systems are designed with humanoid, social cues, users can be easily distracted from the reliability and trustworthiness of an AI system [35, 37, 38, 42, 43].

In this exploratory study, we ask participants to visualize their perceptions of AI through a drawing task. While users' impressions of AI can vary by device, brand, and context, our goal here is to inform a preliminary "baseline" of human perceptions of AI. If we see users' perceptions of AI as being predominantly human-like, this warns us that users may be applying inaccurate heuristics to AI interactions. If most participants view AI as tools, this may lead us to ask how incorporating considerations of embodiment can help build trust and synergy in human-AI collaboration. Applying the mind perception framework in cognitive psychology [13], we discuss users' behavior and design implications in human-AI interaction (HAII), given their perceived images of AI-mediated entities.

1 BACKGROUND & RELATED WORK

1.1 Mind perception and behavior in HAII

Gray et al.'s influential work, *Dimensions of Mind Perception* [13], proposes that how humans perceive other entities internally and how such conceptions translate to external behaviors can be measured through two dimensions: **agency** (the ability to do) and **experience** (the ability to feel). Entities perceived as high agency and high experience are considered to be more human-like and similar to the self, while low agency, low experience entities are less humanoid and more objectified [14, 39, 40]. Moreover, when a character is perceived with more experience than agency (e.g., an infant), individuals view that subject as more benign and vulnerable, and demonstrate greater tendencies to protect, forgive, and "come to the rescue of" that entity. [4, 17]. Conversely, for those viewed as having more agency than experience (e.g., a political leader), people show less empathy toward those agentic doers and perceive them as less sincere [16]. Indeed, previous work has found that participants tended to perceive greater agency than experience in a physical robot and thus may even punish it for moral wrong-doing [22, 23]. This raises the question of whether today's users have already developed their own mind perception standards toward intelligent agents.

A popular approach in today's HAII research defines AI through technical capabilities [1, 44]. While this approach resonates with the concept of agency in the mind perception framework, this working definition may fail to capture the full picture of human perceptions of AI and how these may influence behavior. As discussed above, individuals' interactions with others are not only determined by the absolute level of their agency, but more importantly, the relationship between agency vs. experience. Under the mind perception framework, individuals are more likely to trust the vulnerable regardless of their ability, while they are less willing to comply with a capable but less empathetic actor. In this regard, solely pursuing technical advances may not be sufficient to build pleasant HAII experiences, let alone support effective human-AI collaborations.

1.2 Counterbalancing agency and experience through design

A key challenge then concerns how we can balance the perceived agency and experience of AI entities to create user-friendly experiences while diminishing the risk of overlooking deficits in HAII. Recent research has demonstrated that users' perceptions of capability and trustworthiness of AI are malleable [11, 25, 28]. For instance, participants' perceived credibility of AI applications can be significantly moderated through framing [20, 24, 25] or offering explanations [2, 7]. Though limited work in HAII has attempted to manipulate the perception of experience in AI, the notions of building "social robots" [6, 10] or even "robots with soul" [5, 19] have long been promoted in the field of human-robot interaction (HRI). Scholars have proposed embedding anthropomorphic cues in robot design (e.g., adding facial expressions, showing nonverbal behavior, and sensing micro emotions) [5, 19]. On the other hand, while the emerging research area of embodied AI offers opportunities for control over the physical appearance of AI entities [29], ethical and security issues have raised questions of whether humanoid design remains an appropriate design strategy for future AI-mediated entities [8, 12, 18].

In light of these efforts to manipulate perceived agency and experience of AI-mediated entities, we propose the question: **How do people visualize AI?** In abundant explainable AI research, the assumption is that users' perceptions of AI agency are not high enough to make it trustworthy [2, 31, 33, 41]. The motivation to design anthropomorphic robots implies that the perceived experience of mechanical entities is too low [5]. However, it remains unknown whether these assumptions reflect the reality of users' mind perception of AI, nor how perceptions of agency and experience may interact. To address this question, we utilize Gray et al.'s mind perception model coupled with a drawing prompt to capture and analyze users' mind perception toward current AI. Furthermore, we apply the mind map as a tool to inform HAII design.

2 METHODS AND MATERIALS

2.1 Participants and Procedures

The present research recruited N = 36 student participants from the Communication and Information Science student participant pool (SONA) at Cornell University, allowing us to recruit student participants both from STEM and non-STEM background. All participants successfully completed the study, and were included the sample. Nine participants identified as male and 27 participants identified as female. The mean age of the participants was 21.00 ($S.D.$ = 3.16). The self-reported ethnicities of our participants are reported in Appendix A (Table 5). After consenting, participants completed a survey rating the perceived agency and experience of 7 entities (Self, AI, Apple Siri, Amazon Alexa, Google Assistant, Computer, and Robot). Similar to [30], we chose three of the most common AI agents in consumer products (i.e., Siri, Alexa, and Google Assistant) to compare whether brand identities further influence users' perception of AI entities. They were also asked to draw a portrait for each entity, as follows: "Please take a piece of paper on the table and draw a portrait of [Name of the entity]. You can also add text annotations to help elaborate what you have in mind." We did not provide any other instructions, i.e., we did not explicitly instruct participants to visualize AI entities through their physical devices. All rating and drawing questions were presented in a randomized order. After completing all the drawing and rating questions, participants filled out a short exit survey, reporting demographic data, knowledge related to AI and previous experiences interacting with AI-mediated agents.

2.2 Measurement

2.2.1 Mind perception. We adopted the original scales of mind perception measurement [13], including 7 items measuring the degree of agency ($M = 3.59; S.D. = .95$) and 11 items measuring the dimension of experience ($M = 3.75; S.D. = 1.12$). We followed the original study protocol by asking participants to perform pair-wise comparisons using 5-point Likert scales. That is, for each set of rating questions, participants saw two randomly paired entities and rated which of the two entities were more capable of performing each of the items described (see Figure 5 in Appendix). Each participant completed pair-wise comparisons for all possible pairs, i.e., 21 sets of mind perception ratings.

2.2.2 Drawings of entities. Participants were provided with a pile of letter size paper and a box of 12 colored markers. They were asked to specified the entity they were drawing on the corner of each piece of portrait to ensure the authors would not mistake the entity of each portrait. Each participant drew all 7 entities. In total, we collected 252 drawings.

2.2.3 Domain knowledge. We asked participants to self-evaluate their level of expertise in domains related to AI, including (1) artificial intelligence, (2) computer science, (3) robotics, and (4) machine learning, using 7-point Likert scales (1 = Not knowledgeable at all; 7 = Extremely knowledgeable). Overall, the majority of participants did not consider themselves as domain experts on these relevant areas ($M = 3.62; S.D. = 1.53$).

2.2.4 Prior experiences. We asked participants whether they have used and interacted with the three branded AI-mediated entities (Siri, Alexa, and Google Assistant). All participants used at least one of them previously. Furthermore, we asked them to estimate the frequency of interacting with these AI entities when they used a device with such applications on a 7-point Likert scales (e.g., how frequently they would call Siri up when they were using an Apple device; 1 = Never, 4 = Half the time, 7 = Always; $M = 1.97; S.D. = .81$).

2.2.5 Impression of AI in popular media. Finally, we asked participants to describe a memorable piece of content from popular media (e.g., movies, online videos, social media content) that influences their impression of AI using open-ended text. Furthermore, we asked participants to rate the "uncanniness" of the AI entity they described using existing 7-point Likert-scales [15] ($M = 4.11; S.D. = 2.03$).

2.3 Data Analysis

We began our analysis by coding the drawings qualitatively before looking at the survey responses. Applying an open coding approach [9], each author went through all the drawings independently, and took notes on observed themes. We then regrouped to discuss how themes identified by the two coders converged and/or differ. Together, we developed a qualitative codebook for the drawing data (see Figure 7 in Appendix). Next, we plotted participants' drawings of each entity according to their average survey ratings on the dimensions of agency and experience [13]. This allowed us to examine whether participants' perceived agency and experience of these entities reflect on their envisioned, visualized forms. Finally, we explored whether participants' domain knowledge, prior experiences, and impressions of AI influenced their mind perception and how these factors related to with their drawings' positions on the mind maps.

3 RESULTS

3.1 Overview of AI mind perception & physical embodiment

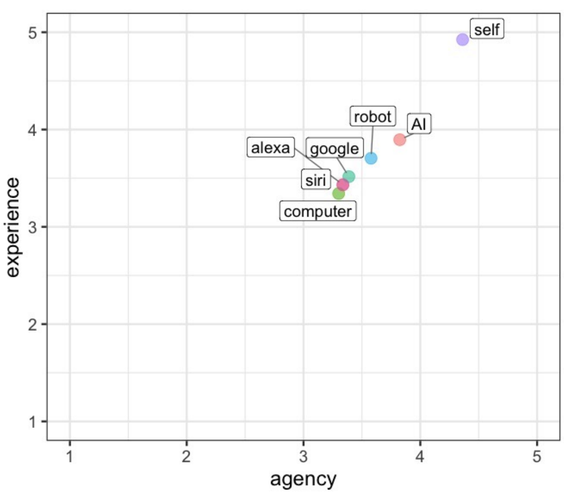


Figure 2: Mind perception of all entities.

An overview of participants' drawing of AI is shown in Figure 1 and the mind perception ratings of the seven entities are shown in Figure 2. Replicating previous literature, participants rated themselves as having highest agency and highest experience, followed by AI and then Robot. Mind perceptions of the three branded AI assistants were nearly identical, while Computer was perceived as having the least "mind". Through an one-way ANOVA test,[1] we verified that the degree of agency ($F = 10.60, p < .001$) and experience ($F = 8.14, p < .001$) differs significantly by entities. Following up with a Tukey HSD test to examine pair-wise difference, we saw AI rated as marginally more agentic and experiential than Robot, and significantly greater in both dimensions than Siri, Alexa, and Google.

Participants' drawings of AI (Figure 1) and AI-mediated entities vary by the extent of embodiment, personification, and their relationships with physical devices. **Despite the term AI *per se* being**

[1]As the ratings for "self" are much higher than all other entities, we performed the ANOVA test both including and excluding the "self" category (i.e., one test comparing across all entities and one test comparing among non-human entities). However, the results remain the same. The statistics here report test results among non-human entities.

an abstract concept, participants associated it with physical forms and objects. In fact, only 3 out of all drawings stated AI "has no tangible form." Based on our initial review of the drawings, we developed a coding scheme, informed by existing literature [3], to evaluate whether participants' visions of AI included the following forms of embodiment: body schema (i.e., showing body parts, such as limbs and/or torso as in human figures), emotion (i.e., showing emotional cues such as through facial expression), and social identity (i.e., showing cues that represent one's identities, such as gender or race). Based on participants' drawings, we also coded for "voice" in the drawings, as that is also a common medium for AI embodiment.[2] Indeed, participants' drawings of AI entities show how they can be embodied in various forms (body schema: 23.2%, emotion: 18.8%, social: 15.2%, voice: 47.8%) and can own more than more types of embodiment. However, participants who vision AI entities as "more embodied" (i.e., being embodied through more forms and tangible elements) did not perceive them as more agentic (Pearson's $\rho = -.17, p = .05$) or having more experience (Pearson's $\rho = -.09, p = .30$)[3]. A large portion of participants associated AI entities with specific devices (*Device-based embodiment*; see more in Section 3.3.1). Among participants who envisioned AI entities *without* specific devices (*Device-free embodiment*; see more in Section 3.3.2), mind perception for AI shows significantly higher agency ($t = 3.99, p < .001$) and experience ($t = 3.97, p < .001$). In drawings of AI entities (but not in the drawings of Robot and Computer), some participants considered their interactions with AI entities to extend beyond one-on-one communication with the AI *per se*. This allowed us to distinguish two key themes in participants' drawings —they may present **AI by and of itself** or demonstrate the processes, infrastructure, and even organizations that go into the loop of AI development and implementation, which we call **infrastructure-in-the-loop**. Below, we elaborate further on these two concepts.

3.2 AI in techno-ecosystem: Infrastructure-in-the-Loop

A handful of drawings (17%) illustrate AI entities through a sequence or network of events. In essence, participants denote the path from inputs (e.g., images) to outputs (e.g., prediction, classification, pattern identification), accompanied by some or all of the following elements: Depiction of users offering data through devices and/or interfaces, which get stored in certain data management systems or, specifically, "the cloud." Such storage is often linked to individuals (e.g., programmers) and/or organizations (e.g., tech companies) that construct and implement AI applications. Eventually, the system sends output decisions and/or actions back to users. In this view, human-AI interaction is beyond individual human and AI. Instead, the full path of interactive experience is grounded on and supported by a holistic techno-ecosystem and establishes a one(user)-to-many(system) relationship. For instance, when users call up Google, they are interacting with "a group of men connected to help."

[2]Again, two authors coded the elements of physical embodiment in each drawing independently. The inter-rater reliability is 89.9%.

[3]The Pearson's coefficients here represent the correlation between the number of embodied cues (through body schema, emotion, social identity, and/or voice) versus ratings for perceived agency and experience respectively

Dependent variable	df	Sum Sq	Mean Sq	F-value	p-value
Experience	5	47	9.31	8.14	< .001***

Table 1: ANOVA test on the relationship between perceived experience and entities

	AI	Alexa	Computer	Google	Robot
Alexa	< .001***	-	-	-	-
Computer	< .001***	.951	-	-	-
Google	.003**	.966	.539	-	-
Robot	.432	.088	.006**	.444	-
Siri	< .001***	1.000	.955	.963	.084

Table 2: Tukey HSD test on the significance of pair-wise difference in perceived experience between entities

Dependent variable	df	Sum Sq	Mean Sq	F-value	p-value
Experience	5	47	8.95	10.6	< .001***

Table 3: ANOVA test on the relationship between perceived agency and entities

	AI	Alexa	Computer	Google	Robot
Alexa	< .001***	-	-	-	-
Computer	< .001***	.998	-	-	-
Google	< .001***	.993	.922	-	-
Robot	.062	.071	.022*	.263	-
Siri	< .001***	1.000	.999	.992	.070

Table 4: Tukey HSD test on the significance of pair-wise difference in perceived agency between entities

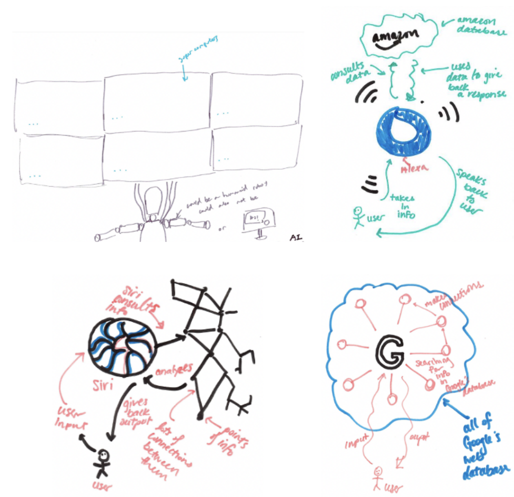


Figure 3: Examples of drawings of AI as "infrastructure in the loop."

We further asked whether this sophisticated view of AI is specific to more knowledgeable and experienced users. However, participants who drew infrastructure-in-the-loop did not self-identify as more AI savvy ($t = -.61, p = .54$) nor more experienced ($t = .96, p = .34$) than those who drew AI alone. Finally, we examine how this view of AI maps to mind perception. As shown in the blue area *below* the diagonal in Figure 4, those who portrayed AI in a user-to-ecosystem relationship assigned greater agency than experience to AI-mediated entities.

3.3 AI by and of itself

Among those who visualized AI as an individual entity, we saw a noticeable gap in the degree of anthropomorphism. Participants either presented AI entities by illustrating devices (e.g., iPhone, Amazon Echo) or they drew humanoid representations—more specifically, female figures —of AI, Siri, Alexa, and Google. However, this does *not* indicate the non-humanoid presentations of AI were considered as less personified. Through other forms of embodiment and sensory modalities (e.g., voice), participants described AI as "having emotions and personalities." For example, participants mentioned Siri as "funny," "approachable," and "more friendly," as "she tells better jokes," while Alexa is "more put together." Below, we break down the two forms of presentation to discuss **device-based vs. device-free embodiment**.

3.3.1 Device-based embodiment: AI lives in devices. Around half of all drawings (47%) visualized AI and AI-mediated entities through devices (e.g., computers, smartphones, tablets, smart home devices, etc.). Through their drawings and notations, participants conveyed the idea that AI entities "live" or "hide" in these devices, where they "do all the thinking." In other words, the capability to feel (experience) and do (agency) is embedded in the devices as such. More specifically, participants described Alexa as "AI being trapped in a gadget" and that "Siri lives in iPhone." Distinct from previous embodiment literature [3], where different forms of embodiment are leveled in a continuous fashion (having body schema → having emotions → having social identity), here, various types of embodiment are independent. For example, participants perceived the "positivity" and gender identity of Siri without ever seeing "her" smiley facial face. On the mind map, these drawings fall along the diagonal and near the center of the plot, which means device-based AIs are perceived as equally agentic and experiential, with both dimensions rated at moderate levels.

3.3.2 Device-free embodiment: AI owns its face. On the other extreme, we saw that participants came up with their own visions of AI entities (in 13.4% of all drawings). Unlike the illustrations for Robot, which were "sort of human-looking in structure, but not really" (see more in the next section), these images of AI are completely "human-like." Some are further annotated with demographic details (e.g., "Siri is 25 years old"). In this regard, we examine whether these humanoid representations of AI are also perceived as having more human-like minds (i.e., having high experience and high agency). We saw these drawings fall in the triangular area *above* the diagonal of the mind map (see Figure 4), which indicates participants tended to perceive AI and AI-mediated entities as having more experience than agency. Again, we ask whether the tendency to imagine AI "with a face" is elicited by certain characteristics of the participants. But once again, the participants who rendered AI as fully humanoid did not show greater domain knowledge nor more prior experience than those who did not do so.

3.4 Contrasting AI: Human has hair, Robot lacks a social role, and Computer is "just a computer"

In this part of the analysis, we contrast general and branded AI entities to other characters in the present study (i.e., participants themselves, Robot, and Computer). To begin with, we identified hair as an element that is crucial to constructing a human figure. In all self-portraits, participants clearly illustrated the color, style, and length of their hair. Even a bald participant who drew a portrait of themselves wearing a mask added facial hair to the figure. While eyes and mouths are commonly illustrated in drawings of all human(-like) figures, hair was only depicted in self-portraits and imagined humanoid representations of AI, but not in any of the Robot drawings. This resonates with the aforementioned comment, which suggests that robots have the structure of human body but not the details. Moreover, in drawings of Robot, the gender, social role, and technical capability of robots are unspecified, nor did we see any traits of interaction with human users (i.e., participants presented robots *in solo*). By contrast, participants frequently showed back-and-forth conversations and/or actions between agents and users in drawings of AI-mediated entities. Finally, drawings of Computer are the most consistent across all participants. These images feature the device alone, typically with a screen and keyboard. Participants did not show users interacting with computers nor did they specify their technical capacities.

3.5 AI in pop culture has more experience than agency

After drawing their own visions of AI entities, participants were asked to describe their impressions of AI in popular media, mostly inspired by movies (e.g., The Terminator, iRobot, Her), a few from social media (e.g., Lil Miquela), and news coverage of tech products (e.g., Siri, Alexa, and Google). According to participants, highly humanoid robots dominated images of AI in the popular cultural sector. Such entities are depicted as being capable of "mimic[king] human emotions, behavior, habits" and "whatever humans can do," which leads to two common story plots: Either AI "takes over humanity," or a man may "fall[s] in love with a robot who can feel him." While we did not see these dramatic elements reflected in participants' drawings (only 6 out of all drawings for AI entities include the theatrical components above), participants' ratings for AI in popular media showed higher experience than agency when AI was described as a character rather than a product (situated above the diagonal in the mind map, see Figure 6d in Appendix).

4 DISCUSSION

In this project, we applied the mind perception framework to capture participants' perceived images of AI beyond its technical capability. We found that users tended to imagine AI in three ways: (1) a physical device, (2) a humanoid figure, or (3) a full techno-ecosystem that affords the development and implementation of AI. Different from previous work, which found that mind perceptions suggested both high experience and high agency for personified characters but low agency and low experience in inanimate objects [17, 23, 40], participants assign more experience than agency when associating AI with a human face and body, while emphasizing agency over experience when including infrastructures in the loop. We also see that participants' perceptions of the humanoid, social identities of AI need not be based on lower-level, physical embodiment (e.g., AI does not need to be embedded in body schema to elicit

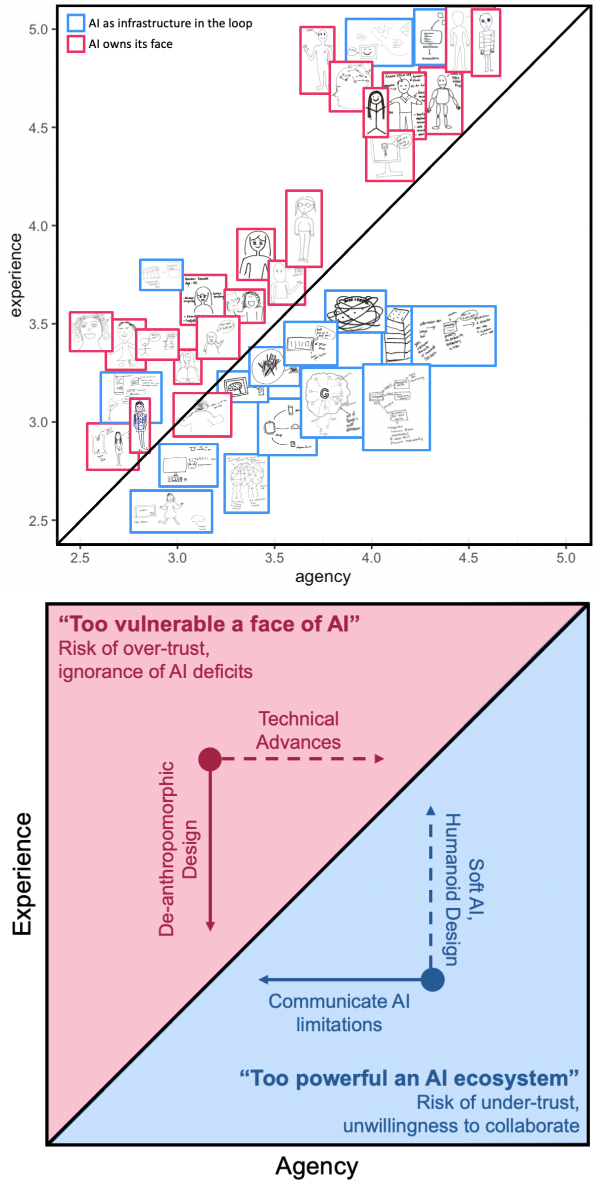


Figure 4: Top: Drawings of AI with imaginary humanoid figures (red outline) and infrastructure in the loop (blue outline). Bottom: Applying the mind map as a tool to diagnose risks in HAII and corresponding design solutions.

flesh and blood in users' minds). Below, we consider the conceptual and practical implications of our current findings.

4.1 Design guidelines: Mind map as a tool to inform HAII design

In this section, we discuss how the mind map from our study could be applied to inform HAII design strategies (Figure 4). As discussed in Section 1, when users' mind perception toward an AI-mediated entity falls above or below the diagonal line (imbalance between agency and experience), relevant HAII risks arise accordingly. Therefore, we use this tool to inform design decisions on which direction to move users' perceived agency and/or experience toward AI applications.

De-anthropomorphic design for "vulnerable AI" (experience > agency). When mind perception falls in the red triangular area above the diagonal (see Figure 4), individuals view an entity as vulnerable and benign —indeed, this is the area where humans' mind perception for children and infants falls [4, 13, 17]. Participants' drawings further show that they tend to envision highly humanoid images when they perceive AI in this fashion. Here, we see risks of users over-trusting AI and forgiving its deficits or improper usage [42]. To resolve these issues, we can either consider increasing perceived agency or decreasing perceived experience. To achieve the former, technical advances of AI provides an obvious path, though it likely requires more time and resources to accomplish. As a result, reducing humanoid cues in content design to reduce the perceived experience of an AI application can be a more feasible alternative [6]. In fact, given that users already hold highly experiential impression toward entities in this area, adding human touches (through visual and sound features) can increase the risk of deception (i.e., can cause users to mistakenly treat AI-mediated agents as overly human).

Highlight limitations of "powerful AI" (agency > experience). On the other hand, the blue triangular area below the diagonal (Figure 4) signals challenges for establishing trust and engagement. Furthermore, as users recall the full ecosystem that goes into the building and making of AI, their perception of agency tends to overpower the extent of perceived experience in machine-mediated entities here [15]. Recent HAII literature has proposed a number of effective ways to communicate the limitations of an AI system [1]. In this regard, clarifying what an AI application cannot afford offers a more promising and achievable solution to amend users' impression of overly powerful AI entities.

4.2 Empirical and theoretical implications

Because we did not show participants any visual references (i.e., we did not show participants any physical or virtual agent in the study, which is a common study setup in prior work), our study demonstrates that users possess their own "baseline" mind perception toward AI entities. This finding first raises the need to differentiate the effects of a specific agent's physical embodiment from users' existing perceptions toward that type of agent. This leads to several theoretical connections with previous research on users' mental models of AI. To begin with, we found that consumer products of AI (i.e., Siri, Alexa, and Google Assistant) received lower ratings in both perceived agency and experience than the more general concepts of AI and robot, aligning with Sundar's dual-process framework of the Theory of Interactive Media Effects for Human-AI Interaction (HAII-TIME, [34]). Specifically, this theoretical model suggests that when users view AI as a symbol, they impose their own beliefs about the capability of AI; but when users engage in direct interaction with AI entities, they base their evaluations on the AI's specific attributes. Therefore, the limitations of current AI products (e.g., given the current state-of-the-art of natural language processing, Siri cannot completely understand user speech) can easily reduce users' perceived capability of AI entities when they actually use and interact with them. According to the Stereotype Content Model, [27] individuals tend to view AI as more competent but less warm when it operates independently, while AI showing alignment with human interests is perceived as more friendly and socially close. While the two opposite scenarios resonate well with the images of AI as tech infrastructures and as humanoid figures in our study, this also implies perceived AI independence may help explain individuals' baseline impressions of intelligent machines. Considering the affective or experiential dimension, [32] has found participants may even view humans as "higher in goodness" but weaker than AI when carrying out technical tasks. Finally, it is worth noting that, as shown in Figure 2, all of the AI entities remain lower than human in both the agentic and experiential dimensions. Based on the Uncanny Valley, users' possessing images of AI that are humanoid to – but not quite human – can trigger particularly uneasy feelings during interactions [26]. Together, our current work adds to previous research to depicted a more comprehensive view of humans' impression of AI. Future work will investigate how these perceptions may influence users' willingness to interact with, trust, and collaborate with AI systems.

5 CONCLUSION, LIMITATIONS AND FUTURE WORK

This mixed-methods study captures users' mind perception of AI entities. The drawing data first show that participants model the physical embodiment of AI in their mind through devices, imaginary human figures, or infrastructures-in-the-loop. Informed by participants' drawings and mind perception ratings, we apply the mind map as a tool to pinpoint potential risks in HAII and explain how users' existing images of AI can help inform design solutions. Specifically, when AI is associated with a vulnerable figure (experience > agency), perceived experience of AI should be reduced through de-anthropomorphic design; when users view an AI product through a full, powerful AI ecosystem, conveying its limitations to reduce perceived agency can be a possible solution to enhance engagement. Beyond learning from the current small study, we encourage future work to empirically examine the effectiveness of our proposed design implications. Furthermore, given our current methodology, we acknowledge that our study design may have encouraged participants to concretize their conception of AI. Therefore, we encourage future studies to probe users' perception through other approaches, such as through language (e.g., interview) or neuroimaging. Finally, given the current mixed-methods design, we conducted the study with a restricted sample size, consisting of student participants who mostly identified as female. In this regard, we encourage future work to consider survey designs that allow researchers to study users' mind perceptions on a larger scale. In summary, users imagine a range of tangible forms of AI entities, and we posit HAII research and development should be cognizant of such existing AI images and exploit them as starting points for design improvement.

REFERENCES

[1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300233

[2] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. 2021. Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–16. https://doi.org/10.1145/3411764.3445717

[3] Frank Biocca. 2006. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. *Journal of Computer-Mediated Communication* 3, 2 (June 2006), 0–0. https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

[4] Angela S. Book, Vernon L. Quinsey, and Dale Langford. 2007. Psychopathy and the Perception of Affect and Vulnerability. *Criminal Justice and Behavior* 34, 4 (April 2007), 531–544. https://doi.org/10.1177/0093854806293554

[5] C. Breazeal, C.D. Kidd, A.L. Thomaz, G. Hoffman, and M. Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Edmonton, Alta., Canada, 708–713. https://doi.org/10.1109/IROS.2005.1545011

[6] Cynthia Breazeal and Brian Scassellati. 1999. A Context-Dependent Attention System for a Social Robot. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence - Volume 2* (Stockholm, Sweden) *(IJCAI'99)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1146–1151.

[7] Carrie J. Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–24. https://doi.org/10.1145/3359206

[8] Corinne Cath. 2018. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, 2133 (Nov. 2018), 20180080. https://doi.org/10.1098/rsta.2018.0080

[9] Juliet M. Corbin and Anselm L. Strauss. 2015. *Basics of qualitative research: techniques and procedures for developing grounded theory* (fourth edition ed.). SAGE, Los Angeles.

[10] Brian R. Duffy. 2003. Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42, 3-4 (March 2003), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

[11] Andrea Ferrario, Michele Loi, and Eleonora Viganò. 2020. In AI We Trust Incrementally: a Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions. *Philosophy & Technology* 33, 3 (Sept. 2020), 523–539. https://doi.org/10.1007/s13347-019-00378-3

[12] Amelia Fiske, Peter Henningsen, and Alena Buyx. 2019. Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research* 21, 5 (May 2019), e13216. https://doi.org/10.2196/13216

[13] Heather M. Gray, Kurt Gray, and Daniel M. Wegner. 2007. Dimensions of Mind Perception. *Science* 315, 5812 (Feb. 2007), 619–619. https://doi.org/10.1126/science.1134475

[14] Kurt Gray, Joshua Knobe, Mark Sheskin, Paul Bloom, and Lisa Feldman Barrett. 2011. More than a body: Mind perception and the nature of objectification. *Journal of Personality and Social Psychology* 101, 6 (2011), 1207–1220. https://doi.org/10.1037/a0025883

[15] Kurt Gray and Daniel M. Wegner. 2012. Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition* 125, 1 (Oct. 2012), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

[16] Kurt Gray and Daniel M. Wegner. 2012. Morality takes two: Dyadic morality and mind perception. In *The social psychology of morality: Exploring the causes of good and evil.*, Mario Mikulincer and Phillip R. Shaver (Eds.). American Psychological Association, Washington, 109–127. https://doi.org/10.1037/13091-006

[17] Kurt Gray, Liane Young, and Adam Waytz. 2012. Mind Perception Is the Essence of Morality. *Psychological Inquiry* 23, 2 (April 2012), 101–124. https://doi.org/10.1080/1047840X.2012.651387

[18] Arno Hartholt, Ed Fast, Adam Reilly, Wendy Whitcup, Matt Liewer, and Sharon Mozgai. 2019. Ubiquitous Virtual Humans: A Multi-platform Framework for Embodied AI Agents in XR. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, San Diego, CA, USA, 308–3084. https://doi.org/10.1109/AIVR46125.2019.00072

[19] Guy Hoffman and Wendy Ju. 2014. Designing Robots With Movement in Mind. *Journal of Human-Robot Interaction* 3, 1 (March 2014), 89. https://doi.org/10.5898/JHRI.3.1.Hoffman

[20] Yoyo Tsung-Yu Hou and Malte F. Jung. 2021. Who is the Expert? Reconciling Algorithm Aversion and Algorithm Appreciation in AI-Supported Decision Making. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–25. https://doi.org/10.1145/3479864

[21] Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. 2006. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in

human–robot interaction. *International Journal of Human-Computer Studies* 64, 10 (Oct. 2006), 962–973. https://doi.org/10.1016/j.ijhcs.2006.05.002

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

[23] Minha Lee, Peter Ruijten, Lily Frank, Yvonne de Kort, and Wijnand IJsselsteijn. 2021. People May Punish, But Not Blame Robots. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 715, 11 pages. https://doi.org/10.1145/3411764.3445284

[24] Mengqi Liao and S. Shyam Sundar. 2021. How Should AI Systems Talk to Users when Collecting their Personal Information? Effects of Role Framing and Self-Referencing on Human-AI Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445415

[25] Bingjie Liu. 2021. In AI We Trust? Effects of Agency Locus and Transparency on Uncertainty Reduction in Human–AI Interaction. *Journal of Computer-Mediated Communication* 26, 6 (Nov. 2021), 384–402. https://doi.org/10.1093/jcmc/zmab013

[26] Maya B Mathur and David B Reichling. 2016. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition* 146 (2016), 22–32.

[27] Kevin McKee, Xuechunzi Bai, and Susan Fiske. 2021. Understanding human impressions of artificial intelligence. (2021).

[28] Siddharth Mehrotra. 2021. Modelling Trust in Human-AI Interaction. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*. 1826–1828.

[29] Tønnes F. Nygaard, Charles P. Martin, Jim Torresen, Kyrre Glette, and David Howard. 2021. Real-world embodied AI through a morphologically adaptive quadruped robot. *Nature Machine Intelligence* 3, 5 (May 2021), 410–419. https://doi.org/10.1038/s42256-021-00320-3

[30] Anastasia K Ostrowski, Jenny Fu, Vasiliki Zygouras, Hae Won Park, and Cynthia Breazeal. 2021. Speed Dating with Voice User Interfaces: Understanding How Families Interact and Perceive Voice User Interfaces in a Group Setting. *Frontiers in Robotics and AI* 8 (2021), 730992–730992.

[31] Urja Pawar, Donna O'Shea, Susan Rea, and Ruairi O'Reilly. 2020. Explainable AI in Healthcare. In *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*. IEEE, Dublin, Ireland, 1–2. https://doi.org/10.1109/CyberSA49311.2020.9139655

[32] Daniel B Shank, Madison Bowen, Alexander Burns, and Matthew Dew. 2021. Humans are perceived as better, but weaker, than artificial intelligence: A comparison of affective impressions of humans, AIs, and computer systems in roles on teams. *Computers in Human Behavior Reports* 3 (2021), 100092.

[33] Donghee Shin. 2021. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies* 146 (Feb. 2021), 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

[34] S Shyam Sundar. 2020. Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication* 25, 1 (2020), 74–88.

[35] Nathan L. Tenhundfeld, Hannah M. Barr, Emily O'Hear, Andrew Atchley, and Jenna E. Cotter. 2021. Effects of Human-Likeness on Robot Use in High-Risk Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 65, 1 (Sept. 2021), 1049–1053. https://doi.org/10.1177/1071181321651061

[36] Giulia De Togni, Sonja Erikainen, Sarah Chan, and Sarah Cunningham-Burley. 2021. What makes AI 'intelligent' and 'caring'? Exploring affect and relationality across three sites of intelligence and care. *Social Science & Medicine* 277 (2021), 113874. https://doi.org/10.1016/j.socscimed.2021.113874

[37] Michelle M.E. van Pinxteren, Ruud W.H. Wetzels, Jessica Rüger, Mark Pluymaekers, and Martin Wetzels. 2019. Trust in humanoid robots: implications for services marketing. *Journal of Services Marketing* 33, 4 (Aug. 2019), 507–518. https://doi.org/10.1108/JSM-01-2018-0045

[38] Alan R. Wagner, Paul Robinette, and Ayanna Howard. 2018. Modeling the Human-Robot Trust Phenomenon: A Conceptual Framework based on Risk. *ACM Transactions on Interactive Intelligent Systems* 8, 4 (Nov. 2018), 1–24. https://doi.org/10.1145/3152890

[39] Adam Waytz, Kurt Gray, Nicholas Epley, and Daniel M. Wegner. 2010. Causes and consequences of mind perception. *Trends in Cognitive Sciences* 14, 8 (Aug. 2010), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

[40] Daniel M. Wegner and Kurt Gray. 2017. *The mind club: who thinks, what feels, and why it matters.* Penguin Books, London.

[41] Katharina Weitz, Dominik Schiller, Ruben Schlagowski, Tobias Huber, and Elisabeth André. 2019. "Do you trust me?": Increasing User-Trust by Integrating Virtual Agents in Explainable AI Interaction Design. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. ACM, Paris France, 7–9. https://doi.org/10.1145/3308532.3329441

[42] Jin Xu. 2018. Overtrust of Robots in High-Risk Scenarios. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New Orleans LA USA, 390–391. https://doi.org/10.1145/3278721.3278786

[43] Jin Xu and Ayanna Howard. 2018. The Impact of First Impressions on Human-Robot Trust During Problem-Solving Scenarios. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Nanjing, 435–441. https://doi.org/10.1109/ROMAN.2018.8525669

[44] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376301

6 APPENDIX

A SELF-REPORT ETHNICITY OF PARTICIPANTS IN THE PRESENT STUDY

Ethnicity	Count (%)	Ethnicity	Count (%)
Biracial/Multi-racial	5 (13.9%)	Black American	4 (11.1%)
Caucasian	13 (36.1%)	East Asian	11 (30.6%)
Hispanic	5 (13.9%)	Native American	1 (2.8%)
Pacific Islander	0 (0.0%)	South Asian	3 (8.3%)
South East Asian	1 (2.8%)		

Table 5: Participants' self-report ethnicity in the present study

B MIND PERCEPTION SURVEY

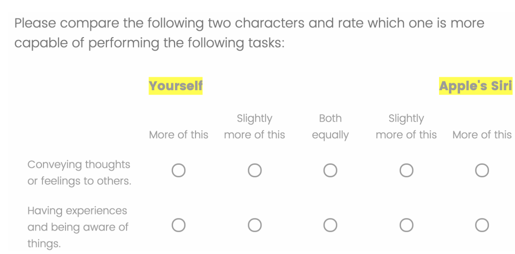
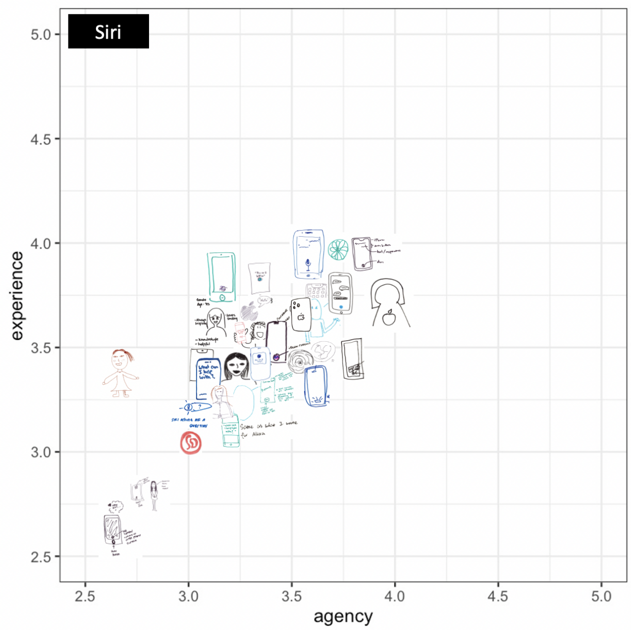


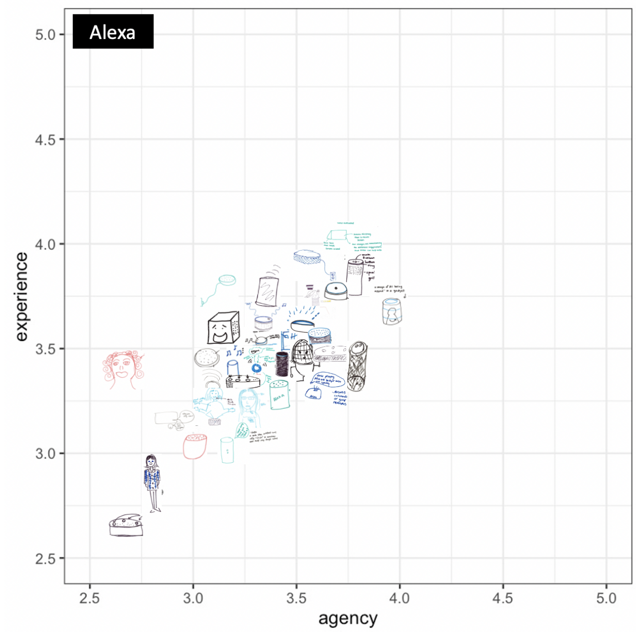
Figure 5: Snapshot of the pair-wise rating questions

(Continued on next page)

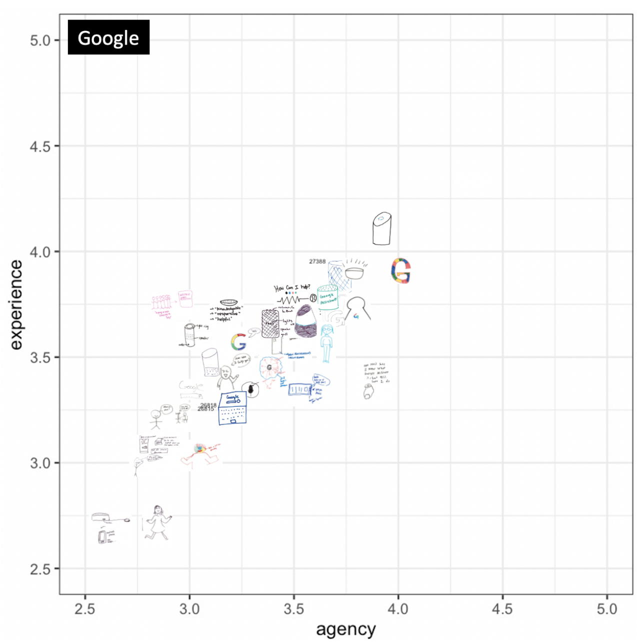
C DRAWINGS OF OTHER AI ENTITIES



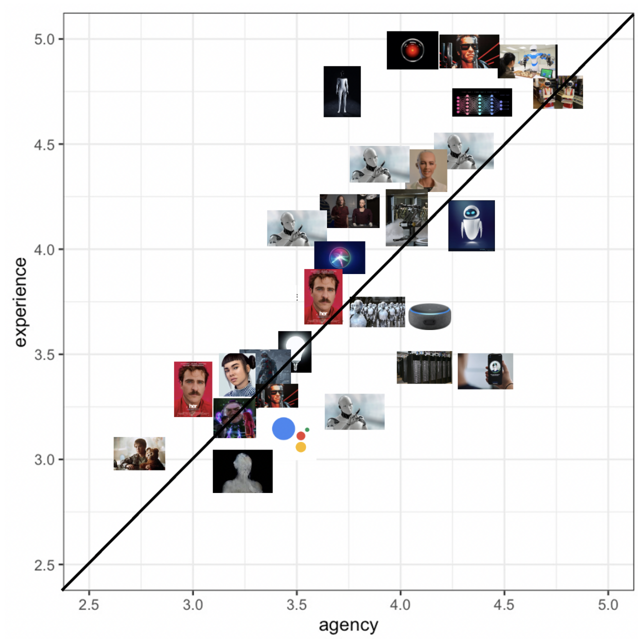
(a) Siri



(b) Alexa



(c) Google Assistant



(d) Pop Culture

Figure 6: Drawings of consumer AI products (Siri, Alexa, and Google Assistant) and AI in popular culture

D QUALITATIVE CODEBOOK

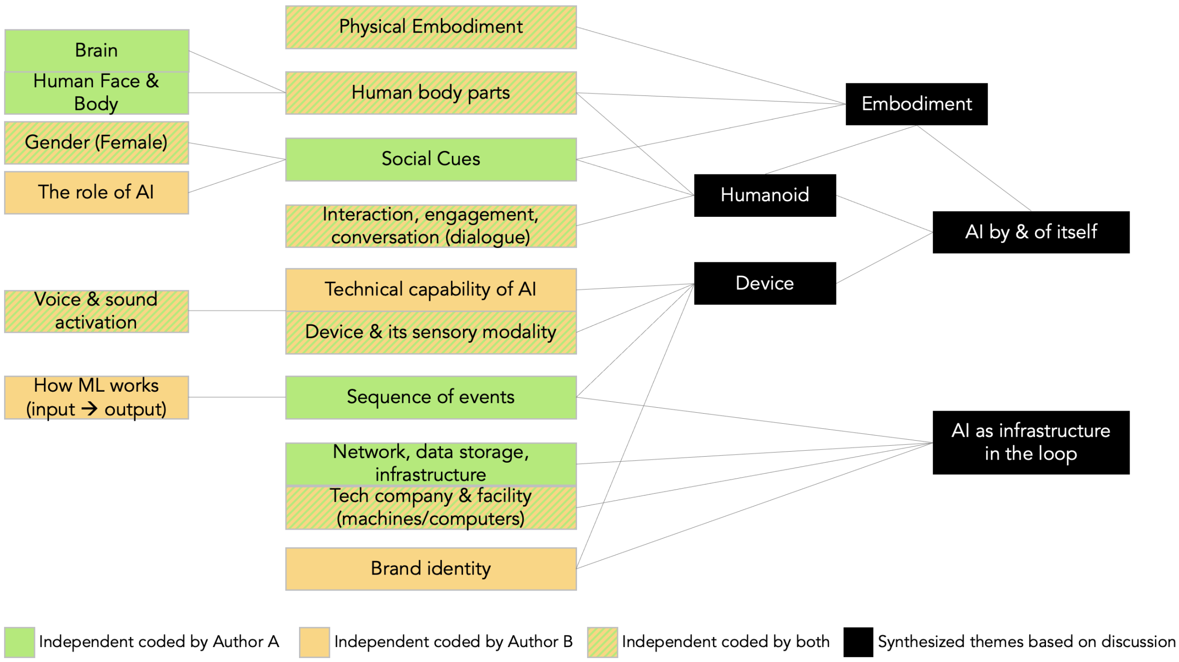


Figure 7: Emerging themes based on qualitative coding of the drawing data