How do employees imagine AI they want to work with: A drawing study

Samantha Straka*
Martin J. Koch*
Astrid Carolus
Marc Erich Latoschik
Carolin Wienrich
samantha.straka@uni-wuerzburg.de
martin.koch@uni-wuerzburg.de
astrid.carolus@uni-wuerzburg.de
marc.latoschik@uni-wuerzburg.de
carolin.wienrich@uni-wuerzburg.de
University of Würzburg
Würzburg, Bavaria, Germany

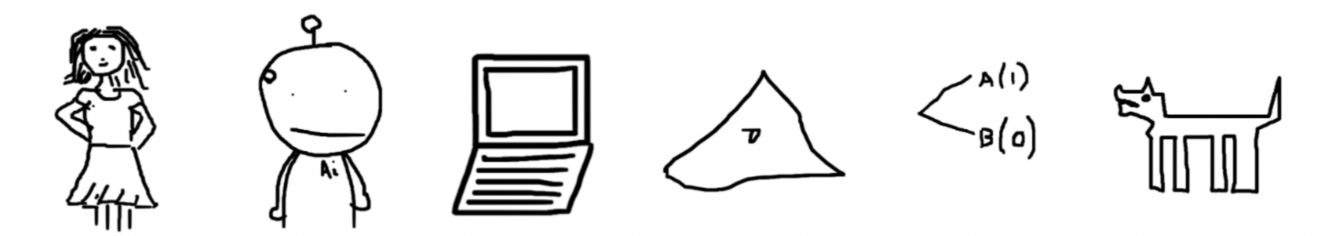

Figure 1: From left to right: representative samples from the six different categories found in AI drawings: *human-shaped, robot-shaped, hardware-shaped, simple shapes, computer-science related, animal-shaped.*

ABSTRACT

Perceptions about AI influence the attribution of characteristics and the interaction with AI. To find out how workers imagine an AI they would like to work with and what characteristics they attribute to it, we asked 174 working individuals to draw an AI they would like to work with, to report five adjectives they associate with their drawing and to evaluate the drawn and three other, typical AI representations (e.g. robot, smartphone) either presented as male or female. Participants mainly drew humanoid or robotic AIs. The adjectives that describe AI mainly referred to the inner characteristics, capabilities, shape, or relationship types. Regarding the evaluation, we identified four dimensions (warmth, competence, animacy, size) that can be reproduced for male and female AIs and different AI representations. This work addresses diverse conceptions of AI in the workplace and shows that human-centered AI development is necessary to address the huge design space.

*Both authors contributed equally to this research.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585631

CCS CONCEPTS

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

KEYWORDS

artificial intelligence, representation, warmth and competence, drawings, user-centered

ACM Reference Format:
Samantha Straka, Martin J. Koch, Astrid Carolus, Marc Erich Latoschik, and Carolin Wienrich. 2023. How do employees imagine AI they want to work with: A drawing study. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3585631

1 INTRODUCTION

Imagine working together with artificial intelligence (AI). What do you see before your inner eye? Do you have a nice little helper in mind that makes your job easier or a big machine that scares you? Initial general surveys show that perceptions provide at least a skewed picture of the principles, opportunities, and risks of AI, often based on science fiction narratives [21, 50, 60]. These narratives sketch AI as an uncaring threat to humanity that has lost control of its own creation [6]. In the context of work, the more realistic and often published threat of AI eliminating jobs [15] has additionally led to a rather negative perception of this technology

[39, 60]. With such a negative perception, it will be very challenging to use AI systems in the work context in an enjoyable, motivating, and efficient way. However, little is known about how people envision an AI system they would like to work with. Moreover, there is little empirical research on which perceptual dimensions may be essential for designing AI-System in the context of work. Therefore, we explore how 174 people, working in different domains, imagine an AI that they *want* to work with and what associations and perceptions they associate with it. We further aim to identify latent dimensions that can be used for the future evaluation and design of AI representations. Thus, our work contributes initial insights into how we can design AI systems that are enjoyable, motivating, and efficient to work with.

2 RELATED WORK

The appearance of an AI at the interface can range from simple effects like the execution of a requested operation to simple text displays to humanoid and human-looking robots or virtual agents [50]. However, little research has been conducted on what AI can and should look like for specific purposes. Phillips and colleagues [33] conducted three studies in which participants were asked to draw a household robot, a military robot, a generic robot, a humanoid robot, or an AI. Participants drew robots differently depending on the context with more human-like depictions in the household context, while military robots and AI showed fewer facial features. The drawings of AI did yield a diverse range of depictions, going from completely anthropomorphic robots to abstract drawings of a network. In sum, this reveals that the context of usage influences the expected appearance of the depicted entities. Carolus and Wienrich [5] asked participants how they imagine the bodies of intelligent voice assistants (i.e. smart speakers). They found that most participants imagined human beings or (humanoid) robots, with few imagining other representations (e.g. animals, objects, or abstract representations). These few studies suggest that human characteristics are often attributed to AI systems. According to the Media Equation Approach [30, 36], this has severe consequences because numerous studies with computers [29], smartphones [4], or smart speakers [51] showed that people also assign human-like characteristics to technical entities. These attributions consequently influence the perception and interaction. It is, therefore, interesting to look at which properties influence the perception of (humanoid) others.

In human-human interaction, warmth and competence (or similar constructs) determine how we perceive others [1, 32, 38, 52]. Warmth includes evaluations that describe the perceived intent another human might have (e.g. morality, trustworthiness, and friendliness). Competence comprises traits that describe the perceived ability to enact the intent (e.g. efficacy, skill, and intelligence) [14]. Stereotypes (e.g. religious affiliation, origin, age, and profession) and behaviour influence how warm and competent an individual is perceived which leads to specific emotional and behavioral reactions [8]. Additionally, attractiveness [57, 59], facial expression [23, 28, 44, 49, 58], and choice of clothing and style [19, 43, 53] tend to influence how we perceived others.

Similarly, the appearance of artificial entities like robots influences the experienced closeness [12], perceived aggression, intelligence, and animacy [40]. Robots with a human-like appearance foster perceived attractiveness [45, 55], trust [46, 47], preference [11, 17], likability/sympathy [13, 42], empathy [37], sociability [24], warmth, and competence [41]. For AI, perceived warmth predicts preference [16, 26] and anthropomorphism seems to predict credibility [55], performance expectations and preference (in highly controlled contexts) and threat (in less controlled contexts) [54], effort expectancy [18], likeability [48], trust [31], forgiveness of failure [25] and more risky financial decisions [10]. Overall, it can be said that the appearance of an AI has a strong influence on the emotional and behavioral reactions of users and also impact their interaction with artificial counterparts.

In summary, research shows that people have very different perceptions of AI systems, and letting them draw freely is a promising way to interrogate these perceptions (for further discussion see [5]). However, there are few studies so far that have used this method and none that specifically address the context of work. It also turns out that many internal and external features influence the perception of human and artificial counterparts. However, there is no systematic classification of the features. Therefore, this paper addresses two research questions: How do people working in different domains imagine an AI system they would like to collaborate with (Part I)? Which latent features influence the perception of AI in the work context (Part II)?

3 EMPIRICAL STUDY

3.1 Data collection & sample

To answer the research questions, empirical data was collected from 174 English-speaking individuals in an online study using the survey tool SoSci-Survey on August 31, 2022. Participants were recruited using Prolific.org and received a compensation of £3.50 for completion. The average completion time was 20.56 minutes (*SD* = 5.52). Individuals were eligible to participate if they are either full- or part-time employees. The average age was 39.82 years (*SD* = 11.65, ranging from 19 to 73 years). Of all participants, 47.25% identified as female, 51.10% as male, and 1.65% as "other". All individuals who participated came from the United Kingdom.

3.2 Experimental procedure and measures

After the participants were informed about the purpose of the study and gave consent, they were asked to imagine they were working with AI and to draw an AI they would like to work with using their computer and mouse. This procedure was chosen as it was deemed less cumbersome for the participants and we expected clearer images than drawings on paper. They should then freely describe the AI using five adjectives. Then, participants were randomly assigned to one of two experimental conditions (female (A)/male (B) representations of AI). Each participant saw three different AI representations in random order (Figure 2). Half of them either saw a stereotypical female or male set of AI representations depending on the experimental condition. Participants rated the shown AI representations regarding 35 semantic differentials focused on the internal (e.g. dead/alive, evil/good, cruel/caring, incompetent/competent) and external properties (e.g. small/huge, light/heavy, young/old, weak/strong). For each semantic differential, the participants rated how much they attribute these properties to the AI representation on a 7-point scale. The items stemmed from common questionnaires [2, 3] as well as self-constructed items (for a full list, see Appendix A).

3.3 Part I: Exploratory analysis of the drawn AI

We inductively analyzed the drawings of AI through a grounded theory approach with three researchers [7] to identify if there are general categories of representations of AI. All researchers had experience with this approach. The analysis was done in Miro [27], where each drawing was uploaded without any additional information. First, we went through each drawing together, discussed what was pictured, and grouped the drawings on their content. Then, for axial coding, we went through each of these initial groups to find common themes or elements to break them into even smaller, more detailed, and cohesive categories. This was repeated until no smaller categories could be found. Then, each category was labeled with a name. If there were disagreements between researchers on a particular drawing, this drawing was put aside and discussed again later to find a suitable category. If no suitable or agreeable category was found, it was put in the 'Other' - category. The five attributes associated with every drawing were then analyzed in the same manner within each category to identify whether there are emphases in the attribution of certain adjectives depending on the category. If an entry was not showing an accurate word or if it was illegible due to spelling mistakes, it was removed from the analysis.

3.3.1 Results of the drawings. Within the **174 drawings of AI**, four big categories showed AI as robots (50 drawings), humans (65 drawings), simple shapes (20 drawings), and hardware (18 drawings). Representative samples from each category can be seen in figure 1. Robot and human-shaped AI representations differed in the use of right angles to draw the outlines and/or the addition of mechanical parts such as antennas or buttons. AI depicted as hardware was devoid of appendages to physically interact with the real world and was often recognisable as desktop computers, laptops, and smart speakers. No drawing showed a smartphone. Eight drawings showed AI as either animal and 11 were computer-science-related pictures. Animals were either quadruped or had obvious animal-like features such as ears, a tail, or whiskers. Computer-science-related drawings contained any drawings of networks, logical gates, or zeroes and ones. Simple geometric shapes contained circles, rectangles, boxes, triangles, and swirls. One small *'Other'* - group contained the three drawings that did not fit in other groups showing a plant, UI-Buttons and a floating cube between two flat elements.

The category that contained drawings representing AI as a human contained several sub-categories with varying degrees of the depicted human body: one category included only simple *smileys* (n=12), one showed *human faces in a box* (n=4), one included *detailed human faces* (with e.g hair, ears, glasses) (n=21), one included *whole human bodies* (n=19) and one contained only *human upper bodies* (n=9). Most drawings did not explicitly state the gender of the AI. All but eight drawings depicted the human-shaped AI drawings to be smiling, the rest showed a neutral facial expression.

The group of robotic AI representations were further divided into *robotic faces* and *whole body robots*. All robot bodies had appendages to interact with the world (either arms or legs). One robot had six arms, the rest had two or none. Most robots had a neutral (neither smiling nor frowning) or not discernible expression.

3.3.2 Results of the associated attributes. From the **attributes that were associated with each drawing**, five groups were deducted. These groups were the (1) *appearance*, the (2) *inner characteristics*, the (3) *capabilities* of an AI, (4) references to other entities (e.g human-like or robotic), or (5) was not fitting in another group, being labeled as *Other*. Human-, robot- and animal-shaped AI were predominantly described with inner characteristics while the other categories focused on the capabilities of the AI. These categories were then again subdivided into smaller categories. They can be seen in table 1, with the total number of attributes given to each indicated in brackets. For all categories, "Warmth" contains attributes such as (*friendly, happy, kind, pleasant, approachable, likable, trustworthy, nice and warm*), while "Cold" contains the attributes *cold* or *emotionless*. "Capable" contained attributes concerning functionalities and power such as e.g *powerful, functional, working, dependable, productive*, while "Smart" contains the attributes *intelligent, smart, and knowledgeable*. The distinction between "capable" and "smart" was made as "capable" refers to more in-depth functionalities while "smart" is a more general description. These two categories differ for the found categories of drawings, with human-, animal- and robot-shaped AI being perceived as smart and the others as powerful, capable, useful and functional. "Usable" contains attributes concerning the usage and usability of an AI that a user directly experiences while interacting with it, e.g *effective, efficient, usable*. An AI should be effective, efficient, and intuitive, regardless of shape.

3.4 Part II: Exploratory factor analysis and measurement invariance of the 35 semantic differentials

To find latent variables that influence perceptions of AI, the 35 semantic differentials collected for every AI (drawn and given representations) were evaluated with exploratory factor analysis. We also used confirmatory factor analysis to test for measurement invariance with robust estimations (Satorra-Bentler correction) for female and male AI representations and the different classes of AI (robot, human, phone).

3.4.1 Results of the exploratory factor analysis. Even though the χ^2-test became significant ($\chi^2(62) = 108.49$, $p < .001$), other model fit indices show a good model fit (TLI = .98, RMSEA = .04, 90%-CI [.028, .050]) after removing items with low loadings. The loadings for each item and factors can be seen in Table 2. In summary, the exploratory factor analysis helped to reveal four factors that fit the empirical data to a high extent and help to explain a total of 68.10% of variance. The factors encompass the domains "warmth", "competence", "animacy", and "size".

3.4.2 Results of the measurement invariance tests. Measurement invariance describes the "equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups" [35].For AI gender, differential

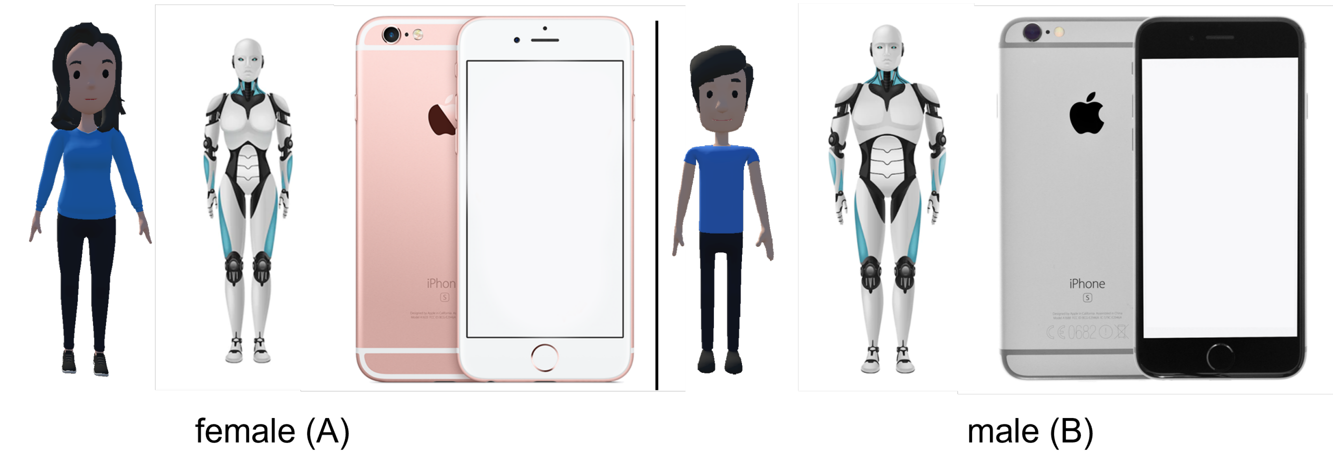

Figure 2: Different AI representations for both experimental conditions (from left to right: humanoid, anthropomorphic robot, phone)

Table 1: The attributes given to each category of drawings. The numbers in brackets indicate the total number of attributes given.

Category	Inner Characteristics	Capabilities	Appearance	References	Other
Human (354)	**190 Attributes:** Warmth (134), Friend (5), Cold (4), Other (47)	**64 Attributes:** Smart (22), Simple (7) , Fast(7), Reliable/Accurate (6), Other (22)	**55 Attributes:** Plain (11), Female (9), Attractive (6), Non-Binary (3), Male (1), Other (25)	**19 Attributes:** Human-Like (12), Computer/Machine (5), Other (2)	**47 Attributes**
Robot (290)	**114 Attributes:** Warmth (70), Fun (9), Cold (7), Other (28)	**69 Attributes:** Smart (29), Capable (15), Usable (14), Logical (5), Fast (4), Other (16)	**48 Attributes:** Metallic (5), Cute (5), Large (4), Strong (4), Mobile (4), Male (2), Other (24)	**44 Attributes:** Robot (17), Human-Like (8), Technical (5), Non-Human (3), Other (11)	**15 Attributes**
Hardware (109)	**32 Attributes:** Warmth (12), Boring (7), Cool (4), Cold (3), Scary (3), Other (3)	**39 Attributes:** Capable (10), Smart (7), Usable (6), Fast (4), All-Seeing (4), Other (8)	**18 Attributes:** Metallic (6), Compact (4), Clean/ Tidy (4), Large (2), Other (2)	**15 Attributes:** Technical (6), Robot (5), Computer (2), Other (2)	**5 Attributes**
Simple Shapes (109)	**21 Attributes:** Warmth (17), Animated (3), Terrible (1)	**36 Attributes:** Capable (11), Usable (10), Smart (5), Fast (2), Other (8)	**16 Attributes:** Rectangular (5), Shiny (5), Stylish (3), Other (3)	**19 Attributes:** Technical (6), Alien/Inhuman (6), Futuristic (3), Alexa/Amazon (2)	**17 Attributes**
C.S -Related (67)	**10 Attributes:** Warmth (2), Cold (2), Witty (2), Straight-Forward (1), Obedient (1), Mysterious (1), Creative (1)	**37 Attributes:** Capable (13), Smart (6), Usable (6), Learning (2), Other (10)	**3 Attributes:** Network (1), Narrow (1), Pattern (1)	**9 Attributes:** Computer (6), Human (1), Network (1)	**8 Attributes**
Animals (40)	**28 Attributes:** Warmth (26), Loyal (2)	**4 Attributes:** Smart (2), Intuitive (1), Connected (1)	**10 Attributes:** Cute (3), Female (2), Attractive (2), Small (1), Furry (1), Young (1)	**3 Attributes:** Cat-Like (1), Lifelike (1), Human-Like (1)	**0 Attributes**

item loading was found for "No own will/Own will", "Incompetent/Competent", and "Unreliable/Reliable" while a differential intercept was found for "Incompetent/Competent". Differential item functioning was found for "Cruel/Caring" regarding item loadings and for "Cruel/Caring", "Unreliable/Reliable", "No own will/Own will", "Awful/Nice", "Has no consciousness/Has consciousness", "Unintelligent/Intelligent", "Not useful/Useful", and "Small/Huge" regarding intercepts between AI representation classes. However, partial invariant models with acceptable model fit were found for both sub-sample comparisons (model fit and comparisons can be found in Appendix B). These results suggest that different representations of AI seem not to be evaluated in the same way.

4 DISCUSSION

Overall, the approach we chose for this study was rather exploratory and user-centred. The participants were asked to draw and evaluate AI they would like to work with without receiving any further cues or information on how an AI might look. Nonetheless, the results are very much in line with the literature on drawings of AI systems [5, 33]. While AI that people want to work with came in diverse shapes, our participants imagine AI mainly as humans or at least humanoid robots/humanoid robot faces. Drawings that are fundamentally different from human and humanoid robotic revealed that AI representations go beyond what can be considered standard, e.g animals, hardware, or abstract representations. Notably, no drawing depicted AI as a smartphone. This is surprising, as it is an everyday tool and uses AI for many functions. However, this is in line with other research that found robots, computers, and humans are the

Table 2: Obliquely rotated factor loadings of the items for AI representation evaluation (factor loadings < .20 are not shown)

	Factor 1	Factor 2	Factor 3	Factor 4
Dislikable/Likable	0.68			
Awful/Nice	0.88			
Evil/Good	0.89			
Dubious/Trustworthy	0.71			
Bad to me/Good to me	0.81			
Cruel/Caring	0.94		-0.23	
Dead/Alive		0.69		
No own will/Own will	-0.30	0.89		
Has no consciousness/Has consciousness		0.87		
Emotionally Unintelligent/Emotionally Intelligent	0.25	0.66		
Unintelligent/Intelligent			0.68	0.21
Not useful/Useful			0.76	
Incompetent/Competent			0.83	
Unreliable/Reliable			0.68	
Small/Huge				1.00
Light/Heavy				0.83

first things that come to mind for many people when thinking of AI [39].

Participants associated different attributes with the categories of drawings. In general, participants mostly attributed aspects referring to inner characteristics, capabilities, the appearance of the AI, or references to other beings (e.g being human-like). Many adjectives directly referred to warmth or other social characteristics such as helpfulness, funniness, or approachability. The less human-like a category is drawn, the fewer attributes refer to warmth but to the capabilities of an AI, e.g computer-science-related representations showed nearly no association with inner characteristics but with capabilities. In contrast, animals were more associated with inner characteristics than with capabilities. This difference in associations might also appear in different working contexts.

Other descriptions described the drawn human-shaped AI to be a friend. It seems that a social relationship is important to many drawers of human AI representations. Especially in the light of the Media Equation Approach and "Computers as Social Actors" [30, 36], it seems reasonable to assume that AI is also perceived as a social actor. Even though some attributes called the AI "scary" or "terrible", most were more positive than negative. This is an expected result, as the drawings and the associated attributes show an AI that our participants *want* to work with.

In the exploratory factor analysis, latent dimensions of warmth, competence, animacy, and size were identified for the total sample. All four factors could be reproduced for the sub-samples of gender (male vs. female) and AI class (humanoid vs. robot vs. phone). However, differential item functioning was found for several items regarding item loading and item intercepts indicating different importance of the items for the latent dimensions and different base levels regarding the items for different AI gender and AI representations. It is known from research on human-human interaction that warmth and competence are two central dimensions on which people perceive their communicative partners [8] and that they are central for emotional and behavioral reactions that humans show towards human interactive partners [9]. Research related to interactions with robots and AI also shows that warmth and competence seem to be central dimensions that influence our interaction [16, 26, 34, 41]. Animacy has also been considered an important dimension in research on robots [40]. Similarly, anthropomorphism has been discussed in relation to robots [42], virtual agents [20] and AI [5, 10, 18, 22, 25, 31, 48, 54, 55]. The question arises whether the "human-likeness" (i.e. anthropomorphism) has a positive effect because of the more familiar appearance [56] or because we perceive robots, virtual agents and/or AI which are more "human-like" to also be more animate. Possibly, higher anthropomorphism goes hand in hand with higher animacy which could be a potential mediator for the effects of anthropomorphism. So far, both constructs are rather loosely defined and can be difficult to distinguish.

Even though it was possible to reproduce the same configuration of item loadings and latent dimensions in sub-samples of different AI representation gender and different classes of AI representations (human, robot, phone), we found metric and scalar measurement noninvariance. Some items showed differential item functioning regarding the height of item loadings and item intercepts. At first, this poses a practical problem as it disallows the comparisons of different AI representation gender and classes regarding the latent dimensions. Besides its relevance for the research practice of mean comparisons, the presence of measurement noninvariance is also interesting for our knowledge about the assessment of AI representations. We must assume that some characteristics (i.e. items) differ in their importance for the evaluation of AI representations and that different evaluative standards regarding some characteristics (i.e. items) are applied to different classes of AI representations. It might be concluded that AI representations that differ regarding their perceived gender and their general class (human, robot, phone), differ more substantially not only regarding the perceived warmth, competence, and animacy but also regarding the basic evaluation processes. AI representations are, thus, more diversely perceived as expected. It may be necessary to develop different measurement tools to enable the assessment of different AI representations.

4.1 Limitations & Future Work

Several limitations need to be mentioned regarding the present study. First, the quantitative results are based on a rather small sample (N = 174). For the factor analyses, it is difficult to say whether the factorial structure will hold up when further data is collected, in particular in the subsamples of AI classes and gender. Replication with a bigger and more diverse sample might be essential, especially as our complete sample is UK-based. Perceptions of AI seem to be strongly dominated by pop cultural influences and might differ between countries, Western and Eastern cultures, social groups that receive different media, different working domains, or by age or gender. However, for the qualitative analysis of drawings and adjectives, the sample is large and exceeds the size of previous studies.

The evaluation of the AI representations was done via semantic differentials (i.e. two adjectives or short descriptions that are presumed to be exact opposites). This might be problematic: Semantic differentials cannot necessarily be assumed to be metric (no zero point and no identical distances between selection options). Ambivalent adjective pairs such as "Awful/Nice" or "Bad to me/Good to me" might not present "real" opposites and their dichotomy cannot be assumed with certainty. For example, an AI might be "Bad to me" (e.g. because it threatens my feeling of job security) and "Good to me" (e.g, because it helps me with my tasks and reduces job load) at the same time. A further evaluation with independent adjectives/short descriptions instead of semantic differentials might create interesting insights. Especially in the case of ambivalent AI representations, a more accurate picture could emerge. Including the collected adjectives that the participants have generated might identify dimensions given to AI by their potential users.

Most participants drew humans or humanoid robots. This, however, does not mean those humanoid and robotic-humanoid representations are best suited to represent AI in a working environment. More research is needed to explore which options can be implemented at the workplace and to test how different AI representations impact motivation, user intention, user experience, and other work-related variables such as effectiveness and satisfaction. For example, as we found that attributed capabilities and characteristics differ regarding the category of the AI drawing, certain AI representations might be best suited for distinct working contexts that come with specific expectations regarding the role of an AI. Extended reality [50] can be a useful tool to prototype AI directly in a specific working context. Drawings with a computer mouse might have led to simplified or stylized depictions. A further evaluation with pen and paper and in color might stimulate the creativity of the participants and lead to additional insights. Our study does not provide any insights into which of the identified dimensions is central to the actual intention to use an AI or user experience. In the next step, it would be crucial to identify how the AI representation classes differ regarding the identified dimension, how characteristics of an AI representation are related to warmth, competence, animacy, and size, and how these dimensions predict user experience and the intention to use an AI system. Also, it might be interesting to identify how characteristics of the participants such as prior experiences with AI and field of work influence AI drawings and evaluation.

5 CONCLUSION

This study helps to understand how working people imagine AI *they want to work with*, which conceptions they associate with it, and to identify latent dimensions to evaluate different AI representations. We found that (1) most participants either imagined an AI they would like to work with as a human or robot with few notable deviations (animals, hardware, and abstract representations), (2) most adjectives referred to the AI representations' inner characteristics, capabilities, shape, or references to other entities, (3) human-like conceptions were more associated with inner characteristics and device-like conceptions with capabilities, and (4) latent dimensions of warmth, competence, animacy, and size are central for the evaluation of AI representations but produce a measure invariance. Thus, our study contributes to the development of human-centered AI in the context of work by considering the perceptions and associations of potential users. Our work lays the first stepping stone to explore AI representations further and match them with work outcomes and job satisfaction in future studies.

ACKNOWLEDGMENTS

This research was funded by the German Federal Ministry of Labour and Social Affairs [DKI.00.00030.21].

REFERENCES

[1] Solomon. E. Asch. 1946. Forming impressions of personality. *Journal of abnormal psychology* 41, 3 (1946), 258–290. https://doi.org/10.1037/h0055756

[2] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2008. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International journal of social robotics* 1, 1 (2008), 71–81. https://doi.org/10.1007/s12369-008-0001-3

[3] Elizabeth Bouquet, Ville Mäkelä, and Albrecht Schmidt. 2021. Exploring the Design of Companions in Video Games. In *Academic Mindtrek '21: Proceedings of the 24th International Academic Mindtrek Conference*. ACM Press, Tampere, Finland, 145–153. https://doi.org/10.1145/3464327.3464371

[4] Astrid Carolus, Jens F. Binder, Ricardo Muench, Catharina Schmidt, Florian Schneider, and Sarah L Buglass. 2019. Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society* 21, 4 (2019), 914–938.

[5] Astrid Carolus and Carolin Wienrich. 2022. "Imagine this smart speaker to have a body": An analysis of the external appearances and the characteristics that people associate with voice assistants. *Frontiers in Computer Science* 4 (2022), 1–20. https://doi.org/10.3389/fcomp.2022.981435

[6] Stephen Cave, Claire Craig, Kanta Dihal, Sarah Dillon, Jessicae Montgomery, Beth Singler, and Lindsay Taylor. 2018. Portrayals and perceptions of AI and why they matter. https://doi.org/10.17863/CAM.34502

[7] Juliet M. Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.

[8] Amy J. C. Cuddy, Susan T. Fiske, and Peter Glick. 2008. Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. *Advances in Experimental Social Psychology* 40 (2008), 61–149. https://doi.org/10.1016/s0065-2601(07)00002-0

[9] Amy J. C. Cuddy, Peter Glick, and Anna Beninger. 2011. The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior* 31 (2011), 73–98. https://doi.org/10.1016/j.riob.2011.10.004

[10] Yuanyuan Cui. 2022. Sophia Sophia tell me more, which is the most risk-free plan of all? AI anthropomorphism and risk aversion in financial decision-making. *International Journal of Bank Marketing* 40, 6 (2022), 1133–1158. https://doi.org/10.1108/IJBM-09-2021-0451

[11] Maartje M. A. de DeGraaf and Somaya Ben Allouch. 2015. The evaluation of different roles for domestic social robots. In *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, Kobe, 676–681. https://doi.org/10.1109/ROMAN.2015.7333594

[12] Friederike Eyssel, Dieta Kuchenbrandt, Simon Bobinger, Laura de Ruiter, and Frank Hegel. 2012. 'If you sound like me, you must be more human': on the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *HRI '12: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, Holly Yanco, Aaron Steinfeld, Vanessa Evers, and Odest Chadwicke Jenkins (Eds.). ACM Press, New York, New York, USA, 125–126. https://doi.org/10.1145/2157689.2157717

[13] Ylva Ferstl, Sean Thomas, Cédric Guiard, Cathy Ennis, and Rachel McDonnell. 2021. Human or Robot?. In *Proceedings of the 21th ACM International Conference on Intelligent Virtual Agents*. Association for Computing Machiner, New York, NY, USA, 76–83. https://doi.org/10.1145/3472306.3478338

[14] Susan T. Fiske, Amy J. C. Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: warmth and competence. *Trends in cognitive sciences* 11, 2 (2007), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

[15] Carl Benedikt Frey and Michael A. Osborne. 2017. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change* 114 (2017), 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

[16] Zohar Gilad, Ofra Amir, and Liat Levontin. 2021. The Effects of Warmth and Competence Perceptions on Users' Choice of an AI System. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yoshifumi Kitamura, Aaron Quigley, Katherine Isbister, Takeo Igarashi, Pernille Bjørn, and Steven Drucker (Eds.). ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3411764.3446863

[17] Jennifer Goetz, Sara Kiesler, and Aaron Powers. 2003. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003*. IEEE, Millbrae, California, 55–60. https://doi.org/10.1109/ROMAN.2003.1251796

[18] Dogan Gursoy, Oscar Hengxuan Chi, Lu Lu, and Robin Nunkoo. 2019. Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management* 49 (2019), 157–169. https://doi.org/10.1016/j.ijinfomgt.2019.03.008

[19] Susanne-Marie Hüttner and Michael Linden. 2017. Modification of First Impression Formation and "Personality" by Manipulating Outer Appearance. *Psychopathology* 50, 2 (2017), 141–145. https://doi.org/10.1159/000453585

[20] Prem Kalra, Nadia Magnenat-Thalmann, Laurent Moccozet, Gael Sannier, Amaury Aubel, and Daniel Thalmann. 1998. Real-time animation of realistic virtual humans. *IEEE Computer Graphics and Applications* 18, 5 (1998), 42–56. https://doi.org/10.1109/38.708560

[21] Patrick Gage Kelley, Yongwei Yang, Courtney Heldreth, Christopher Moessner, Aaron Sedley, Andreas Kramm, David Newman, and Allison Woodruff. 2019. Happy and Assured that life will be easy 10 years from now.": Perceptions of Artificial Intelligence in 8 Countries. (2019).

[22] Ahyeon Kim, Minha Cho, Jungyong Ahn, and Yongjun Sung. 2019. Effects of Gender and Relationship Type on the Response to Artificial Intelligence. *Cyberpsychology, behavior and social networking* 22, 4 (2019), 249–253. https://doi.org/10.1089/cyber.2018.0581

[23] Sointu Leikas, Markku Verkasalo, and Jan-Erik Lönnqvist. 2013. Posing personality: Is it possible to enact the Big Five traits in photographs? *Journal of Research in Personality* 47, 1 (2013), 15–21. https://doi.org/10.1016/j.jrp.2012.10.012

[24] Dingjun Li, P. L. Patrick Rau, and Ye Li. 2010. A Cross-cultural Study: Effect of Robot Appearance and Task. *International Journal of Social Robotics* 2, 2 (2010), 175–186. https://doi.org/10.1007/s12369-010-0056-9

[25] Xingyang Lv, Yue Liu, Jingjing Luo, Yuqing Liu, and Chunxiao Li. 2021. Does a cute artificial intelligence assistant soften the blow? The impact of cuteness on customer tolerance of assistant service failure. *Annals of Tourism Research* 87 (2021), 103114. https://doi.org/10.1016/j.annals.2020.103114

[26] Kevin R. McKee, Xuechunzi Bai, and Susan T. Fiske. 2022. Warmth and competence in human-agent cooperation. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022)*. arXiv, Auckland, 1–30. https://doi.org/10.48550/arXiv.2201.13448

[27] Miro. 2023. *Miro*. Miro. https://miro.com/

[28] Joann M. Montepare and Heidi Dobish. 2003. The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior* 27, 4 (2003), 237–254. https://doi.org/10.1023/A:1027332800296

[29] Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of social issues* 56, 1 (2000), 81–103.

[30] Clifford Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. In *Conference companion on Human factors in computing systems - CHI '94*, Catherine Plaisant (Ed.). ACM Press, New York, New York, USA, 72–78. https://doi.org/10.1145/259963.260288

[31] Dongfang Niu, Jacques Terken, and Berry Eggen. 2018. Anthropomorphizing information to enhance trust in autonomous vehicles. *Human Factors and Ergonomics in Manufacturing & Service Industries* 28, 6 (2018), 352–359. https://doi.org/10.1002/hfm.20745

[32] Guido Peeters. 1983. Relational and informational patterns in social cognition. In *Current issues in European Social Psychology*, Willem Doise and Serge Moscovici (Eds.). Cambridge University Press, Cambridge, 201–237.

[33] Elizabeth Phillips, Daniel Ullman, Maartje M.A. deGraaf, and Bertram F. Malle. 2017. What does a robot look like?: A multi-site examination of user expectations about robot appearance. In *Proceedings of the Human Factors and Ergonomics Society*, Vol. 49. HFES, Boston, Massachusetts, 1215––1219. https://doi.org/10.1177/1541931213601786

[34] Pooja Prajodand, Mohammed Al Owayyed, Tim Rietveld, Jaap Jan Van Der Steeg, and Joost Broekens. 2019. The Effect of Virtual Agent Warmth on Human-Agent Negotiation. In *AAMAS '19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, Noa Agmon, Matthew E. Taylor, Edith Elkind, and Manuela Veloso (Eds.). Curran Associates Inc, Red Hook, NY, 71–76.

[35] Diane L. Putnick and Marc H. Bornstein. 2016. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental review : DR* 41 (2016), 71–90. https://doi.org/10.1016/j.dr.2016.06.004

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

[37] Laurel D. Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. 2009. How anthropomorphism affects empathy toward robots. In *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Matthias Scheutz, François Michaud, Pamela Hinds, and Brian Scassellati (Eds.). ACM Press, New York, New York, USA, 245–246. https://doi.org/10.1145/1514095.1514158

[38] Seymour Rosenberg, Carnot Nelson, and P. S. Vivekananthan. 1968. A multidimensional approach to the structure of personality impressions. *Journal of personality and social psychology* 9, 4 (1968), 283–294. https://doi.org/10.1037/h0026086

[39] Neil Selwyn, Beatriz Gallo Cordoba, Mark Andrejevic, and Liz Campbell. 2020. AI for Social Good? Australian Attitudes Toward AI and Society. https://www.monash.edu/data-futures-institute/news/ai-for-social-good-australian-public-attitudes-toward-ai-and-society

[40] Valerie K. Sims, Matthew G. Chin, David J. Sushil, Daniel J. Barber, Tatiana Ballion, Bryan R. Clark, Keith A. Garfield, Michael J. Dolezal, Randall Shumaker, and Neal Finkelstein. 2005. Anthropomorphism of Robotic Forms: A Response to Affordances?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 49. HFES, Orlando, Florida, 602–605. https://doi.org/10.1177/154193120504900383

[41] Steven J. Stroessner and Jonathan Benitez. 2019. The Social Perception of Humanoid and Non-Humanoid Robots: Effects of Gendered and Machinelike Features. *International Journal of Social Robotics* 11, 2 (2019), 305–315. https://doi.org/10.1007/s12369-018-0502-7

[42] S. Shyam Sundar, T. Franklin Waddell, and Eun Hwa Jung. 2016. The Hollywood Robot Syndrome media effects on older adults' attitudes toward robots and adoption intentions. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE Press, Christchurch, 343–350. https://doi.org/10.1109/HRI.2016.7451771

[43] G. R. Thornton. 1944. The effect of wearing glasses upon judgments of personality traits of persons seen briefly. *Journal of Applied Psychology* 28, 3 (1944), 203–207. https://doi.org/10.1037/h0055862

[44] Alexander Todorov and Jenny M. Porter. 2014. Misleading first impressions: different for different facial images of the same person. *Psychological science* 25, 7 (2014), 1404–1417. https://doi.org/10.1177/0956797614532474

[45] Fang-Wu Tung. 2016. Child Perception of Humanoid Robot Appearance and Behavior. *International Journal of Human-Computer Interaction* 32, 6 (2016), 493–502. https://doi.org/10.1080/10447318.2016.1172808

[46] Iis P. Tussyadiah, Florian J. Zach, and Jianxi Wang. 2020. Do travelers trust intelligent service robots? *Annals of Tourism Research* 81 (2020), 102886. https://doi.org/10.1016/j.annals.2020.102886

[47] 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 (2019), 507–518. https://doi.org/10.1108/JSM-01-2018-0045

[48] Katja Wagner, Frederic Nimmermann, and Hanna Schramm-Klein. 2019. Is It Human? The Role of Anthropomorphism as a Driver for the Successful Acceptance of Digital Voice Assistants. In *Proceedings of the 52nd Hawaii International Conference on System Sciences (Proceedings of the Annual Hawaii International Conference on System Sciences)*, Tung Bui (Ed.). Hawaii International Conference on System Sciences, Wailea-Makena, Hawaii, 1286–1395. https://doi.org/10.24251/hicss.2019.169

[49] Mirella Walker and Thomas Vetter. 2016. Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of personality and social psychology* 110, 4 (2016), 609–624. https://doi.org/10.1037/pspp0000064

[50] Carolin Wienrich and Marc Erich Latoschik. 2021. eXtended Artificial Intelligence: New Prospects of Human-AI Interaction Research. *Frontiers in Virtual Reality* 2 (2021), 94. https://doi.org/10.3389/frvir.2021.686783

[51] Carolin Wienrich, Clemens Reitelbach, and Astrid Carolus. 2021. The Trustworthiness of Voice Assistants in the Context of Healthcare Investigating the Effect of Perceived Expertise on the Trustworthiness of Voice Assistants, Providers, Data Receivers, and Automatic Speech Recognition. *Frontiers in Computer Science* 3 (2021), 53.

[52] Bogdan Wojciszke, Roza Bazinska, and Marcin Jaworski. 1998. On the Dominance of Moral Categories in Impression Formation. *Personality and Social Psychology Bulletin* 24, 12 (1998), 1251–1263. https://doi.org/10.1177/01461672982412001

[53] Jane E. Workman and Kim K. P. Johnson. 1991. The Role of Cosmetics in Impression Formation. *Clothing and Textiles Research Journal* 10, 1 (1991), 63–67. https://doi.org/10.1177/0887302X9101000109

[54] Yang Yang, Yue Liu, Xingyang Lv, Jin Ai, and Yifan Li. 2022. Anthropomorphism and customers' willingness to use artificial intelligence service agents. *Journal of Hospitality Marketing & Management* 31, 1 (2022), 1–23. https://doi.org/10.1080/19368623.2021.1926037

[55] Kyung Hyan Yoo and Ulrike Gretzel. 2009. The influence of virtual representatives on recommender system evaluation. In *15th Americas Conference on Information Systems 2009*. Curran, Red Hook, NY, 4510–4518.

[56] Robert B. Zajonc. 1968. Attitudinal effects of mere exposure. *Journal of personality and social psychology* 9, 2, Pt.2 (1968), 1–27. https://doi.org/10.1037/h0025848

[57] Leslie A. Zebrowitz, Judith A. Hall, Nora A. Murphy, and Gillian Rhodes. 2002. Looking Smart and Looking Good: Facial Cues to Intelligence and their Origins. *Personality and Social Psychology Bulletin* 28, 2 (2002), 238–249. https://doi.org/10.1177/0146167202282009

[58] Leslie A. Zebrowitz, Masako Kikuchi, and Jean-Marc Fellous. 2007. Are effects of emotion expression on trait impressions mediated by babyfaceness? Evidence from connectionist modeling. *Personality and Social Psychology Bulletin* 33, 5 (2007), 648–662. https://doi.org/10.1177/0146167206297399

[59] Leslie A. Zebrowitz and Gillian Rhodes. 2004. Sensitivity to "Bad Genes" and the Anomalous Face Overgeneralization Effect: Cue Validity, Cue Utilization, and Accuracy in Judging Intelligence and Health. *Journal of Nonverbal Behavior* 28, 3 (2004), 167–185. https://doi.org/10.1023/B:JONB.0000039648.30935.1b

[60] Baobao Zhang and Allan Dafoe. 2019. Artificial intelligence: American attitudes and trends. https://doi.org/10.2139/ssrn.3312874