



# Attitudes towards Social Robots (ASOR): Revisiting the Scale with Four Types of Robots

Anna Dobrosovestnova  
Human-Computer Interaction Group,  
TU Wien  
Vienna, Austria  
anna.dobrosovestnova@tuwien.ac.at

Ralf Vetter  
Artificial Intelligence and Human  
Interfaces, University of Salzburg  
Austria  
ralf.vetter@plus.ac.at

Astrid Weiss  
Human-Computer Interaction Group,  
TU Wien  
Vienna, Austria  
astrid.weiss@tuwien.ac.at

## ABSTRACT

Our work is theoretically grounded in the notion of sociomorphing contending that not all dimensions of experienced sociality with robots pertain to projection of human-like mental states i.e. anthropomorphism. To investigate dimensions of attributed sociality, we deployed the Attributes towards Social Robots scale (ASOR) in a video-based online study (n=202) with four different robots (Starship Delivery Robot, Telenoid, Blossom, Vector). The four robots were rated slightly differently which aligned with our expectations because of the differences in appearances and how they were contextualised in the videos. However, further evaluation of the statistical properties of the scale and the solicited qualitative feedback to the items pointed to limitations of the scale.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI; HCI theory, concepts and models.**

## KEYWORDS

experienced robot sociality, attitudes towards social robots, scale, different robot embodiments

### ACM Reference Format:

Anna Dobrosovestnova, Ralf Vetter, and Astrid Weiss. 2024. Attitudes towards Social Robots (ASOR): Revisiting the Scale with Four Types of Robots. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion)*, March 11–14, 2024, Boulder, CO, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3610978.3640573>

## 1 INTRODUCTION

Anthropomorphizing robots is commonly discussed in Human-Robot interaction (HRI) and social robotics (SR) in relation to developing robots with “social presence” and “social behaviors”, and efforts at understanding why people treat robots as social actors (c.f. [3], [2]). While acknowledging diverse definitions and assessments of anthropomorphism and sociality c.f. [7], [1], our work is rooted in a theoretical contention that not all experiences of sociality with robots amount to attribution of human-like mental states to robots. This theoretical anchoring is informed by our ongoing work concerning experienced sociality with autonomous delivery robots.

Relevant to this report findings from observations and interviews with passersby were: i) people engage in pro-social behaviors towards functional robots and experience them as social actors of a kind, ii) people rely on anthropomorphic language to describe these interactions, simultaneously recognizing how such language fails to represent their experience. The lack of linguistically precise tools for describing social experiences with robots has been pointed out by Seibt and colleagues in reference to the so called descriptive problem [5, 6]. To assert that not all experiences of sociality converge to imagined or projected human-like capacities – even though both lay people and researchers are prone to describe them as such (c.f. [8]) – Seibt et al. proposed *sociomorphing* as a better fit for understanding the *actual* asymmetric social capacities people attribute to and experience with robots [6]. Sociomorphing can manifest in types of experienced sociality. To descriptively anchor these, they developed the Ontology of Asymmetric Interactions (OASIS) theoretical framework based on the interdisciplinary evidence from empirical, conceptual, and phenomenological research (ibid.).

Motivated to explore how people perceive robot sociality on these theoretical grounds, we deployed the Attitudinal Stances towards Social Robots (ASOR) scale, developed by Damholdt et al. [4], in an in-between online study with four robot conditions including three social and one functional robot. We chose ASOR because of its conceptual bridge to the OASIS framework through the notion of sociomorphing [4, p.29]. Our research goals were to: (i) Investigate whether ASOR achieves different outcomes for four different types of robots; ii) Assess statistical properties of the scale to compare the results with Damholdt et al.’s; iii) Probe limitations of the scale via qualitative feedback to the items.

## 2 ATTITUDINAL STANCE TOWARDS ROBOTS SCALE

The ASOR scale was developed to gauge how social robots can influence five dimensions of social relatedness. These include: (1) *Socio-practical relatedness* which obtains when “the agent is perceived as an interaction partner capable of training or acting in “accordance with a norm”” [4, p.28]; (2) *Intimate-personal relatedness* obtains “when a person is attached to an item”; (3) *Moral relatedness* obtains “when the agent is perceived as a moral agent or as a moral patient.”; (4) *Psychological relatedness* refers to the situations when “the agent is perceived as having feelings and emotions and the perceiver is engaged in the processes of social cognition that are described by accounts of emotional contagion, sympathy, or compassion.”; (5) *Mental relatedness* obtains when “the agent is perceived as having intentions and beliefs, and the perceiver is



This work is licensed under a Creative Commons Attribution International 4.0 License.

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-0323-2/24/03.  
<https://doi.org/10.1145/3610978.3640573>

(consciously or unconsciously) engaged in the processes of social cognition that are described by theories of ‘mind-reading’.” (ibid.). Along these five dimensions, 37 items were generated and statistically assessed with a sample of 339 participants. A factor analysis yielded a three-factor solution consisting of 25 questionnaire items. The resulting 3 factors were defined as follows: (1) *ascription of mental capacities (AMC)* assessing whether the respondent perceives the “existence of an emotional and mental life, hereunder the robot’s self-understanding and the social obligations towards robot.”, (2) *ascription of socio-practical capacities (APC)* encompassing items that gauge whether the respondent believes “the robot will act in the respondent’s best interest and will act consistently over time” [4, p.48]; (3) *ascription of socio-moral status (AMS)* including the items that “reflect the status of the robot as a social agent but also the specific expected role and status it would have to the respondent” (ibid.). The limitations of the scale, as outlined by the authors, included: the need for further establishing cross-cultural validity, construct validity and predictive validity of the scale; and only moderate internal consistency of the AMS subscale suggesting the need for further elaboration.

### 3 REVISITING THE ASOR SCALE

#### 3.1 Study set-up and participants

Our study consisted of the original 25-items ASOR scale and four experimental conditions with four robots: the humanoid robot Telenoid, the small creature-like robot Vector, the DIY social robot Blossom, and the commercial autonomous delivery robot by Starship. A sample of  $n = 202$  participants (aiming at 50 per robot type) was recruited via the online research platform Prolific ©(2023). Balanced sample (50% male and 50 % female participants) and “All countries” settings were used. Participants were paid a reward of £15.38/hr for their participation in the study (the study took on average about 8 minutes to complete). One participant was excluded because of unrealistically short survey completion time.

After reviewing the informed consent information, participants were asked a demographic question (age). The resulting mean age of participants was 27.88 years ( $SD=7.77$ ). Participants were then taken to the video of one of the four robots depicting the given robot performing task(s). They were then directed to the pages displaying the ASOR scale items. The items were presented in the same order as they were presented in the original study by Damholdt et al.<sup>1</sup>. For consistency, we also preserved the 4-points Likert scale from 1 (not at all) to 4 (to a high degree) used in the original study<sup>2</sup>. As we were interested in participants’ feedback to the individual items, we included an option to respond “not applicable/other” (NA) for each of the 25 items of the ASOR scale. When the NA option was selected, participants were asked to further elaborate on their choice in an open response text field. Towards research aim (iii), we also recruited six additional participants among colleagues and friends. These participants had various disciplinary backgrounds. They were instructed to provide more extended feedback to the items they found challenging to interpret or respond to. They were asked

<sup>1</sup>We thank Malene Damholdt for sharing the video of the Telenoid robot and providing the original structure of the questionnaire.

<sup>2</sup>Please refer to the article by Damholdt et al. [4] for further details on why 4-points scale was chosen.

to fill in the survey for all 4 robots with a break between several days up to one week between each survey. Only qualitative data (feedback to the items) from these six participants were included in the analysis. These participants were not compensated monetarily.

#### 3.2 Stimulus videos

The videos were chosen to showcase (social) affordances of the robots and depicted tasks aligning with realistic use case scenarios. For the Telenoid robot, the same video as in the original ASOR study by Damholdt et al. [4] was used. The video, lasting 77 seconds, depicted a conversation between a woman and a *Telenoid*. The original video was in Danish; for our study we incorporated English subtitles. In the conversation, the woman expressed her concerns about her daughter’s slow progress at school. In response, the robot suggested to assist with tutoring. The *Starship* robot video (80 seconds long) depicted a Starship robot delivering a package to its destination. The video started with the robot introducing itself in English, followed by it driving away. The video included a scene where the robot asked a pedestrian to press the traffic light button at the street crossing and thanking the person for granting the request. The video of the *Vector* robot (72 seconds long) consisted of interactions between a woman and the robot. These included: the robot “waking up” following the woman’s command, the robot responding to the question about the identity of the woman, the robot “purring” in response to being “petted” on its back, and reporting the weather. The video of the *Blossom* robot (73 seconds long) depicted a woman inquiring the robot: “Hey Blossom, how are you?”. In response, the robot moved its body and the woman commented jokingly “You seem happy today” and asked whether it would like to watch a movie. The rest of the video depicted the woman and the robot in front of the laptop screen, with the robot moving in reaction to the animated film.

### 4 ANALYSIS AND FINDINGS

#### 4.1 Insights on the four robots

To analyze how people assessed the four robots on the original ASOR scale, we computed the scales by averaging the items. Through that, the subscales become continuous variables. Table 1 shows the means and standard deviations for each robot and subscale:

**Table 1: Ratings for each original ASOR scale for each robot**

Robot	AMC		APC		AMS		
	n	mean	SD	mean	SD	mean	SD
Starship	51	1.90	.66	2.81	.50	2.82	.65
Telenoid	50	2.10	.67	2.86	.56	<b>2.59</b>	.65
Blossom	50	2.10	.66	<b>2.53</b>	.64	2.96	.59
Vector	50	1.97	.66	2.84	.54	3.11	.57

All four robots were rated rather low on the ascription of mental capacity, but higher on the ascription of socio-practical capacity and social-moral status. The results for the AMC and APC are in line with our expectations. For the AMS scale, we were surprised by the high ratings and by the fact that Telenoid was rated the lowest. We assume that the results on the AMS subscale are potentially

inverse. For this subscale calculation, we reversed the items, as suggested by Damholdt et al. However, when looking at the semantic core of the original AMS items (see Table 2), reversing seems counter-intuitive. A One-Way ANOVA with a subsequent Tukey HSD post hoc test partly supported our assumptions. There was a significant effect of the robot type on the ascription of AMS:  $F(3, 197)=6.54, p<.00 (\eta^2=.09)$ . The post hoc tests revealed significant differences between the Telenoid group and the Blossom and Vector group, but not Starship which to us was slightly unexpected, but it is to our conviction also related to the fact that the items of this subscale did not translate to the Starship use case well because they are tailored to social (by design) robots. Another significant difference was found for the APC subscale:  $F(3,197)=3.8, p<.01 (\eta^2=.06)$ . Here the post hoc tests revealed significant differences between the Blossom group and the Telenoid and Vector group, but again not Starship. The  $\eta^2$  values indicated medium effects, however the actual differences in the mean ratings were small.

## 4.2 Evaluating the statistical properties of ASOR

Towards the research goal (ii) of exploring the statistical properties of the ASOR scale, we performed a *reliability analysis* with listwise deletion on all data ( $n = 201$ ). Both the AMC ( $n = 196, \alpha = .88$ ) and APC ( $n = 187, \alpha = .80$ ) subscales had satisfying reliability. The AMS subscale had low reliability ( $n = 197, \alpha = .56$ ). Given no AMS item stood out for improving reliability by deletion, this pointed to uncertainty regarding the dimensionality of the ASOR scale and confirmed the need for further elaboration of the AMS construct, as suggested by Damholdt et al.

We then conducted an *exploratory principle factor analysis* on the 25 ASOR items with varimax rotation to test if the factor structure as described by [4] would be replicated in our data. Pairwise deletion was applied to retain the maximum of possible data points considering the “not applicable” option. At most, 5 participants were excluded for an item (see Table 2). We performed the Kaiser-Meyer-Olkin test ( $KMO = .845$ ) to measure the sampling adequacy and Bartlett’s test of sphericity ( $BTS = [\chi^2] (300) = 1860.635, p = <.001$ ). The results indicated that our data were appropriate for analysis.

An initial inspection of the eigenvalues showed 6 factors with eigenvalues above 1. The first four factors explained a combined 51.7% of the total variance (27.3%, 10.6%, 8.2%, 5.6%) and had eigenvalues between 6.8 and 1.4, while the fifth and sixth factor explained 4.4% and 4.2% of the total variance and had eigenvalues slightly above 1 (1.09 and 1.04). The scree plot indicated retaining a 4 or 6 factor solution would be justified. Thus, the factor analysis did not replicate the expected 3 factor solution. We investigated the rotated component matrix to identify which items constituted the factors in our solution, and to establish which items caused ambiguity compared to the original ASOR scale.

We specified three criteria for factor interpretation: (a) An item needed a loading above 0.6 to be allocated to a factor. (b) An item was allocated to a factor if the difference between its highest loading and any other loading was at least 0.2. (c) A factor was suitable for interpretation if at least four items adhered to the first two criteria. The overall structure of Factor 1 for the most part aligned with the AMC as in [4] and included six items that primarily gauged people’s

expectations about the robot having mental and emotional states (AMC B07, AMC B08, AMC B09, AMC C01, AMC C02, AMC C03). The remaining five of the original AMC items (AMC B05, AMC B06, AMC C04, AMC D01) loaded elsewhere than the core group. The clustering of the items under Factor 2 (APC A04, APC A06, APC A07) suggested a semantic core that aligned with the ascription of the socio-practical capacities as predicted by Damholdt et al. However, five items of the original APC subscale (APC A01, APC A02, APC B02, APC B03, APC D03, APC D05) loaded elsewhere, thus deeming Factor 2 overall unsuitable for interpretation per our criteria. The clustering under Factor 3 (AMC B06, AMC D01), based on the semantic core of the two items, suggested the underlying construct tapped into the dimension of perception of the robot as a moral patient, though again the factor included only two items. Factor 4 consisted of only two items (APC A02, AMC C04). Factor 5 consisted of the grouping of the three items gauging ascription of socio-moral status (AMS A03, AMS A05, AMS D04). Semantically, these items were associated with a feeling of discomfort experienced in relation to the robot. On Factor 6, only one item loaded sufficiently (APC D05). In sum, these results suggest that, though there is some alignment between the factor structures resulting from our study and the one of [4], only 1 factor (AMC) performs as predicted.

## 4.3 Qualitative feedback to the items

For reporting, we clustered the qualitative feedback to the items based on the kind of issues raised by the participants (these included both participants online and the additional participants): **Domain specificity:** Qualitative feedback to items APC A01 and APC A02 suggested it mattered to participants whether a specific robot was developed to perform the task in question. For instance, in the case of the Starship robot, participants indicated they struggled to see how it would make sense to take medical or financial advice from it: “This is too far away from its intended use to answer for me.” **Lack of transparency about how the robot arrives at a decision:** In response to items gauging whether participants trust the information provided by the robot (e.g., APC A07), participants indicated it was a challenge for them to respond because they lacked information with regard to the process leading to a specific output: “I consider the robot as just ‘front end’ in this case and would probably try to disregard how good/bad it is.” **Lack of affordances to perform an action:** Selected items assumed specific capabilities related to the robot design e.g., dialogue-based interactions (e.g., APC A02, APC A07). E.g., for Blossom, participants pointed out they did not see how the dialogue assuming items were relevant. **Item containing two questions:** In the feedback to the item APC A07 participants pointed out that the item in fact contained two, and not one question. **Lack of interpretation anchoring with regards to mode of enactment:** One of the common problems identified in qualitative feedback was that participants did not know whether the items gauged if the robot could have a genuine phenomenological experience, or merely mimic or simulate e.g., an emotion or a mental state (e.g., AMC B08, AMC B09). **Recognition of dependency on external actors such as programmers:** Similarly, for the items that were implicitly entangled with the notion of agency (e.g., AMC B07, AMC C02), participants indicated that it was not clear to them whether the item probed the robot independently carrying out a

Item	Item Code*	Factor loading
1. Do you believe that the robot can feel sad?	AMC B08	F1 (0.854)
2. Do you believe that the robot can be happy?	AMC C01	F1 (0.847)
3. Do you believe that the robot can have hobbies and interests?	AMC C02	F1 (0.739)
4. Do you believe that the robot can feel pain?	AMC B09	F1 (0.744)
5. Do you think that the robot can wish for something?	AMC B04	NF
6. Do you believe that the robot has a sense of justice?	AMC B07	F1 (0.616)
7. You know who you are. Do you think that the robot knows who it is?	AMC C04	F4 (0.621)
8. Do you think that the robot can understand a joke?	AMC C03	F1 (0.652)
9. Do you think that you would help fulfill the robot's wish if it asked you to (e.g. take it with you on holiday)?	AMC B05	NF
10. Do you think that it matters how people treat the robot?	AMC B06	F3 (0.711)
11. I think that I would feel sorry for the robot if I saw others being evil to it (e.g. kick or speak badly to it).	AMC D01	F3 (0.775)
12. Do you think that you will trust what the robot tells you?	APC A04	F2 (0.778)
13. Do you believe that the robot will give you the best possible advice?	APC A06	F2 (0.694)
14. Do you think that you can give an order to the robot?	APC D03	NF
15. Do you think that the robot gives correct information, and that it always gives the same answer to a specific question (e.g. ingredients for meat sauce)?	APC A07	F2 (0.691)
16. Do you think that you would take the robot's advice on choice of medication?	APC A01	NF
17. If you owned the robot, would you consider giving it a human name?	APC B03	NF
18. Do you think that you would take the robot's advice on personal finances (e.g. choice of pension fund)?	APC A02	F4 (0.728)
19. Do you believe that the robot would be able to recognise you, if it had met you more than once?	APC D05	F6 (0.718)
20. Do you think that you would like doing things with the robot (e.g. play Ludo, chit-chat, learn a new language, or cook)?	APC B02	NF
21. If you had the robot at home, would you store it in a broom cupboard?	AMS B01	NF
22. Would you feel embarrassed if other people saw you hugging the robot?	AMS D04	F5 (0.705)
23. Do you think that you would feel insecure being with robot, if it were more intelligent than you?	AMS A03	F5 (0.719)
24. I think I would be annoyed, if the robot cut me off during a conversation.	AMS D02	NF
25. Do you think that you will be afraid of the robot?	AMS A05	F5 (0.652)

\*The first three letters of the item codes indicate the original ASOR subscales

**Table 2: Item wording, codes, and primary factor with loading. NF stands for not assignable to factor**

decision, or it would be a result of someone else programming it to behave in a specific way: *“If it was programmed to do so.”* **Lack of context to interpret the question:** One of the items that yielded a lot of feedback was AMS A03. Participants indicated the term intelligence is just too broad, and it was a challenge for them to respond to this question without any additional context provided: *“What do you mean here? Being in one room with it or in any other kind of relationship? It depends. Intelligence is also subjective until some point, I guess”*. Participants experienced similar issue with the item APC A06, indicating they did not know what “best possible” meant, and what kind of advice that would be: *“Advice on what? advice on where to cross the road or advice on whether I should take out a loan and invest it in bitcoins?”*.

## 5 CONCLUSION

We conclude ASOR is a promising – though imperfect – tool to assess dimensions of attributed sociality. Based on our experience, there are several weaknesses of the scale beyond those identified by Damholdt et al. First, though the descriptive problem is emphasized, and distinction between modes of simulations (functional replication, imitation, mimicking, displaying, approximating) is made in the context of the OASIS [6], many ASOR items rely on anthropomorphic language conflating these differences and leading to conflicting interpretations. Second, the weak(er) performance of the APC and AMS subscales may have to do, in line with the

descriptive problem, with the ambiguity of certain terms (e.g., intelligence, recognition), but, also with mixing of perspectives. Some items probe the expectations with respect to the robot capabilities, while others, assuming specific robot capabilities, probe anticipated human reactions in underdetermined situations. The challenge with the APC scale may be further aggravated by the fact that no established unified theory exists with respect to socio-practical relatedness. Simply put, while we believe socio-practical relatedness will play an increasingly important role when it comes to embedding robots in everyday lives, we are yet to formulate sound conceptual grounding concerning what socio-practical relatedness actually is and what it involves. Moreover, ASOR faces challenges common to other scales in HRI, such as probing post-hoc rationalizations, and, as is the case with our study, doing so based on a one-off exposure to a video stimuli as opposed to the actual phenomenological experience of interacting with a robot. Recognizing the challenges laid out above, our aim in a follow-up study is to introduce and test concrete incremental steps towards the improvement of the ASOR scale at the level of individual items and semantic cores of the subscales.

## ACKNOWLEDGMENTS

The first author is a recipient of a DOC Fellowship of the Austrian Academy of Sciences (DOC/25861).

## REFERENCES

- [1] Luisa Damiano and Paul Dumouchel. 2018. Anthropomorphism in Human–Robot Co-evolution. *Frontiers in Psychology* 9 (2018). <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00468>
- [2] 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](https://doi.org/10.1016/S0921-8890(02)00374-3)
- [3] Julia Fink. 2012. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *Social Robotics (Lecture Notes in Computer Science)*, Shuzhi Sam Ge, Oussama Khatib, John-John Cabibihan, Reid Simmons, and Mary-Anne Williams (Eds.). Springer, Berlin, Heidelberg, 199–208. [https://doi.org/10.1007/978-3-642-34103-8\\_20](https://doi.org/10.1007/978-3-642-34103-8_20)
- [4] Malene Flensburg Damholdt, Christina Vestergaard, Marco Nørskov, Raul Hakli, Stefan Larsen, and Johanna Seibt. 2020. Towards a new scale for assessing attitudes towards social robots: The attitudes towards social robots scale (ASOR). *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems* 21, 1 (Jan. 2020), 24–56. <https://doi.org/10.1075/is.18055.fle>
- [5] Johanna Seibt. 2014. Varieties of the 'As If': Five Ways to Simulate an Action.. In *Robophilosophy*. 97–104.
- [6] Johanna Seibt, Christina Vestergaard, and Malene F. Damholdt. 2020. Sociomorphing, not anthropomorphizing: towards a typology of experienced sociality. *Culturally Sustainable Social Robotics—Proceedings of Robophilosophy* (2020), 51–67.
- [7] Sam Thellman, Maartje de Graaf, and Tom Ziemke. 2022. Mental State Attribution to Robots: A Systematic Review of Conceptions, Methods, and Findings. *J. Hum.-Robot Interact.* 11, 4, Article 41 (sep 2022), 51 pages. <https://doi.org/10.1145/3526112>
- [8] Sam Thellman and Tom Ziemke. 2019. The intentional stance toward robots: Conceptual and methodological considerations. In *The 41st Annual Conference of the Cognitive Science Society, July 24-26, Montreal, Canada*. Cognitive Science Society, Inc., 1097–1103. <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1338970>