A Database for Kitchen Objects: Investigating Danger Perception in the Context of Human-Robot Interaction

Jan Leusmann
LMU Munich
Germany
jan.leusmann@ifi.lmu.de

Carl Oechsner
LMU Munich
Germany
c.oechsner@ifi.lmu.de

Johanna Prinz
LMU Munich
Germany
johanna.prinz@campus.lmu.de

Robin Welsch
Aalto University
Finland
robin.welsch@aalto.fi

Sven Mayer
LMU Munich
Munich, Germany
info@sven-mayer.com

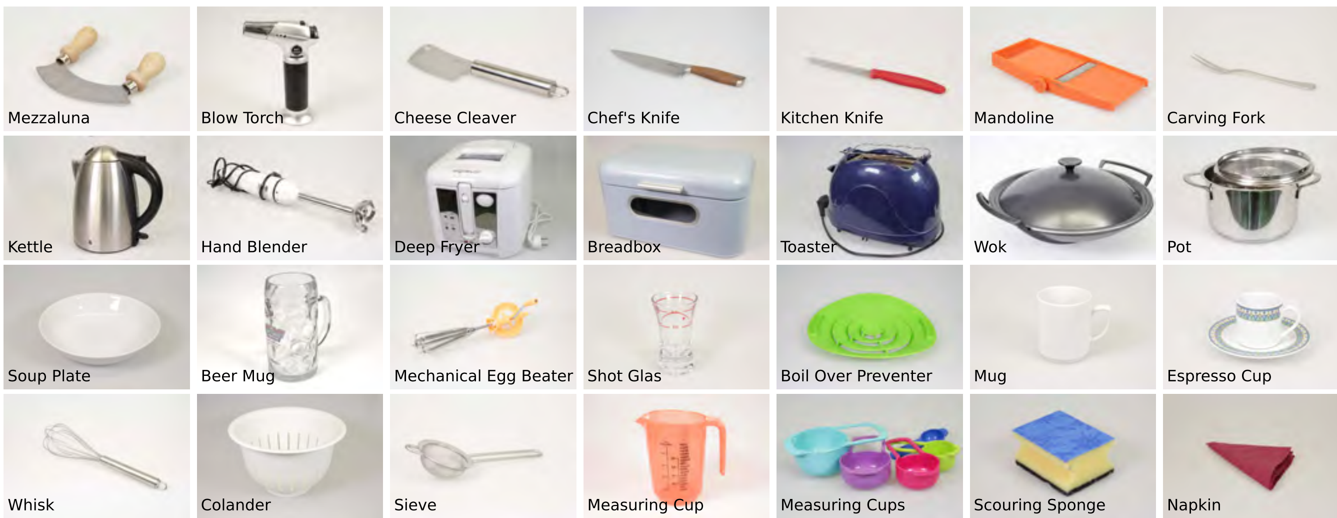


Figure 1: Subset of images from kitchen objects from our database, sorted by their perceived danger perception across the three Scenarios (mezzaluna having the highest danger perception, napkin having the lowest).

ABSTRACT

In the future, humans collaborating closely with cobots in everyday tasks will require handing each other objects. So far, researchers have optimized human-robot collaboration concerning measures such as trust, safety, and enjoyment. However, as the objects themselves influence these measures, we need to investigate how humans perceive the danger level of objects. Thus, we created a database of 153 kitchen objects and conducted an online survey (N=300) investigating their perceived danger level. We found that (1) humans perceive kitchen objects vastly differently, (2) the object-holder has a strong effect on the danger perception, and (3) prior user knowledge increases the perceived danger of robots handling those objects. This shows that future human-robot collaboration studies must investigate different objects for a holistic image. We contribute a wiki-like open-source database to allow others to study predefined danger scenarios and eventually build object-aware systems: https://hri-objects.leusmann.io/.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

human-computer interaction, human-robot interaction, dataset, kitchen, robots, bayesian mixed models

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

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585884

ACM Reference Format:
Jan Leusmann, Carl Oechsner, Johanna Prinz, Robin Welsch, and Sven Mayer. 2023. A Database for Kitchen Objects: Investigating Danger Perception in the Context of Human-Robot Interaction. 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, 9 pages. https://doi.org/10.1145/3544549.3585884

1 INTRODUCTION

Today's voice assistants (e.g., Siri) can answer questions and access APIs to control smart home environments. However, they cannot support users through direct actions with the physical world. Domestic robots like vacuum robots have the capability to interact with the users' environment. However, they are still only capable of solving one specific task. In the future, collaborative robots (cobots) can bridge this gap, either interacting with users directly or with physical objects around them. The humans' role during this interaction can range from only observing the other actor to working cooperatively together [23]. To interact with the world, the cobot will use common tools (e.g., a drill) or directly move objects (e.g., a dinner plate). In most cases, the robot manipulates the environment, and especially in more collaborative settings interacts with the same objects as humans or does a hand-over [22]. Augmented Reality systems can be used to show the intent of the robot and the future position of objects [26, 34]. While we know that interactions with robots carrying objects are associated with subjective feelings of danger [25, 30, 32], it is unclear how users' danger perception differs between various objects and which objects are especially dangerous when being carried by a robot as compared to a human.

We know that danger perception is influenced by different situational or personal factors [9, 10, 14, 29]. In particular, people holding objects of varying danger has an effect on danger perception, whereas holding a dangerous object also increases the perceived danger [1, 35]. Robots additionally are associated with a feeling of eeriness and scariness themselves that may likewise be exacerbated when carrying dangerous objects [25, 30, 32]. However, from a technical standpoint, cobots are designed to ensure users' safety by including collision detection via, e.g., force sensors [6, 8, 13]. Thus, users are not actually in danger but may still perceive the situation as dangerous. This is further complicated by the fact that future cobots may be able to carry and hand over any kind of object in the physical space. To design human-robot interaction (HRI) effectively, danger elicited by cobots has to be understood in tune with the type of object that is maneuvered by the robot. To the best of our knowledge, there is currently no extensive open-source image database of kitchen utensils, especially not in the context of danger perception in HRI. As direct interaction between robots and humans is still very novel, the perception of danger is still influenced by a multitude of things. By understanding how the wielded object affects this interaction we can remove one element of this danger perception equation.

We aim to understand how danger is perceived for various objects to allow for optimal HRI. In this work, we will focus on the danger levels of kitchen objects as a prominent example of HRI, cf. [5, 21, 31, 33]. In the kitchen, some of these objects can oftentimes be dangerous. Kitchen objects like knives or a hot pan can directly hurt users. Thus, we usually treat them more carefully than other items. In detail, we investigated the danger perception of objects from three different Scenarios: when the object is used by oneself, when used by another human close by, or when used by a cobot close by. Therefore, we curate a dataset of 153 unique kitchen objects. Our dataset will also enable future research in the domain of domestic cobots, e.g., kitchen cobots. In an online survey with 300 participants, we gain an understanding of the differences in the perception of danger from kitchen objects. In this survey, participants had to rate how dangerous they feel objects under the three Scenarios: *self*, *other*, and *robot*. With this, we can understand if users perceive cobots similar to other humans. The difference between a robot and a human close by using an object is of special importance to understanding long-term effects.

Our investigation of 153 different kitchen objects showed that they are perceived vastly differently, ranging from 0 to 100 on a 101-point danger scale. The danger levels are generally consistent in the three interaction Scenarios; however, our results showed that the same object used by a cobot is more dangerous than when used by *self* or *other*. Interestingly, when the cobot uses an object, fire is perceived as very dangerous, even more than sharpness, but sharp objects like knives are still among the most dangerous objects. Our results highlight the need to investigate not only a single object to understand optimal human-robot interaction. Therefore, we open-source our database, allowing others to investigate their scenarios. Moreover, as our current database is only a small step in identifying a large number of objects that occurs in the HRI context, the next step is to enlarge the database now that we have shown there is a large impact of the object on the perceived dangers.

2 DATASET

To the best of our knowledge, there is no extensive dataset of items with pictures, especially not kitchen objects. Most datasets in the kitchen domain are video-based datasets, mostly used for action segmentation tasks [7, 12, 17]. Other extensive datasets collected 3D-models of general objects [3, 18, 27] However, none of these is an extensive dataset with labeled images of kitchen objects together with an investigation of the perceived danger level.

Thus, we create an extensive open-source dataset of objects in the following. In this initial step, we focus only on kitchen objects as they already comprise many items and cover a wide range of object properties. However, we will extend the list with more everyday objects in the future. We aim to set a foundation for HRI research to investigate how various items impact human-robot collaboration. In the first step, we curate a list of names of kitchen objects and representing images.

2.1 Item Generation

We created the first set of kitchen items in a 3-hour workshop with four authors. Here, we started by investigating different kitchen item lists online. During this, we also discussed the granularity of items, e.g., when to group items. We decided to group items as long as the functionality and intended use were the same. Thus, we grouped all items in one group even if they have a different color, form, and size, such as a red plate and a blue plate. However, when the functionality of the intended us is different, we would put them in different groups, e.g., paring knife and table knife.

With this initial list, we asked 30 additional people if they were missing any kitchen items. The people were from four different continents (South America, North America, Europe, and Asia) to ensure a diverse set of cooking styles and cultures were represented in the dataset.

In a second workshop, we discussed all item names with a North American citizen to ensure consistent language. As we created this

Figure 2: The vision of a future kitchen, where cobots support users with their cooking tasks. Source: Oechsner et al. [21] under Creative Commons license.

initial item list for human collaboration scenarios, we removed it with the following exclusion criteria: The item is typically not moved and, thus, typically has a dedicated spot in the kitchen, e.g., microwave or fridge. This resulted in our final set of 153 items. In this stage, we also added alternative names for the items. Finally, we took pictures of all the items enabling everyone to understand the item even if the name was unclear. See Figure 1 for 28 example items.

3 STUDY DESIGN

We conducted an online survey to understand how the perception of the 153 kitchen objects differs. We asked 300 participants to rate their perceived danger of each item. As items can be used in various scenarios, we investigate three Scenarios: *self*, *other*, and *robot*. In Scenario *self*, the item is used by one's self. From a different perspective, during *Human-Human Collaboration* when cooking together, humans perceive the item from a bystander Scenario (condition: *other*). Here, the perceived danger should change. In line with this is *Human-Cobot Collaboration*, where the item is being used by a cobot, not a human (condition: *robot*). See such an example in Figure 2.

3.1 Questionnaire

We designed a four-part questionnaire to discover the perception of various kitchen items, see Section A.1. We informed participants beforehand that this questionnaire was about the perception of different kitchen objects. In the first part, we asked participants about their demographics (e.g., age, gender), and technical expertise using the affinity for technology interaction questionnaire [11], cooking expertise, and cooking frequency per week. For the next three parts, each participant was assigned ten random kitchen objects out of the 153 kitchen objects from our dataset. However, we ensured that each item would have at least 15 results. We informed participants before each part about what they should envision for the next ten presented kitchen objects. In the second part, we asked the participants about their *familiarity* with the shown kitchen object, whether they *own* it, and lastly, about how *dangerous* they feel that object is when they are using it themselves. The third part (condition: *other*) showed the same ten objects in the same order and just asked them how dangerous they find that item to be when someone else next to them used it. Lastly (condition: *robot*), we asked about the danger perception when a robot next to them would use that object. Before that, we showed the participants a picture of a potential vision of such a robot in the kitchen being a regular 6-Degree-of-Freedom robotic arm hanging from the ceiling, see Figure 2. Lastly, we asked participants for general feedback via a text field.

Participants answered the questions with sliders ranging from 0 (Strongly Disagree) to 100 (Strongly Agree), with the initial position in the middle. We did not display the numbers or ticks on these sliders, c.f. Matejka et al. [20] as such sliders have been shown to lead to more precise responses [24]. Thus, we chose them over standard Likert scales.

3.2 Preprocessing

During and after the survey, we rejected and removed participants' responses, which either (1) were obvious low-effort responses, or (2) had a miss-match between demographic data in prolific and our survey. To find low-effort answers, we automatically screened the responses. We flagged responses found by our automatic approach and then manually looked at the responses to decide whether we would reject that participant. Our questionnaire mainly consisted of questions that had to be answered via a slider to select a value between 0 and 100. Thus, we first looked if all questions had been answered with the same value, so if people just took the slider and dragged it completely to the left (0) or the right (100). Next, we handpicked a set of very obviously dangerous (e.g., chef's knife, cleaver) and not dangerous items (e.g., sponge, napkin, dish towel) and screened responses with unfitting danger perception scores. When screening the responses, we would never reject participants for just one outlier, but we would always consider the global scope of the response. We also excluded responses where the age and sex provided via prolific did not match their responses in our survey.

3.3 Participants

We recruited all our participants via Prolific[1]. In total, we accepted responses from 300 participants (154 female, 142 male, and 4 non-binary) balanced for six continents (Africa, Asia, Europe, North America, Oceania, and South America), resulting in 50 participants from each. On average, the participants were between 19 and 73 years old (M = 32.6, SD = 11.3). Out of the 300 participants, 175 stated to work full-time, 69 part-time, and 56 stated a different employment status. On average, participants took 10.1 (min = 4.0, max = 57.2, SD = 5.0) minutes to finish the survey. The average household size of participants was 3.21 persons (min = 1, max = 10, SD = 1.48). The majority of participants have a bachelor's degree or higher (149 bachelor's, 42 master's, and 11 Ph.D.).

4 RESULTS

In the following, we will present the results from the questionnaire responses collected in October 2022. Every participant gave results for 10 out of the 153 randomly selected objects. We collected survey responses iteratively to make sure we had a balanced amount of responses for every object. Every object got at least 15 responses (M = 19.477, max = 30.00, SD = 3.424). We then created one dataset

[1] https://www.prolific.co/

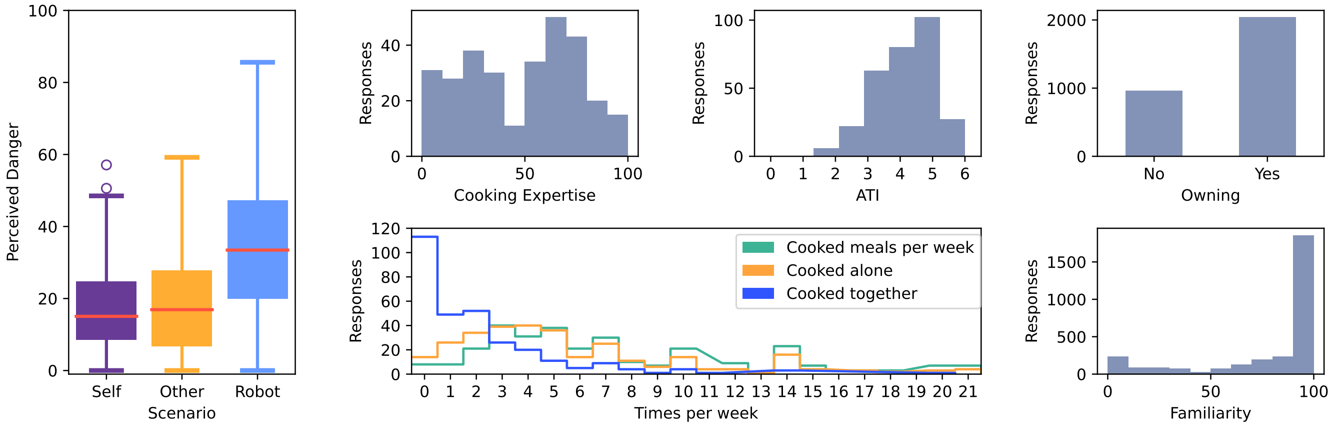


Figure 3: (A) Perceived Danger for each SCENARIO (B) Histogram of Cooking Experience (C) ATI scores (D) Distribution of owned objects (E) Cooking frequency, split up in meals cooked per week in general, meals cooked per week alone, and together. (F) Histogram of Familiarity with the objects

including all danger scores for the three SCENARIOS and the other measurements, c.f. Figure 3. We use a Bayesian approach to data analysis in the form of Bayesian linear mixed models (BLMM). The Bayesian approach has been taken up lately in HCI [15, 16, 19, 28] as it presents several advantages to classical statistics. Kay et al. [16] explain the advantages of Bayesian statistics in HCI that are also relevant to our study. We can incorporate prior knowledge about the to-be-estimated parameters, get probability functions for each parameter instead of point estimates, and the inferences are more robust to non-normality in the data. We can incorporate a large number of factors into our analysis and, most importantly, are not testing binary hypothesis which would mandate correction procedures. To model the data, we fitted several linear mixed models using brms [2] models in STAN [4] using 4 chains and 20.000 iterations with a warm-up of 10% for each. Effects were considered meaningful when there was a particularly low probability ($p_b <=$ 2.5%) of the effect being zero or the opposite. In addition to the median of the parameter, we calculated the High-Density Interval (HDI) at 95% of the posterior distribution for all parameters, which indicates the possible range of effects given the data, alongside the median of the respective parameter.

4.1 Danger

We first investigated how the SCENARIO as fixed effect affects danger perception in a mixed effects model with random intercepts for participants and stimuli with each random slope for SCENARIO. *Self* was specified as a baseline with treatment contrasts. We found that the SCENARIO robot has a distinguishable effect on the perceived danger ($p_b < 0.001$, $Median = 16.272$, $HDI_{95\%} = [13.866, 18.770]$). Thus, objects carried by robots increased ratings by about 16 points in our model. There is also a distinguishable effect from the SCENARIO other on the perceived danger ($p_b = 0.005$, $Median = 1.656$, $HDI_{95\%} = [0.422, 2.934]$). Therefore, objects carried by others elicited slightly larger ratings of perceived danger, see Figure 4.

To explore the effects of our other variable (familiarity, owning, technical expertise, cooking experience, and cooking frequency), we expanded the model by one fixed-effects level and also estimated the interaction for each SCENARIO. The general results of the questions leading to these factors are depicted in Figure 3. We found a distinguishable effect between the parameter estimating the *robot* × familiarity interaction ($p_b < 0.001$, $Median = 0.065$, $HDI_{95\%} = [0.026, 0.103]$). In the expanded model for owning, we found a distinguishable effect for the parameter on owning objects and *robot* ($p_b < 0.001$, $Median = 4.909$, $HDI_{95\%} = [2.321, 7.595]$). The model that included technical expertise was similar to the only-scenario fixed effect model, i.e., we did not find distinguishable effects for technical expertise and *robot* $p_b > .025$. There was also no distinguishable interaction effect for the model including cooking experience, i.e., *robot* × cooking experience parameter($p_b = 0.251$, $Median = 0.024$, $HDI_{95\%} = [-0.047, 0.093]$). Note that in this model, we found that the parameter for *other* × cooking experience($p_b =$ 0.001, $Median = 0.063$, $HDI_{95\%} = [0.023, 0.105]$) was distinguishable. However, whether the difference of *self* to *other* SCENARIO changed with cooking experience was not relevant to our study. Likewise, therefore the model with *robot* ×cooking frequency, the effect was centered around zero ($p_b = 0.202$, $Median = 0.160$, $HDI_{95\%} = [-0.128, 0.536]$). From this, we can state that the SCENARIO *robot*, together with the familiarity and ownership of the given object, has an effect on the perceived danger.

4.2 Object Danger

Figure 4 depicts the perceived danger levels for all our 153 objects. The top five most dangerous objects from *self* SCENARIO are a mandoline, mezzaluna, paring knife, pressure cooker, and blow torch. The five least dangerous items from *self* SCENARIO are: whisk, scouring sponge, napkin, serving spoon, and kitchen scale. The top five most dangerous objects from *other* SCENARIO are a mezzaluna, matchbox, bread knife, knife sharpener, and cheese cleaver. The five least dangerous items from *other* SCENARIO are dish cloth, kitchen paper, piping bag, measuring spoon, and tea spoon. The top five most dangerous objects from *robot* SCENARIO are a matchbox, blow torch, mezzaluna, bread knife, and kitchen knife. The five

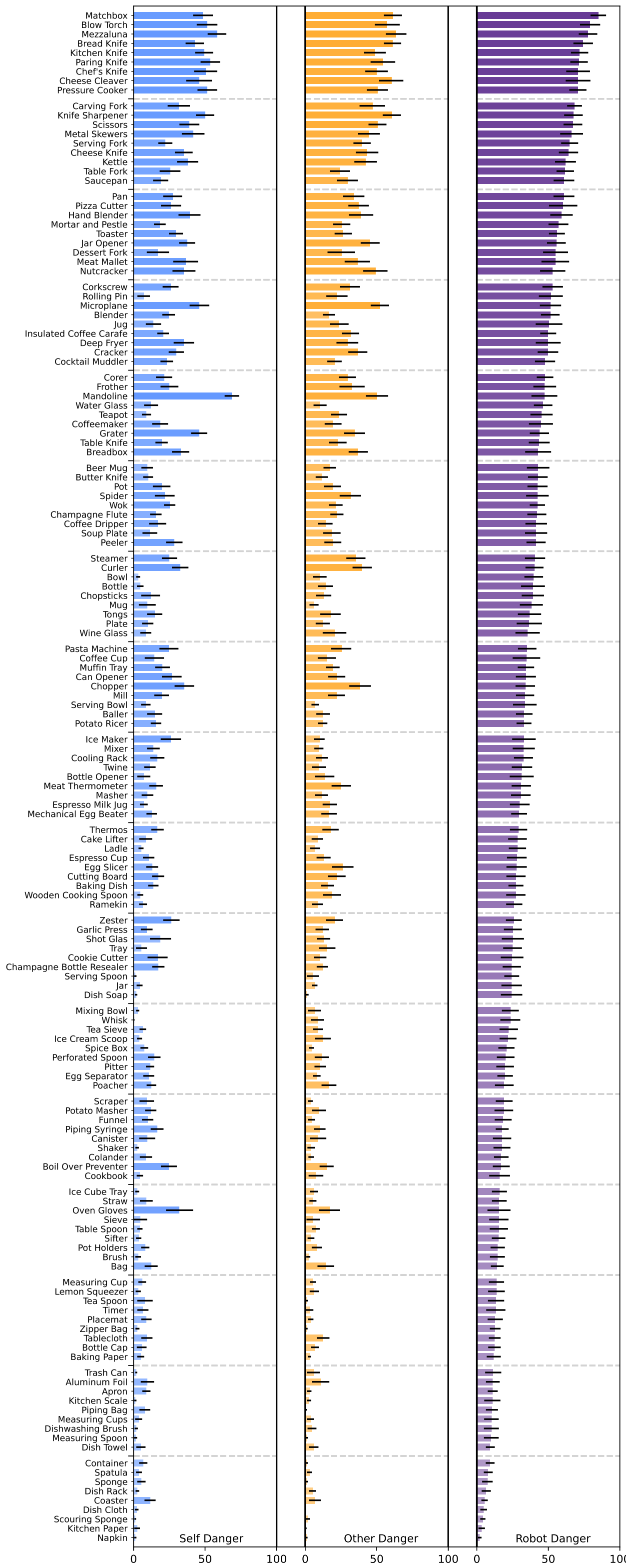
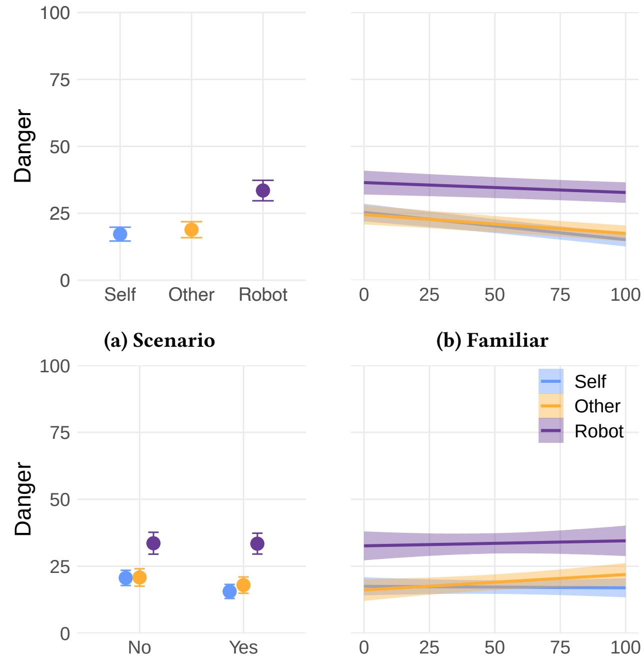


Figure 4: The mean danger levels for the different objects. Error bars denote +- 1 *SD* / n (for visibility). All danger levels are sorted by robot danger.



(a) Scenario (b) Familiar

(c) Owning (d) Cooking Expertise

Figure 5: Results of predicted values for fixed effects in the first simple SCENARIO model (a), with distinguishable effects on the perceived danger per SCENARIO. (b) Object familiarity decreases perceived danger for when humans handle that objects, but not for robots. (c) Owning objects increases the difference of perceived danger between humans and robots. (d) Cooking expertise increases the slope of perceived danger when other humans handle that object.

least dangerous items from *robot* SCENARIO are napkin, kitchen paper, scouring sponge, dish cloth, and coaster. Objects with small perceived danger values are all items, which are not sharp, fire, and can not break when dropped. Dangerous items are mostly sharp or entail fire; thus, there are direct risks of being hurt.

4.3 Object Danger Difference

We used the posterior predictive distribution of our simple SCENARIO model and predicted the difference of random slopes for *robot* vs. *self* for each of the objects (random intercept) ignoring individual variation between and within subjects and the fixed effect. We visualized this difference in Figure 6. We found that there are significantly more items with a negative *robot-self* difference without zero-crossings (18), i.e., people have a higher danger perception when using these items themselves, than a positive difference (4). Potentially obvious, but also interestingly, is that all objects used to grate or slice are rated with way higher *self*-danger than *robot*-danger.

5 DISCUSSION

As expected, we found that the perceived danger of items varied substantially; see Figure 4. Soft, dull, and unbreakable items, such as a napkin or containers, are among the items that are perceived as the least dangerous. In contrast, sharp and fire-related items are rated as the most dangerous, e.g., kitchen knives, and blowtorch. This emphasizes the importance that whenever investigating users interacting with items, it is crucial to ensure that a range of different danger levels is investigated to produce generalizable results.

We found that the danger level for interacting with an item (*self*) and someone else interacting with an item close by (*other*) are descriptively similar. However, the perceived danger when a *robot* is working with objects is much more pronounced, see Figure 3a. Thus, when investigating HRI scenarios, it is important to ensure that users are comfortable with the interaction, as everything is perceived as much more dangerous.

At the same time, we uncovered that not all danger levels increase when a robot uses an item compared to oneself; see Figure 4. We find three categories. First, objects that are relatively more dangerous for *self* as compared to the *robot* (all objects with HDI's outside not crossing 0 and being negative). These are, in particular, objects that are sharp or have sharp edges, including mandoline, grater, mezzaluna, or knife sharpener. Second, objects that are indifferent to Scenario (crossing the zero line), e.g., tablecloth, mill, or muffin tray. And lastly, objects that are more dangerous when being carried by the robot as compared to carrying them oneself (positive; not crossing the zero-line), e.g., bowl, water glass, or bottle. Six of the ten items with the highest positive difference are objects that will most likely break when dropped (porcelain, glass), highlighting that the difference in danger perception is not equal to the generally perceived danger, as these objects are rated with overall danger levels around the mean.

Overall, we found that familiarity with an item decreases perceived danger no matter the scenario, see Figure 5b. In detail, we found that more familiarity does not decrease the perceived danger the same in the *robot* condition as compared to the *self* condition. Interestingly and against intuition, we found that the affinity to technology measured using the ATI questionnaire [11] did not impact the perceived danger scores. This illustrates that robots, especially in domestic scenarios, are still such a new technology that even people with higher technical interest and expertise can not fully envision this form of future interaction possibilities, indicating the research potential and also the necessity for cobots in domestic settings in the future.

6 CONCLUSION

We envision that in the future, our open-source database not only contains kitchen objects but also other household objects and even tools for fabrication, such as power drills. Therefore, we plan to add additional objects and create guidelines for others to contribute to this wiki-like database. Future work could also add 3D Models for each kitchen object to enable comparison studies in AR/VR and the real world. In this work, we investigated the difference in danger perception when an object is used by oneself, another person in close vicinity, or a cobot nearby. We found that currently, users' danger perception of objects is way higher when a cobot uses an object compared to a human. We additionally found that the difference in danger perception between self-use and robot-use can be categorized into three clusters: (1) Objects that are perceived as relatively more dangerous when used oneself (sharp), (2) objects where it does not matter whether the object is wielded by oneself or a robot (dull), and (3) objects that are perceived more dangerous when used by a robot (fragile). While we performed this investigation on kitchen objects, we hypothesize that our findings apply to all other fields where robots use various objects to accomplish tasks. Thus, our next steps will be to enlarge our database by incorporating items outside of the kitchen context and also verify our results in an actual kitchen scenario via a lab study that can consider other factors such as proxemics, context, and robot dynamics.

7 OPEN SCIENCE

We encourage the reader to build upon this research. Therefore, our experimental setup, collected data, and analysis are online available as a wiki-like open-source database: https://hri-objects.leusmann.io/. Further, we strongly encourage contributions to extend the database.

ACKNOWLEDGMENTS

During this work, Jan Leusmann was supported by Honda Research Institute Europe. Carl Oechsner was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project ID 425412993-SPP 2199.

REFERENCES

[1] Adam T. Biggs, James R. Brockmole, and Jessica K. Witt. 2013. Armed and Attentive: Holding a Weapon Can Bias Attentional Priorities in Scene Viewing. *Attention, Perception, & Psychophysics* 75, 8 (Nov. 2013), 1715–1724. https://doi.org/10.3758/s13414-013-0538-6

[2] Paul-Christian Bürkner. 2017. Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80 (Aug. 2017), 1–28. https://doi.org/10.18637/jss.v080.i01

[3] Berk Calli, Arjun Singh, James Bruce, Aaron Walsman, Kurt Konolige, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M Dollar. 2017. Yale-CMU-Berkeley dataset for robotic manipulation research. *The International Journal of Robotics Research* 36, 3 (2017), 261–268. https://doi.org/10.1177/0278364917700714

[4] Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 76 (2017), 1. https://doi.org/10.18637/jss.v076.i01

[5] Elizabeth Cha, Jodi Forlizzi, and Siddhartha S. Srinivasa. 2015. Robots in the Home: Qualitative and Quantitative Insights into Kitchen Organization. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. Association for Computing Machinery, New York, NY, USA, 319–326. https://doi.org/10.1145/2696454.2696465

[6] MichałCzubenko and Zdzisław Kowalczuk. 2021. A Simple Neural Network for Collision Detection of Collaborative Robots. *Sensors* 21, 12 (Jan. 2021), 4235. https://doi.org/10.3390/s21124235

[7] Dima Damen, Hazel Doughty, Giovanni Maria Farinella, Antonino Furnari, Jian Ma, Evangelos Kazakos, Davide Moltisanti, Jonathan Munro, Toby Perrett, Will Price, and Michael Wray. 2022. Rescaling Egocentric Vision: Collection, Pipeline and Challenges for EPIC-KITCHENS-100. *International Journal of Computer Vision (IJCV)* 130 (2022), 33–55. https://doi.org/10.1007/s11263-021-01531-2

[8] Alexander Dietrich, Thomas Wimbock, Alin Albu-Schaffer, and Gerd Hirzinger. 2012. Integration of Reactive, Torque-Based Self-Collision Avoidance Into a Task Hierarchy. *IEEE Transactions on Robotics* 28, 6 (Dec. 2012), 1278–1293. https://doi.org/10.1109/TRO.2012.2208667

[9] Mary Douglas. 2002. *Risk Acceptability According to the Social Sciences.* Routledge, London. https://doi.org/10.4324/9780203708781

[10] Mary Douglas and Aaron Wildavsky. 1983. *Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers.*

[11] Thomas Franke, Christiane Attig, and Daniel Wessel. 2018. *A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale.* Taylor & Francis. https://doi.org/10.1080/10447318.2018.1456150

[12] Georgios Georgakis, Md Alimoor Reza, Arsalan Mousavian, Phi-Hung Le, and Jana Kosecka. 2016. Multiview RGB-D Dataset for Object Instance Detection.

In *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, Stanford, CA, USA, 426–434. https://doi.org/10.1109/3DV.2016.52
[13] Amin Hamed, Mehdi Tale Masouleh, and Ahmad Kalhor. 2020. Design & Characterization of a Bio-Inspired 3-DOF Tactile/Force Sensor and Implementation on a 3-DOF Decoupled Parallel Mechanism for Human-Robot Interaction Purposes. *Mechatronics* 66 (April 2020), 102325. https://doi.org/10.1016/j.mechatronics.2020.102325
[14] Mary B. Harris and Kari C. Miller. 2000. Gender and Perceptions of Danger. *Sex Roles* 43, 11 (Dec. 2000), 843–863. https://doi.org/10.1023/A:1011036905770
[15] Matthew Kay, Steve Haroz, Shion Guha, and Pierre Dragicevic. 2016. Special Interest Group on Transparent Statistics in HCI. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1081–1084. https://doi.org/10.1145/2851581.2886442
[16] Matthew Kay, Gregory L. Nelson, and Eric B. Hekler. 2016. Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4521–4532. https://doi.org/10.1145/2858036.2858465
[17] Hilde Kuehne, Ali B. Arslan, and Thomas Serre. 2014. The Language of Actions: Recovering the Syntax and Semantics of Goal-Directed Human Activities. In *Proceedings of Computer Vision and Pattern Recognition Conference (CVPR'14)*. IEEE. https://doi.org/10.1109/CVPR.2014.105
[18] Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. 2011. A large-scale hierarchical multi-view RGB-D object dataset. In *2011 IEEE International Conference on Robotics and Automation*. 1817–1824. https://doi.org/10.1109/ICRA.2011.5980382
[19] Jisoo Lee, Erin Walker, Winslow Burleson, Matthew Kay, Matthew Buman, and Eric B. Hekler. 2017. Self-Experimentation for Behavior Change: Design and Formative Evaluation of Two Approaches. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6837–6849. https://doi.org/10.1145/3025453.3026038
[20] Justin Matejka, Michael Glueck, Tovi Grossman, and George Fitzmaurice. 2016. The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5421–5432. https://doi.org/10.1145/2858036.2858063
[21] Carl Oechsner, Sven Mayer, and Andreas Butz. 2022. Challenges and Opportunities of Cooperative Robots as Cooking Appliances. (2022), 7. https://ceur-ws.org/Vol-3154/paper14.pdf
[22] Valerio Ortenzi, Akansel Cosgun, Tommaso Pardi, Wesley Chan, Elizabeth Croft, and Dana Kulić. 2021. Object Handovers: A Review for Robotics. *IEEE Transactions on Robotics* 37, 6 (Jan. 2021), 1855–1873. https://doi.org/10.1109/TRO.2021.3075365 arXiv:2007.12952 [cs, eess]
[23] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. (2023).
[24] Ulf-Dietrich Reips and Frederik Funke. 2008. Interval-Level Measurement with Visual Analogue Scales in Internet-based Research: VAS Generator. *Behavior Research Methods* 40, 3 (Aug. 2008), 699–704. https://doi.org/10.3758/BRM.40.3.699
[25] David A. Robb, Muneeb I. Ahmad, Carlo Tiseo, Simona Aracri, Alistair C. McConnell, Vincent Page, Christian Dondrup, Francisco J. Chiyah Garcia, Hai-Nguyen Nguyen, Èric Pairet, Paola Ardón Ramírez, Tushar Semwal, Hazel M. Taylor, Lindsay J. Wilson, David Lane, Helen Hastie, and Katrin Lohan. 2020. Robots in the Danger Zone: Exploring Public Perception through Engagement. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/3319502.3374789
[26] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2019. Communicating and Controlling Robot Arm Motion Intent through Mixed-Reality Head-Mounted Displays. *The International Journal of Robotics Research* 38, 12-13 (Oct. 2019), 1513–1526. https://doi.org/10.1177/0278364919842925
[27] Ruben Sagues-Tanco, Luis Benages-Pardo, Gonzalo López-Nicolás, and Sergio Llorente. 2020. Fast Synthetic Dataset for Kitchen Object Segmentation in Deep Learning. *IEEE Access* 8 (2020), 220496–220506. https://doi.org/10.1109/ACCESS.2020.3043256
[28] Abhraneel Sarma and Matthew Kay. 2020. Prior Setting in Practice: Strategies and Rationales Used in Choosing Prior Distributions for Bayesian Analysis. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376377
[29] Chauncey Starr. 1969. Social Benefit versus Technological Risk. *Science* (Sept. 1969). https://doi.org/10.1126/science.165.3899.1232
[30] Jan-Philipp Stein and Peter Ohler. 2017. Venturing into the Uncanny Valley of Mind—The Influence of Mind Attribution on the Acceptance of Human-like Characters in a Virtual Reality Setting. *Cognition* 160 (March 2017), 43–50. https://doi.org/10.1016/j.cognition.2016.12.010
[31] Rainer Stiefelhagen, Christian Fogen, P. Gieselmann, Hartwig Holzapfel, Kai Nickel, and Alex Waibel. 2004. Natural Human-Robot Interaction Using Speech, Head Pose and Gestures. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, Vol. 3. IEEE, Sendai, Japan, 2422–2427. https://doi.org/10.1109/IROS.2004.1389771
[32] Megan K. Strait, Cynthia Aguillon, Virginia Contreras, and Noemi Garcia. 2017. The Public's Perception of Humanlike Robots: Online Social Commentary Reflects an Appearance-Based Uncanny Valley, a General Fear of a "Technology Takeover", and the Unabashed Sexualization of Female-Gendered Robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Lisbon, 1418–1423. https://doi.org/10.1109/ROMAN.2017.8172490
[33] Kristina Tornbjerg, Anne Marie Kanstrup, Mikael B. Skov, and Matthias Rehm. 2021. Investigating Human-Robot Cooperation in a Hospital Environment: Scrutinising Visions and Actual Realisation of Mobile Robots in Service Work. In *Designing Interactive Systems Conference 2021*. ACM, Virtual Event USA, 381–391. https://doi.org/10.1145/3461778.3462101
[34] Chao Wang, Anna Belardinelli, Stephan Hasler, Theodoros Stouraitis, Daniel Tanneberg, and Michael Gienger. 2023. Explainable Human-Robot Training and Cooperation with Augmented Reality. https://doi.org/10.1145/3544549.3583889 arXiv:2302.01039 [cs]
[35] Jessica K. Witt and James R. Brockmole. 2012. Action Alters Object Identification: Wielding a Gun Increases the Bias to See Guns. *Journal of Experimental Psychology: Human Perception and Performance* 38, 5 (2012), 1159–1167. https://doi.org/10.1037/a0027881

A APPENDIX

A.1 Survey

A.1.1 Demographics.

(1) How old are you? (number field)
(2) How do you describe yourself?
   - Male
   - Female
   - Non-binary
   - Self-described (text field)
(3) What is the highest degree you have received?
   - High school graduate
   - Some college but no degree
   - Bachelor's degree
   - Master's degree
   - Doctoral degree
   - Vocational education
(4) What is your current primary occupation? (text text)
(5) Which country do you currently reside in? (text text)

A.1.2 Cooking Habits & Expertise.

(1) I consider myself an expert cook. (101-point sliders from strongly disagree to strongly agree)
(2) How many people live in your household? (number field)
(3) How many times do you prepare meals in total in an average week? (meals per week) (number field)
(4) How many times do you prepare meals alone, on average, per week? (meals per week) (number field)
(5) How many times do you prepare meals together with others, on average, per week? (meals per week) (number field)

A.1.3 Affinity for Technology Interaction. Next, we asked participants to fill in the Affinity for Technology [11] Interaction questionnaire. (6-point Likert scale ranging from completely disagree to completely agree)

A.1.4 Kitchen utensil familiarity and perception when using the item. We explained: *In the following you will be presented a kitchen item and questions about how familiar you are with the item and about* *your perception when using this item.* Additionally, we stated: *Now, envision you are using this item. Then answer the questions below.* With this, we displayed the name of the object as well as the image of the object.

(1) I own this item.
   - Yes
   - No
(2) I am very familiar with this item. (101-point sliders from strongly disagree to strongly agree)
(3) I find the situation dangerous when I am using this item. (101-point sliders from strongly disagree to strongly agree)

A.1.5 Kitchen utensil perception when others use the item. Now, envision a person next to you is using this item. Then answer the questions below.

(1) I find the situation dangerous when somebody next to me is using this item. (101-point sliders from strongly disagree to strongly agree)

A.1.6 Kitchen Robot Questions. We displayed the image of Figure 2 by Oechsner et al. [21] and the text *Imagine yourself being in a kitchen environment where robot arms are available to support you with all kitchen tasks. These robot arms are able to grab and move every item you can find in a kitchen. Kitchen utensil perception when a robot uses the item. In the following, you will be presented the same 10 kitchen items with questions regarding your danger perception when such a robot next to you is using this item.* Then again we asked the following question for each object in the current survey:

(1) I find the situation dangerous when a robot next to me is using this item. (101-point sliders from strongly disagree to strongly agree)

A.1.7 General Feedback.

(1) This is the end of the survey. If you have any further feedback, this is the last spot where you can let us know. (text field)

A.2 Additional Insights

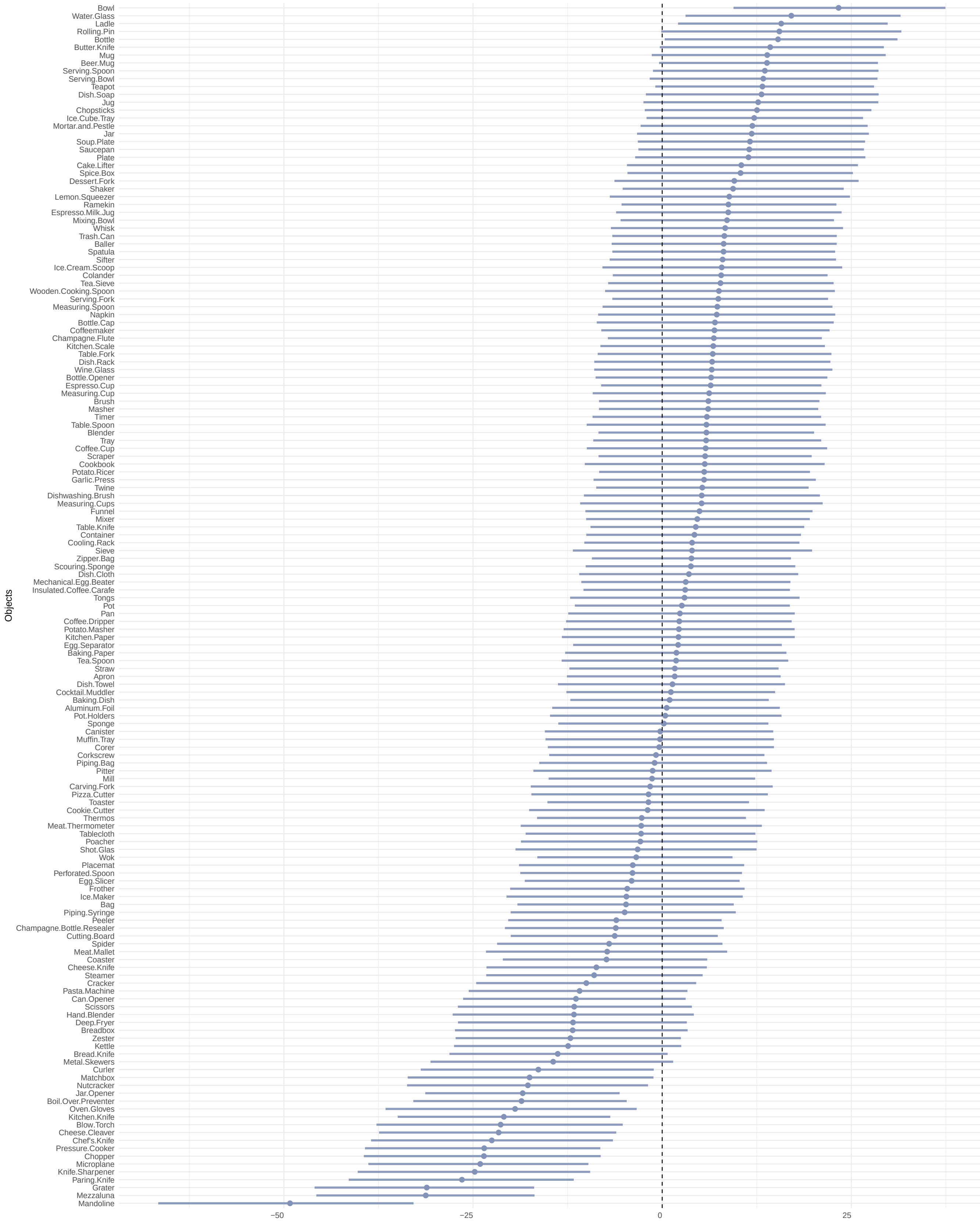


Figure 6: The difference between the perceived danger from ones own SCENARIO and the perceived danger from the robot SCENARIO. Objects with negative means are perceived as relatively more dangerous when used *self*, than by a *robot*.