Evaluating the Effectiveness of Iconography for Representing Robot Mental States in the Build-A-Bot Platform*

Benjamin Dossett[1], *Student Member, IEEE*, Weston Laity[1], Maisey Toczek[1], Robel Mamo[1], Jordan Sinclair[1], Nicole Train[2], Daniel Pittman[2], *Member, IEEE* and Kerstin S. Haring[1], *Member, IEEE*

***Abstract*— Robot designers and Human-Robot Interaction (HRI) practitioners can face challenges when people form a mental model of a robot that is not appropriate. Although the field of robotics would benefit significantly from a broad representation of designers, there is currently no comprehensive method of including many people in the design process and no theory of what expectations a robot design feature might elicit. We seek to address these challenges through the creation of a robot design platform, an online tool similar to a character creation interface in a video game, where users create a robot design. By collecting a large number of robot designs from users, we seek to be able to identify aspects of a robot's design that influence the mental models humans ascribe to the robot. To maximize the universal usability of the platform, we conducted a three-part survey to assess which icons should be used to visually represent the mental states ascribed to the robots created by users on the platform. In our assessment, we found nine icons that met our criteria for use in the platform and others that should be further evaluated.**

I. INTRODUCTION

Research has postulated that humans create a mental model of a robot during their initial interaction with it [1]. A mental model is a representation that allows humans to understand, reason about, and make predictions in a particular situation or experience. It allows humans to reason about and simulate the behavior of a system, even if they haven't directly experienced it [2].

The initial mental model of a robot is based on surface clues (i.e., the appearance and exterior of the robot), assuming that surface clues are the first available information about a robot [3]. These mental models often manifest through the attribution of mental states to robots in a similar way to how humans attribute mental states to other humans, pets, or even inanimate objects. After forming a mental model, people will then inadvertently form certain expectations towards a robot based on the model [4]. The expectations often concern the capabilities that a robot should (or should not) display in an interaction with a human. This means that robot designers and Human-Robot Interaction (HRI) practitioners can encounter challenges when people form a mental model of a robot that is not appropriate.

*This work was supported by the University of Denver's Professional Research Opportunities for Faculty (PROF) under grant # 142101-84994

[1]Benjamin Dossett, Weston Laity, Maisey Toczek, Robel Mamo, Jordan Sinclair, and Kerstin S. Haring are with the Ritchie School of Computer Science and Engineering, University of Denver, Denver, CO 80210, USA, kerstin.haring@du.edu

[2]Daniel Pittman and Nicole Train are with the Department of Computer Sciences, Metropolitan State University of Denver, Denver, CO 80204, USA dpittma8@msudenver.edu

Inappropriate expectations and misconceptions towards a robot are not the only challenges robot designers and developers encounter. Robots must often be designed with a form factor in mind. A form factor is a hardware design aspect that defines and prescribes the size, shape, and other physical specifications of robot components to facilitate functions such as movement, locomotion, and transport of the hardware necessary for the robot to function. Faced with these challenges, there might be little consideration during the design process of the mental model that a human may form of the robot. The lack of a comprehensive and predictive theory of what expectations a robot design feature elicits makes it very challenging for robot designers to consider the mental models possibly formed by a certain robot design. In many cases, this consideration is important to the designer's goals for the robot, as research on robot mental models shows that a robot's form can lead to misinformed and misguided assumptions about a robot's capabilities [3], [5]. Although frameworks have been developed to characterize social robots [6] and to build shared mental models in human-robot teaming [7] there is currently no comprehensive way to predict what mental model a robot design could generate.

A related issue with robot design is that the field of robotics, as with many other STEM fields, suffers from under-representation of people other than white and Asian men [8]. Many fields of technology, including robotics, benefit significantly from a broad representation of designers and developers [9], as well as a participatory design process [10]. We postulate that robotics and robot design will greatly benefit from a broad demographic of robot designers.

We seek to address these challenges through the creation of the Build-a-Bot platform, an online tool in which users can create a robot design using an interface similar to a character creation interface in a video game. As part of creating a design on the platform, users are asked to design their robot such that it could exhibit one of several different mental states, which are randomly assigned when a user begins a new design. By collecting a large number of designs from users, we seek to be able to identify aspects of a robot's design that influence human mental models of the robot. The platform does not restrict users to physically feasible designs in order to eliminate the bias of having to design around a form factor first. In addition, the platform is designed to be usable by as many different people as possible, including those who are not currently well represented in the field of robotic design.

As part of making the platform usable to as large a group of people as possible, the usability and intuitiveness of the user interface (UI) are particularly important [11]. In the early iterations of the platform prototypes, the robot design interface was largely text-based and in English. This led us to believe that such an interface excludes a significant portion of the potential user demographic that we seek to include in the robot design process. To increase the universal usability of the interface, we chose to utilize icons and significantly reduce the amount of text in the interface. Although there are widely accepted icons for certain actions (e.g., the floppy disk for "save", the house for navigating to "home"), as the goal of the platform is to create a large database of broad robot designs, we also needed a way to represent the mental states or emotions that should be reflected in a robot's design. Using icons and emoticons instead of large amounts of text can be a very expressive and powerful approach to communicating across a wide range of demographics [12], [13], [14], [15], [16], [17], [18]. Although there are some studies that look at robots using emoticons, there is no complete set of emoticons or icons that communicate the robot's experience or the robot's agency [19]. With the existing icons that we found in extensive web searches, it is currently unclear how those icons communicate the mental states that we wanted to include in our robot designs (e.g., the experience of being recognized or conscious). We decided that it was necessary to assess the icons that we plan on using to represent these mental states before we start collecting data on robot designs.

In this work, we present the design and results of a survey with the goal of providing insight into human perception and identification of unfamiliar icons in the context of social robotics and mind perception in robots. We provide a set of icons that we believe effectively represent mental states that humans may attribute to robots. We describe details of the icon selection process and discuss how the use of these icons as a part of a larger icon-based interface can promote universal usability and allow a more diverse set of users to participate in the robot design process.

II. Related Work

A. Icon-Based Interfaces

Increased icon-related research and developments in UI/UX have led to the development of icon-based interfaces that prioritize icon usage as a way to increase universal usability for illiterate (have not learned to read or write) or non-literate (people in cultures without written language) individuals and across languages [20]. Despite using almost exclusively icons, these interfaces allow for complex interactions and data collection. For example, in 1996, a team of American and South African professors and engineers developed an "Icon User Interface design for handheld computers that allowed non-literate [traditional animal] trackers to enter complex data" [21] into their device. The creators of this interface found that, in the hands of indigenous trackers, the interface could collect complex and rich data that could not be collected in any other way.

Businesses are also developing icon-based interfaces to open doors to people who may not be able to navigate the typical interfaces presented on smartphones, tablets, and desktops. For example, the Anou Cooperative [20], an artisan-owned cooperative based in Fes, Morocco, uses an icon-based interface that has been developed so that artisans can list their products on the online platform. Many artisans in rural parts of Morocco are non-literate, so the interface was designed with the goal of being accessible to all its users. Artisans can upload pictures, size information, and pricing with minimal text presented by the interface itself. In this example, icons are used to redesign a series of complex steps with the end user in mind.

In any case, icon-based interfaces are not an emerging design practice. HCI specialists have been creating pictographic and iconographic representations of data or processes within computer systems since the 1980s [22]. Icons such as a floppy disk representing the "save" action have been staples of UI design for decades [23] and remain integral elements in today's designs. Icons are critical to the development of accessible interfaces [24], and icon-based interfaces go beyond this in terms of universal usability for the general public [21]. We strive to make the Build-A-Bot platform as accessible as possible to all users. This is achieved by deliberately using icons and designing an interface for universal usability.

To create universal usability for interfaces, developers must be thoughtful and proactive about accessible design in today's digital world [24]. In the iterative design process of the interface, we noticed that the screen describing the robot design requirements to the user used only text (see Figure 2). We decided that it would increase universal usability to use icons to represent the mind-attribution target to the user. Additionally, an appropriate icon can be continuously displayed during the design process in the robot character creation interface to remind the user what to design for.

Character creation has been shown to be a powerful tool for participatory web-based robot design with adolescents [25]. The creation of virtual characters can also be a human-centered design approach to understanding robot perceptions [26]. This suggests that the implementation of widely available technologies, such as web-based platforms, may provide a better way to understand how people think about robot design.

III. Methodology

A. The Build-A-Bot Platform

The Build-A-Bot platform is hosted on a website (https://www.dubuildabot.com). Users are directed to the homepage, where they can register as new users or log in as existing users. The website also has navigation tabs that allow the user to learn more about the project, the team, and current publications. The website was implemented using the MEAN software stack (MongoDB [27], Express.js [28], Angular [29], and Node.js [30]).

After logging in, a user can directly navigate to the robot design platform by clicking on a button. The platform was implemented by embedding an application made with Unity, a 3D game engine [31]. The platform works similarly to a video game character creation interface, where the user can drag and drop robot parts onto the main screen and assemble them together. The robot builder menu allows for functionalities such as part manipulation (e.g. rotation, scaling, sizing, coloring), undo and redo, removal of parts, saving current progress, and submitting the final robot design (see Figure 1). A submission results in the robot design being added to a database.

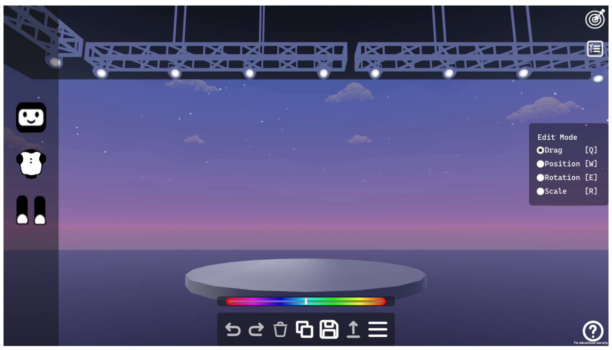


Fig. 1: The main interface of the Build-A-Bot platform, where users build their robot designs.

Before users begin designing a robot, they are presented with a "challenge card". The purpose of the card is to prompt the user to create a robot design that they believe could experience a particular human mental state. A user would be challenged to create a robot that can, for example, experience fear or joy along the experience dimension, or create a robot that can plan or memorize along the agency dimension. These elements, among others, were inspired by dimensions of mind perception [19].

However, the current challenge card contains only text (see Figure 2). This poses two problems for our design: first, we want to use less text to increase universal usability across ages and different languages; and second, we want to keep the challenge card visible throughout the design process so that the user is reminded of what the target design of the robot is. With the current implementation, we cannot achieve either.

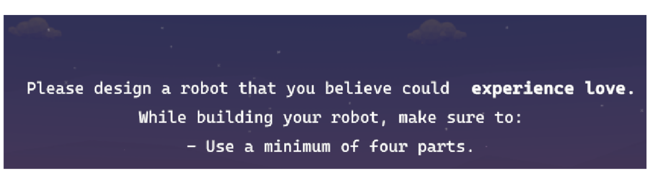


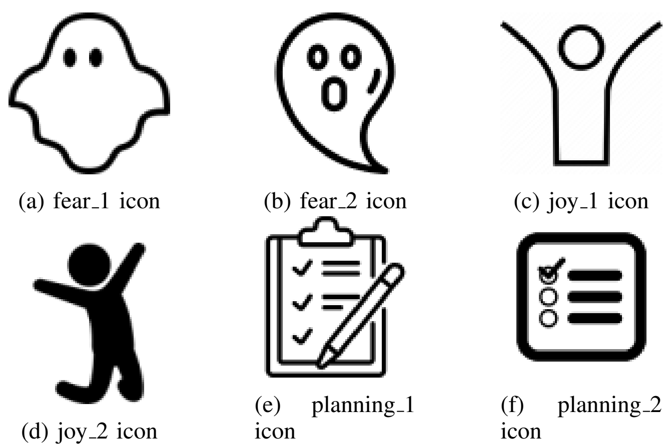
Fig. 2: An example of an existing challenge card, which relies entirely on text.

B. Experimental Design

To evaluate which icons should be used to represent human mental states for the challenge cards in our icon-based interface, we conducted a three-part survey. The survey was designed with three main goals. First, it was necessary to collect data on what users believe a given icon represents. As interface designers, we had a goal in mind for these icons, but this goal may not align with what users think. Second, we needed to verify (or disprove) that the users agreed with the intended descriptions of the icons that we selected. Third, as we have more than one icon option for the majority of human mental states being represented, we needed to have users select a preference between the icons. To evaluate whether our icon selection met these three goals, we designed our survey in three sections, as described in Section III-D.

C. Icon Selection

Our search for icons was based on a list of mind perception dimensions from [19]. Two researchers conducted an online search using these mind perception dimensions as keywords. As there is no existing set of icons to represent these dimensions, the researchers selected two or more icons that appeared to be initially appropriate for representing each dimension. For an icon to be selected, both researchers had to agree that the icon seemed appropriate. The icons were selected from several websites [32], [33], [34], [35], [36], [37], [38], [39], [40] with considerations made to keep a consistent theme for use in our interface. If necessary, these icons were edited to maintain a consistent coloring for cohesive UI design.



(a) fear_1 icon (b) fear_2 icon (c) joy_1 icon

(d) joy_2 icon (e) planning_1 icon (f) planning_2 icon

Fig. 3: A selection of six of the total 40 icons evaluated in our survey. In this selection, there are two icons shown for each of the three mind perception attributes that are anticipated to be presented to a user, fear, joy, and the ability to plan.

D. Survey Design and Evaluation

50 participants were recruited through Amazon Mechanical Turk (MTurk) [41] for the survey. The participants received monetary compensation for their time. The study was approved by the Institutional Review Board (IRB) of the University of Denver.

1) Free Association: The first part of the survey was a free association task. Participants were asked to use between one and five words to describe each icon. Furthermore, participants were asked to rate the strength of their description from zero to ten using a slider. A rating closer to zero means that they felt that their description did not fit the icon well, and a rating closer to ten means that they felt that their description fit the icon very well. 40 individual icons were evaluated. Each participant described all 40 icons. These data were used to see if users agreed on a general theme to describe the icon and if user descriptions matched our assumptions. To evaluate the free response portion, the words provided by the participants were cleaned and stemmed using the `snowballstemmer` Python library and then compiled into a word cloud for each icon. The word cloud allowed us to see whether the most common words associated with an icon matched our intended descriptions. A word cloud that did not use the desired language would indicate that the icon did not match well with our association with that icon. The word clouds also made it possible to compare competing icons, as we could select a corresponding word cloud that we felt best described our intentions. We used the participants' ratings of the strength of their descriptions as an indicator of their confidence, with a higher rating indicating higher confidence.

2) Association Strength: For part two of the survey, participants received an icon and a description of the icon created by us. The participants were then asked to rank the description we provided on a scale of zero to ten, with zero indicating that the provided description does not match the icon at all and ten indicating that the provided description matches the icon perfectly. In this task, the same 40 icons as in the first part of the survey were evaluated. We evaluated the mean rating among all participants for each icon-description pair and also compared the mean rating with the different icons with the same description.

3) Pairwise Comparison: In part three of the survey, participants were presented with two icons and one description. Both icons were meant to represent the same human mental state, and participants were asked which of the two icons, if either, was a better match for the description. For example, two icons representing "recognition" were presented to the user. Participants were asked to move the slider towards the icon they believed to be a better representation of the given description. In the cases where three icons were evaluated for the description, we used three pairwise comparisons that compared each of the three icons with the other two. In total, 24 comparisons were made. Each participant made all 24 comparisons.

For each pair of icons, we evaluated the mean value of the sliders in the question among all participants. A lower mean slider value would indicate that the participants generally preferred the icon displayed on the left side of the slider. A higher mean slider value would mean that participants preferred the icon on the right. A mean slider value close to $M = 5$ (the center) would indicate that participants had little or no preference for one icon over the other. When evaluating these mean slider values, we considered the *winning margin*. The *winning margin* is the difference between the mean slider value and 5 and facilitates the visual representation of how much one icon is preferred over another. We decided that for an icon to have *won* over another, there should be a clear winning margin of at least a 1.5 rating point difference, which we believe would indicate a much better association with the description.

In the cases where we had only two icons for a given attribute, the *winning margin* gave us a clear idea of which icon we should use in our interface. In the cases where there were three icons that targeted the same attribute, we used a pairwise comparison to evaluate which of the three should be selected. A lower *winning margin* would indicate an inconclusive finding and that the icon design needs to be reevaluated.

E. Final Icon Selection

To decide on the final set of icons that were suitable for use on our platform, we first selected icons that had a word cloud from part 1 of the survey that matched our intended association. We then verified that these icons had a high association rating in part 2 of the survey. Finally, we eliminated any icons that lost to any other icons in part 3 of the survey. If all three of these criteria were met, the icon was selected for use.

IV. RESULTS

A. Participants and Demographics

50 participants were recruited. Two of the participants were excluded from the data evaluation due to a technical error that resulted in their survey responses being incomplete. 19 participants identified as women, 29 identified as men, and none identified in the other options, including a "Prefer not to say" option. The average age was $M = 41.2$ ($SD = 9.2$, ranging from 27 to 66). The Institutional Review Board of the University of Denver approved this study as exempt, and all participants agreed to participate.

B. Free Association

The word clouds generated for each of the icons showed that some icons matched the attribute we wanted to express much better than others. Of the 40 icons for which we generated word clouds, 18 icons matched our expected meaning to a sufficient extent, meaning that either our intended association was one of the prominent words in the cloud, or the prominent words in the cloud evoked the same meaning as our intended association. The other 22 icons' word clouds did not match our expected meaning. Figure 4 shows an example of a word cloud that matched the expected meaning, and Figure 5 shows an example of a word cloud that did not match. In total, the presented icons targeted 20 different mental states and we identified ten distinct mental states that had one or more icons whose word clouds matched our intention. In this analysis, there was an overlap where 18 icons were found to be a sufficient match for 10 distinct mental states.

When graphing the participants' self-reported ratings of the strength of their descriptions, we found that participants generally rated their descriptions highly, meaning that they believed their descriptions were accurate. Fig. 6 shows the mean and standard deviation of the ratings the participants gave for the memory icons, including both those in Fig. 4 and Fig. 5.

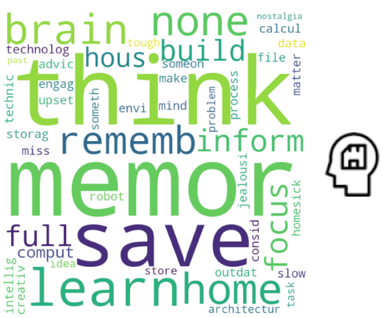


Fig. 4: An example of a word cloud for the "memory_3" icon using word stems that did match our intended association of "memory" for the icon.

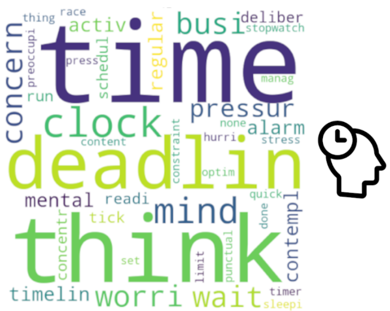


Fig. 5: An example of a word cloud for the "memory_2" icon using word stems that did *not* match our intended association of "memory" for the icon.

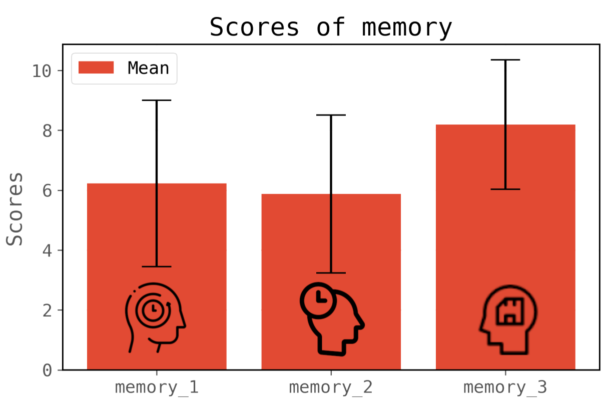


Fig. 6: Plotted mean confidence values for the three "memory" icons evaluated.

C. Association Strength

Participants evaluated 40 icons to see how well they matched a given description. The mean rating value was calculated for each icon. The histogram in Figure 7 shows the count of the rating given, and the histogram is biased to the right. In general, participants rated the given descriptions highly on the given scale from zero to ten for many of the icons we evaluated. Of the 40 icons, 20 had a mean value greater than 8.0, which we decided was a good indication that these icons may be usable in our interface. This decision was made based on the grouping of the icons in the histogram created (see Figure 7). There appeared to be somewhat of a distinct cutoff between a group of icons above 8.0 and a group below, leading to the choice of an 8.0 confidence value from the 1 to 10 scale as a cutoff.

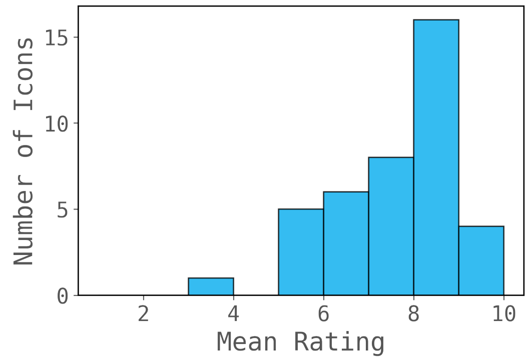


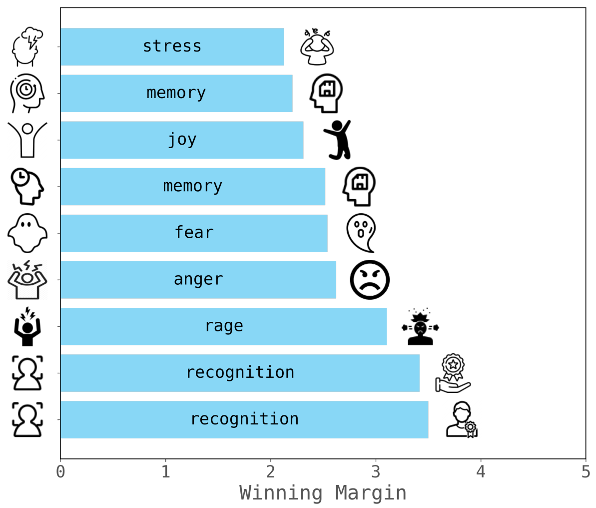
Fig. 7: Histogram showing the distribution of mean ratings for the evaluated icons.

D. Pairwise Comparison

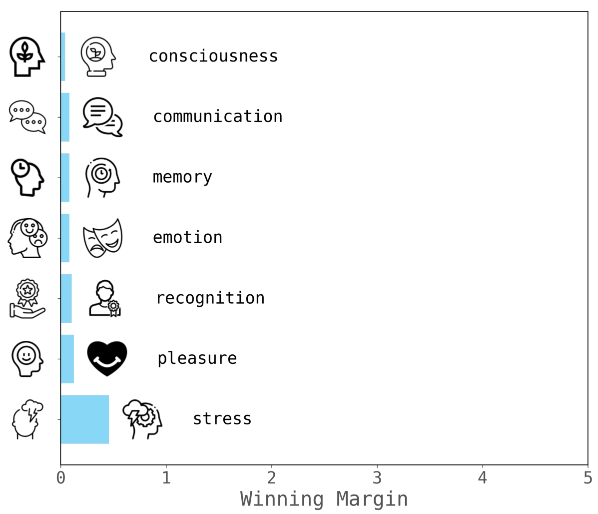
Of the 24 comparisons performed, 12 resulted in a *winning margin* greater than 1.5. These 12 comparisons contained nine distinct icons that had a clear win (> 2.0). Fig. 8a shows the 9 comparisons with the highest *winning margin*. The remaining 12 comparisons resulted in a *winning margin* of less than 1.5. Of those 12, six had a *winning margin* of 0.125 or less. Fig. 8b shows the seven comparisons with the lowest *winning margin*. In the categories with three icons that were compared pairwise, we were able to identify a clear winner among the three. Fig. 9a shows the winning margins for the three comparisons made for the three "memory" icons that were evaluated. The third "memory" icon had a significant win in both comparisons with the other icons, while the comparison of the first and second icons showed no significant preference. This would indicate that the third "memory" icon was preferred given the attribute "memory", and thus would indicate that we should use this icon. We also found cases where two different icons had a significant win over a third targeting the same attribute. For example, three "recognition" icons were evaluated. Fig. 9b shows the winning margins for the three icons. Two distinct icons were rated as having a significant win over the third, which eliminated the third icon from consideration. However, there is no significant preference between the two distinct winning icons, meaning that there is no clear indication of which of the two to use.

E. Icon Selection

To select icons to use on our platform, we used the method described above.

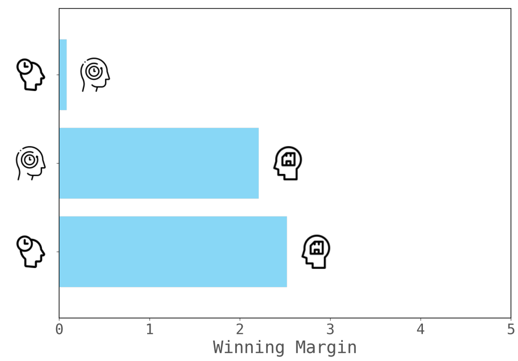


(a) The 9 comparisons with the largest winning margins.

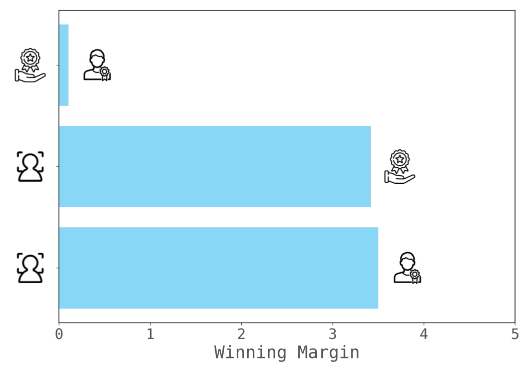


(b) The 7 comparisons with the smallest winning margins.

Fig. 8: Icons were evaluated using comparisons between icons representing the same thing. Comparisons resulted in a “winning margin”, displayed here.



(a) Memory icon winning margins.



(b) Recognition icon winning margins.

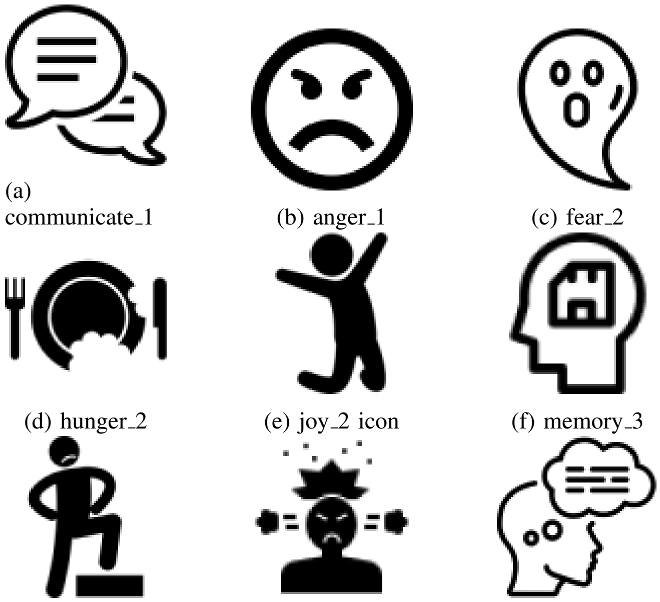
Fig. 9: Different win cases in the pairwise comparison of three icons.

We started with all of the 18 icons that had a word cloud that matched our intended association from part 1 of the survey. We then removed any icons from that list that had an association rating in part 2 of the survey of less than 8.0, which we considered to be the minimum for a strong association as discussed above. This resulted in 15 icons to be considered. We then removed any icons that had one or more losses to another icon in the pairwise comparison section of the survey. This resulted in nine icons remaining: communication_1, anger_1, fear_2, hunger_2, joy_2, memory_3, pride_1, rage_2, and thought_2. These icons were selected for use in our platform, and can be seen in Fig. 10.

V. Discussion

When evaluating the word clouds created from the free associations in the first part of the survey, we used a subjective evaluation of the words in the cloud to decide whether that icon matched our intended description. For example, if an icon produced a word cloud that did not exactly match the word we use for a given mental state, but used similar terminology expressing the same thing, we said that it matched our intention. Although this gives us a metric to work with, there is still a concern that others may interpret the descriptions the participants gave differently than we do, leading to a disconnect between us and the users. Therefore, we needed to interpret these findings together with the other icon assessments. Participants also generally rated their own descriptions highly for how strongly they believed that the descriptions matched the icon. However, some of the icons for which participants rated their descriptions highly were ones that had word clouds that did not match our intentions. For those cases, this would indicate that the icon should not be used, as it generates a confident interpretation that does not match our intention.

The results of the association strength task, in which participants rated the strength of the association between a given icon and a given description, showed a trend toward higher ratings. This means that the participants found that half



(a) communicate_1 (b) anger_1 (c) fear_2
(d) hunger_2 (e) joy_2 icon (f) memory_3
(g) pride_1 (h) rage_2 (i) thought_2

Fig. 10: The nine final icons selected for use in our platform.

of the icons presented matched our intended associations. The 20 icons fell into the group that was rated particularly high (> 8.0) are considered good choices for our interface. Icons that were rated above the rating mean but not high (between $5.0 < M < 8.0$) are not rejected at this point but will require further research to determine if they can still be used with minor modifications for the robot design interface. It is important to note that the 8.0 and above criteria is somewhat subjective as explained in Section IV-C. As this work is exploratory, we did not have a definitive cutoff that could be used to determine whether an icon is strongly associated with a mental state or not. The cutoff selected appears to be appropriate given the exploratory data analysis performed. In future work, we hope to determine common characteristics between icons that have a causal relationship with the ratings given, which could be used to create a more objective measurement of the strength of an association. Selecting or creating an icon that will produce the desired association for every single user at all times might not be possible; however, selecting icons that produce on average the desired association for a majority of users is crucial. This interpretation is supported by the word clouds.

It seems to be the case that a user sometimes associates an icon in a way that is not strictly the same as our intention but conveys the same general idea. This could also result in participants rating some icons lower for the provided descriptions as they would lean towards a different, yet similar description. Therefore, icons in the middle area of mean ratings require further assessments and comparison with alternatives before making a final decision on their use on our platform.

The pairwise comparison between different icons targeting the same association produced interesting complementary results. We were able to identify icons with significant wins over other icons given a description, however, the icons did not match our intended associations as shown in the free association part. Additionally, a small winning margin does not necessarily mean that either icon is a bad choice, only that the two icons were similar in terms of their association. This meant that a nuanced approach was required to select our final set of usable icons.

Our two different methods of determining which icons to use in our platform produced results that were very consistent with each other. This was a good indication that the icons selected using the two methods would produce the intended association in users. Icons that did not meet the criteria we set could potentially still be useful to us but will require further evaluation to see if there are better alternatives that we could use instead.

A. Limitations

One potential limitation of our survey design is that the sliders used to conduct pairwise comparisons had numbers on them from zero to ten. This could potentially lead participants to choose a higher number by default and give less consideration to the actual comparison at hand. From our results, we do not believe that this occurred since the icons on both sides of the slider were chosen by participants. However, in future iterations, we will eliminate the increasing numbers on these sliders and choose a balanced representation to avoid this limitation.

VI. Conclusion

Through our icon assessment, we were able to select a set of nine icons that we believe will be strongly associated with their intended (human) mental state and will be used in the redesigned challenge card shown to a user before starting a new robot design on the platform. By using these nine icons and applying the other insights from this evaluation, we will be able to utilize the icons on the challenge card effectively and we are increasing the universal usability of the robot design platform. This is expected to make this platform usable for a broader audience and result in a more representative robot design database.

References

[1] S. lai Lee, I. Y. man Lau, S. Kiesler, and C.-Y. Chiu, "Human mental models of humanoid robots," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 2767–2772, 2005.

[2] D. Gentner, "Mental models, psychology of," 2001.

[3] K. S. Haring, K. Watanabe, M. Velonaki, C. C. Tossell, and V. Finomore, "FFAB–The form function attribution bias in human–robot interaction," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 4, pp. 843–851, 2018.

[4] V. Lim, M. Rooksby, and E. S. Cross, "Social robots on a global stage: establishing a role for culture during human–robot interaction," *International Journal of Social Robotics*, vol. 13, no. 6, pp. 1307–1333, 2021.

[5] E. Phillips, S. Ososky, J. Grove, and F. Jentsch, "From tools to teammates: Toward the development of appropriate mental models for intelligent robots," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55, p. 1491–1495, Sep 2011.

[6] K. Baraka, P. Alves-Oliveira, and T. Ribeiro, *An Extended Framework for Characterizing Social Robots*, vol. 12, p. 21–64. Cham: Springer International Publishing, 2020.

[7] M. Scheutz, S. A. DeLoach, and J. A. Adams, "A framework for developing and using shared mental models in human-agent teams," vol. 11, p. 203–224, Sep 2017.

[8] NSF, "s&e indicators 2018 — nsf - national science foundation," 2018.

[9] C. Fine, V. Sojo, and H. Lawford-Smith, "Why does workplace gender diversity matter?justice, organizational benefits, and policy," *Social Issues and Policy Review*, vol. 14, pp. 36–72, 2020.

[10] A. Darriba Frederiks, J. R. Octavia, C. Vandevelde, and J. Saldien, "Towards participatory design of social robots," in *Human-Computer Interaction – INTERACT 2019* (D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris, eds.), (Cham), pp. 527–535, Springer International Publishing, 2019.

[11] B. Shneiderman, "Universal usability," *Communications of the ACM*, vol. 43, no. 5, pp. 84–91, 2000.

[12] K.-A. Hsiao and P.-L. Hsieh, "Age difference in recognition of emoticons," in *Human Interface and the Management of Information. Information and Knowledge in Applications and Services: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings, Part II 16*, pp. 394–403, Springer, 2014.

[13] V. A. Pfeifer, E. L. Armstrong, and V. T. Lai, "Do all facial emojis communicate emotion? the impact of facial emojis on perceived sender emotion and text processing," *Computers in Human Behavior*, vol. 126, p. 107016, 2022.

[14] W. Brants, B. Sharif, and A. Serebrenik, "Assessing the meaning of emojis for emotional awareness-a pilot study," in *Companion proceedings of the 2019 world wide web conference*, pp. 419–423, 2019.

[15] D. De Angeli, R. M. Kelly, and E. O'Neill, "Beyond happy-or-not: Using emoji to capture visitors' emotional experience," *Curator: The Museum Journal*, vol. 63, no. 2, pp. 167–191, 2020.

[16] P. Q. Brito, S. Torres, and J. Fernandes, "What kind of emotions do emoticons communicate?," *Asia Pacific Journal of Marketing and Logistics*, vol. 32, no. 7, pp. 1495–1517, 2020.

[17] I. Boutet, M. LeBlanc, J. A. Chamberland, and C. A. Collin, "Emojis influence emotional communication, social attributions, and information processing," *Computers in Human Behavior*, vol. 119, p. 106722, 2021.

[18] S. Bozinovski and A. Bozinovski, "Mental states, eeg manifestations, and mentally emulated digital circuits for brain-robot interaction," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 1, pp. 39–51, 2015.

[19] H. M. Gray, K. Gray, and D. M. Wegner, "Dimensions of mind perception," *Science*, vol. 315, no. 5812, pp. 619–619, 2007.

[20] T. Anou, "The anou - about," 2022.

[21] L. Liebenberg, J. Steventon, N. Brahman, K. Benadie, J. Minye, H. K. Langwane, and Q. U. Xhukwe, "Smartphone icon user interface design for non-literate trackers and its implications for an inclusive citizen science," *Biological Conservation*, vol. 208, pp. 155–162, 2017.

[22] I. J. of Man-Machine Studies, "Icon-based human-computer interaction," *Applied Ergonomics*, vol. 24, pp. 519–543, 1986.

[23] IBM100, "The floppy disk," 2022.

[24] R. M. Gilbert, *Inclusive Design for a Digital World*. Apress, 2019.

[25] S. Bhatia, E. A. Björling, and T. Budhiraja, "Exploring web-based vr for participatory robot design," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '21 Companion, (New York, NY, USA), p. 109–112, Association for Computing Machinery, 2021.

[26] A. S. Kim, E. A. Björling, S. Bhatia, and D. Li, "Designing a collaborative virtual reality game for teen-robot interactions," in *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, pp. 470–475, 2019.

[27] MongoDB, "MongoDB: The Developer Data Platform," 2022.

[28] Express.js, "Express.js." https://expressjs.com/, 2022.

[29] Angular, "Angular," 2022.

[30] Node.js, "Node.js." https://nodejs.org/en/, 2022.

[31] U. Technologies, "Unity Real-Time Development Platform | 3D, 2D VR & AR Engine," 2022.

[32] IconFinder, "Icon finder," 2022.

[33] VeryIcon, "Very icon," 2022.

[34] NicePNG, "Nice png," 2022.

[35] FlatIcons, "Flat icons," 2022.

[36] ShareIcon, "Share icon," 2022.

[37] TheNounProject, "The noun project," 2022.

[38] SmartArtSchool, "Smart art school," 2022.

[39] StockIO, "Stock io," 2022.

[40] ClipIconMax, "Clip art max," 2022.

[41] Amazon, "Amazon mechanical turk," 2023.