Proceedings of the ASME 2023
International Design Engineering Technical Conferences and
Computers and Information Engineering Conference
IDETC/CIE2023
August 20-23, 2023, Boston, Massachusetts

DETC2023-116943

EMPATHIC ABILITY VS. USER SATISFACTION: EVALUATING THE ROLE OF PREFERENCE ACCURACY IN ENGINEERING DESIGN

Erin Johnson[1,*], Cynthia Letting[2], Nicolás F. Soria Zurita[1], Jessica Menold[1,2]

[1]School of Engineering Design and Innovation, The Pennsylvania State University, University Park, PA
[2]Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA
Email: evj5226@psu.edu, cjl5836@psu.edu, nfs5403@psu.edu,jdm5407@psu.edu

ABSTRACT

Prior work has identified the importance of empathy and the effect on design outcomes such as increased quality, originality, and usability. The measurement of empathy in the context of design, however, remains a pervasive issue. While several studies have proposed interventions, tools or methods to enhance designers' empathy, it remains unclear how managers or designers can practically measure empathic abilities. Many of the existing methods are either incredibly time-consuming and thus not feasible in design practice or are not grounded in design work, making the translation and interpretation of instruments or surveys challenging. To address this gap, we study preference accuracy as a measure of empathic ability, and investigate its relationship with design outcomes, specifically user satisfaction. We review a case study implementing this method; fourteen participants were recruited and randomly paired with a partner. Individuals were tasked with designing a chair to meet their partner's needs. Each participant interviewed their partner and completed two preference assessments: one detailing their own design preferences, and one detailing their perceptions of their partner's design preferences. To calculate preference accuracy, the preference assessments for each participant were compared. Results from this study suggest preference accuracy may not be linked with user satisfaction, but may be a useful tool to differentiate empathic abilities across a population of designers.

Keywords: empathy, engineering design, user satisfaction

1. INTRODUCTION

Empathy is interwoven in the design thinking process to solve complex problems [1] and leverage the human experience to bolster innovation. Empathy is often linked with emotional intelligence and defined as a social emotional skill that helps humans accurately identify and appropriately react to the experiences, emotions, or perspectives of others [2]. Within design, empathy is critical as it allows designers to more completely understand users' experiences and design solutions to more effectively meet users' needs [3].

The measurement of empathy can be challenging, and prior work has relied on both self-report surveys [4–8] and physiological measurements [9–11]. The use of self-report surveys measure the level of empathy an individual possesses. The Interpersonal Reactivity Index (IRI) [12], a commonly used survey, compares the level of empathy across time and across differing populations. This survey is used to measure general levels of empathy and recent work suggests it may not be an appropriate measure of empathy within the context of design problems [4]. Physiological measurements, such as facial electromyography and magnetic resonance imaging have been used to measure empathic response. Facial emotion recognition and facial mimicry have been studied in design to establish a link between empathic understanding and design outcomes [9, 10]. Measuring emapthic ability and leveraging the synchrony of physiological signals is incredibly nascent, and additional studies are needed to validate this method. Further, the extensive experimental setup needed to facilitate these measures, often makes them infeasible for real-world designers to implement in practice.

Recently, researchers have explored the creation and use of new techniques to measure empathic ability, specifically during design processes, such as the empathic experience design method and empathic accuracy method [9, 13, 14]. The empathic accuracy method [9] is comprised of an interview between the designer and the user. Both the user and designer review a recording of the interview and highlight moments of emotional importance to describe the user's mental contents. The similarity between user and designer is a measure of the designer's empathic ability. These current empathic measures can be time-consuming and labor-intensive to employ and study. The translation of design methods from research to practice has historically been a chal-

*Corresponding author: evj5226@psu.edu

lenge for our field, and we argue that a less time-consuming and more practical measure of empathic abilities is needed to help design practitioners assess empathic accuracy without disrupting or derailing day-to-day design practice. A more efficient measure of empathic abilities could enable design teams to more effectively leverage empathic interventions and tools in practice.

The aim of this paper is to propose a proxy measure of empathy that leverages the preference accuracy of designers. This method is tested via a case study and the findings of this case study are reviewed, with implications for the use of preference accuracy as a measure of empathic abilities more broadly. The rest of this paper is outlined as follows. Section 2 presents the background knowledge on which the study is based upon. Section 3 presents the research objectives, followed by Section 4 which presents the methods of the study. Section 5 presents the results of the study. A discussion of these results is found in Section 6. Limitations of this work and areas for future studies are presented in Section 7. Finally, the conclusion is presented in Section 8.

2. BACKGROUND

Design thinking and human-centered design practices have been broadly adopted across industries due to the belief that these methods, grounded in empathy, are key drivers of innovation and creativity [1]. The Design Value Index shows that over a ten-year period, companies that invest in design thinking across the organization see a return of 211% compared to companies that rely on traditional practices [15]. Designers' empathic ability is therefore critical to the broader success of the organization. To improve designers' empathic abilities, empathy must first be operationalized and measured.

2.1 Measuring Empathy in Design

Prior work has utilized tools such as self-report surveys, physiological measures and empathic behaviors to quantify empathic abililty. Literature has identified trait empathy as a multidimensional construct that encompasses both cognitive and affective emotional aspects [16–18]. Cognitive empathy refers to the intellectual capacity of an observer to imagine the psychological viewpoint of another and does not necessarily imply an emotional reaction from the observer [19]. In comparison, affective empathy refers specifically to the experiencing of some emotional reaction as a result of observing another [19].

2.1.1 Self-report Surveys. A commonly used self-report survey in design research is the Interpersonal Reactivity Index (IRI). The IRI [12] operationalizes trait empathy as a combination of four constructs: Perspective Taking (cognitive empathy and the tendency to adopt the psychological point of view of others); Fantasy (tendency to transpose oneself imaginatively into the feelings and actions of fictitious characters in books, movies, and plays); Empathic Concern (affective empathy and the feelings of sympathy and concern for others); and Personal Distress ("self-oriented" feelings of personal anxiety and unease in tense interpersonal settings) [12]. These constructs comprise the four subscales that leverage seven Likert-type item statements including "*I try to look at everybody's side of a disagreement before I make a decision*". The empathy score is the summation of responses.

The IRI has been used to compare levels of empathy across time and across populations [4, 6]. In the work conducted by Surma-aho et al. [4], the Perspective Taking (PT) subscale of the IRI was used to study the evolution of perspective taking tendencies over the course of a design thinking training. The findings indicate that perspective taking tendencies increase over time. Similarly, Rasoal, Denielsson, & Jungert [6] utilized the IRI survey to compare empathic abilities across unique populations of students. In their study, the entirety of the IRI was used to compare the level of empathy between students in engineering programs and students in healthcare programs. Their results show engineering students had lower scores across the fantasy and perspective-taking subscales, as compared to healthcare students. Within design research the IRI has been frequently used to assess the empathic abilities of designers. For example, Alzayed et al. [20] investigated the effects empathic composition has in teams. Specifically, the researchers studied the elevation (average) and the diversity (standard deviation) of empathy across design teams. Results found that increased elevation of team empathy positively affected the uniqueness of ideas generated and selected by the team. However, the elegance and usefulness of these ideas were negatively impacted by higher levels of elevation in team empathy. Increased levels of diversity in team empathy was found to positively impact the usefulness of generated ideas. Interestingly, increased levels of diversity in team empathy were associated with less elegant and unique ideas generated and selected by design teams [20].

A commonality between these studies is the limitation of the context of the IRI. The IRI was developed for psychology and intended to measure "everyday" empathy. The scale is too general to accommodate the context specific nature of design. For example, the statement "*I sometimes find it difficult to see things from the 'other guy's' point of view*" [12] is difficult to answer. The "other guy" could be a teammate, end user, or key stakeholder. The generality of these statements produces difficulty in internal consistency and should be paired with other methods when used in the context of engineering design. Hess et al. [21] incorporates design context in their creation of a self-report survey. In this work, the authors iterated on their prior work [22] to create a survey to measure cognitive and affective empathy across three design phases. Cognitive and affective empathy constructs were found to be reliable, though several subconstructs were found to be unreliable. The authors note that iterations of their survey, as well as existing surveys, may be necessary to ensure validity in engineering design contexts.

2.1.2 Physiological Measures. Physiological signals can also be used to measure empathy. Neumann et al. [11] reviewed existing literature and highlight facial electromyography (fEMG), electroencephalogram (EEG), and magnetic resonance imaging (MRI) as common tools to measure empathy. The MRI uses a magnetic field to produce an image of internal brain structures. The volume of activity in these areas during empathic experiences is theorized to measure the level of cognitive and affective empathy. Facial electromyography detects facial muscle activity and provide a threshold of motor mimicry. Increased facial mimicry is believed to indicate empathic responses. Electroencephalograms measure brain waves that reflect psychological processes, including the empathic response [11]. There have been recent studies that use these physiological measures in an effort to quantify empathy in the design process. For example, Chang-Arana et al. [9] measured facial mimicry and correlated the results with designer empathic accuracy scores. The study leveraged video interviews to collect EMG data from the zygomaticus major (lip pulling) muscles as well as empathic accuracy scores. The findings of this study suggest the synchronization of the zygomaticus major does not predict empathic accuracy.

Salmi et al. [10] supported these findings in their exploration of facial expression recognition to measure mimicry as a function of empathic understanding. In the study, participants conducted interviews with users over a video platform. To measure empathic understanding, the participants completed an empathic accuracy task, following the same procedure as [9, 23, 24]. Software was used to read the action units of the Zygmoaticus Major (lip corner puller), Corrugator Supercilii (brow lowerer), and Orbicularis Oculi (cheek raiser) muscles. These readings were used to measure the level of mimicry between the designer and user. Their results found that mimicry does not automatically mean the designer understands the mental contents of the user, meaning cognitive and affective empathy may not be directly linked to user understanding. Though these works have not seen clear links between facial mimicry and empathic understanding, Chang-Arana et al. highlight that synchrony in other physiological signals could result in strong links to empathic understanding [9]. This supposition may not be the most viable option for use in engineering design and design research. Measuring empathy through physiological means is often limited to small samples of participants due to the equipment necessary to scan or collect data. Additionally, these measuring techniques are often time consuming in terms of test protocols and data processing. While physiological measures may quantify empathic understanding and ability, these techniques may not be feasible for design practitioners. Further, we highlight that emerging research findings suggest that these measures may not consistently or accurately measure empathic abilities of designers [10].

2.1.3 Empathic Behaviors. The Empathic Accuracy Method proposed by Chang-Arana et al. [9], begins with a video-recorded contextual interview with users. Users are then asked to review the video and log the thoughts and feelings they experienced during key moments of the video. Next, the designer watches the video and infers the user's thoughts and feelings, pausing the video to describe moments of perceived emotion. An empathic accuracy score is calculated from the similarities between the logged mental contents and the perceived emotions of the user and designer. One downside to this process is that it requires a significant time commitment from both the user, designer, and researchers. Although the researchers proposed a shorter process, the method still requires significant time investment on the part of the user, designer, and researchers. The abbreviated method [14] begins with a video-recorded contextual interview and user reflection of the recording as prescribed by the original procedure. The abbreviated method then deviates from the original procedure by filtering out ten entries through random selection. The designer then infers the user's thoughts and feelings from the selected entries. Filtering out user entries reduces the amount of time for the designer to infer perceived emotions, however a considerable amount of time is still required for the interviews, user review and reflection, and analysis of designer inferences.

We highlight that overwhelmingly, many of these methods are not suitable to design practice, and the assessment of designer's empathic abilities remains a pervasive issue. Further, the linkages between empathic abilities and design outcomes remain split, and it is unclear what the effects of designers' empathic abilities are on key design outcomes. A faster, easier method for measuring empathic ability is needed for design practitioners. Without such a method it becomes impossible for designers, managers, or firms to estimate and improve upon the empathic abilities of designers through targeted interventions, tools, or methods.

3. RESEARCH OBJECTIVE

We propose a new method of leveraging preference accuracy to measure the empathic abilities of designers. This novel method was tested in a case study with designers. We hypothesize that increased preference accuracy will be linked with increased user satisfaction with final designs. Empathy is critical to the accurate identification of user needs and preferences [25]. This hypothesis is based on the tenets of the Kano Model of Customer Satisfaction [26]. This model is comprised of three distinct requirements which influence customer satisfaction: must-be requirements fulfill basic criteria, one-dimensional requirements fulfill functional features, and attractive requirements exceed customer expectations. Matzler and Hinterhuber [27] highlight that attractive requirements are neither explicitly expressed nor expected by the customer, thus, failing to meet the requirement does not necessarily spur feelings of dissatisfaction. We define preference accuracy as the similarity of feature identification and rank order to address the must-be and one-dimensional requirements in Kano's model. In order to understand the viability of using preference accuracy to assess designers' empathic abilities, we test the following hypothesis via case study:
Increased preference accuracy is a predictor of increased user satisfaction.

4. METHODS

To determine the relationship between preference accuracy and user satisfaction, fourteen startup company founders were recruited to participate in a controlled study. This section presents the methodological approaches of our study.

4.1 Participants

Fourteen startup company founders that participated in the National Science Foundation's (NSF) Innovation Corps (I-Corps) program were recruited for this study. NSF introduced the I-Corps program in 2011 to prepare scientists and engineers to extend their focus beyond the laboratory to increase the economic and societal impact of NSF-funded and other basic research projects [28]. Participants were recruited from the Mid-Atlantic I-Corps Hub. Researchers collaborated with individual program directors for recruitment via email, in accordance with the Pennsylvania State University's Institutional Review Board Policies. Participants in the study ranged in age from 20 to 65 years old, see Figure

1. Race, ethnicity, and gender demographics for the sample are shown in Figure 2. The participants specialize in deep tech and/or sustainable technology solutions.

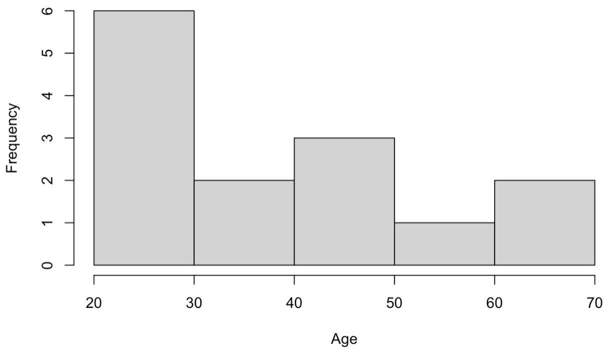

FIGURE 1: DISTRIBUTION OF PARTICIPANTS' AGE

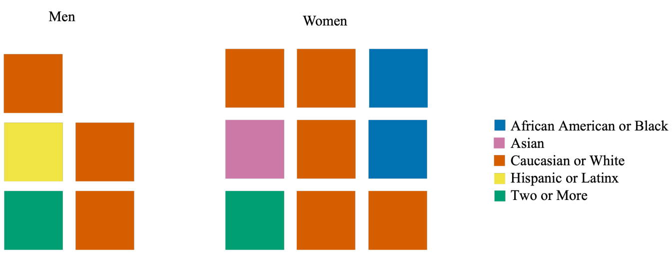

FIGURE 2: DISTRIBUTION OF PARTICIPANTS' RACE/ETHNICITY AND GENDER

4.2 Instruments

A preference assessment sheet was created for this study. A list of preferences were developed by the first, second, and fourth authors based on extensive research of seating solutions. This preference list was pilot tested with users to ensure the list holistically captured all possible seating preferences of individuals. Participants were also provided with blank slots to fill in any additional preferences that were not captured in the initial preference list. No participants made use of this option, indicating the provided list of seating preferences was adequate. The final list is shown in Table 1.

TABLE 1: PREFERENCE LIST

Arm Rest	Leg Rest
Cup Holder	Cushion
Head Rest	Wide
Narrow	Soft
Firm	Recline
Lounge	Fully Adjustable
Easily Stored	Comfortable
Back/Lumbar Support	Lightweight
Rolls	Swivels
Sturdy	Other: ______

4.3 Procedure

This study consisted of six phases, illustrated in Figure 3: Pre-Survey, User Preference Assessment, User and Designer Interviews, Designer Preference Assessment, Ideation, Prototyping, and a Post-Survey.

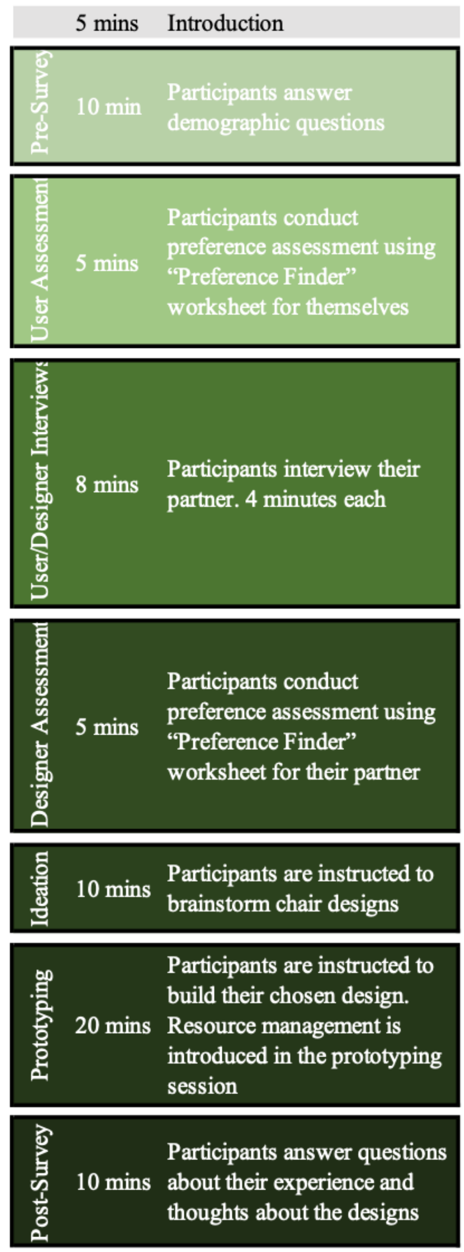

FIGURE 3: TIMELINE OF THE STUDY

We incorporated this study into a design thinking workshop. Before the study began, participants were assigned a unique identifier and were given ten minutes to complete an online survey about general demographic information. At the start of the study, participants were provided with a brief overview of design thinking and the following design challenge: *"How might we redesign the seating experience for our partners using only cardboard?"* Following this reflection, the participants were provided the list of preferences, shown in Table 1, and were instructed to list in order of importance their own preferences for a chair. They were told they did not have to use all of the preferences listed and were given approximately five minutes to determine their ranking of preferences for a chair. Participants were then randomly sorted into pairs and were instructed to interview their partner to determine their partner's preferences for the seating experience. Once interviews were complete, participants completed an empathy map, and sorted notes and observations from their interviews into a 2 by 2 grid as shown in Figure 4. Once the empathy map was complete, participants were once again provided with a preference list and were asked to rank the seating preferences of their partner based on the interviews and observations thus far. They were encouraged to use their interview notes and the empathy map to inform their decisions.

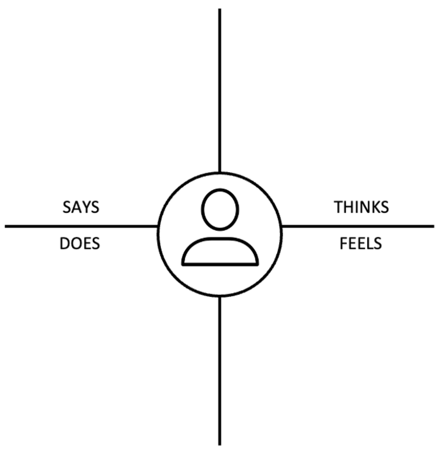

FIGURE 4: EMPATHY MAP

At this stage, the participants were made aware of the prototyping materials that were available. The prototyping materials were limited to cardboard, box cutters, duct tape, and zip ties. The participants were also aware the final prototype would be tested by having their partner sit in the chair for thirty seconds. The participants were then given ten minutes to generate ideas for seating solutions for their partner. Participants were instructed to sketch as many ideas as possible, ideas were captured using idea sheets which provided participants space to sketch and explain each unique idea. At the conclusion of the ten minutes, participants were instructed to select a single idea to prototype. Participants were then given twenty minutes to construct their chairs using the materials provided. When the prototyping session came to a close, the participants relocated the chairs to a central location to facilitate a group debrief and review of each design. The participants were asked to describe the chair and provide a rationale for their design choices based on their partner's unique needs. The participant's partner then tested the chair. At the conclusion of the of the design workshop participants completed a brief survey. Survey items asked participants to reflect upon their level of satisfaction with the chair *that was designed for them*. Responses to this survey question were used as a measure of user satisfaction; this is in alignment with the ISO 9001 standard which defines product quality in terms of user satisfaction [25].

4.4 Data Collection

Data was collected throughout the experiment using Qualtrics for the surveys, and the preference assessment sheets provided in the study.

Pre-Survey: The pre-survey consisted of five demographic questions such as age, gender, race, academic or professional standing, and participation in NSF I-Corps program.

Post-Survey The post-survey was at the conclusion of the study and asked participants to describe their experience with the design challenge as well as evaluate the chair they designed and the chair that was designed for them. Satisfaction was asked using a 5-point Likert Scale ranging from 1, completely disagree, to 5, completely agree. Participants responded to the following prompt: *"I am satisfied with the chair that was built for me"*. The responses from this question were used as a measure of user satisfaction.

Preference Assessment: The rank order of each participant were assigned to the features on the preference list. The items listed as "Most Important" were assigned rank of 1 with subsequent ranks following the order of importance. The features not identified by the participants, were assigned a rank of 21. The number 21 was used to signify the furthest distance from the most important feature, rank number 1, and allowed the researchers to account for unequal identification of features within the pair. For example, one participant may only rank five features while their partner may rank seven. An example of this scenario is shown in Table 2.

TABLE 2: EXAMPLE OF PREFERENCE ASSESSMENT COMPARISON

	Participant A (User)	Participant B (Designer)
Arm Rest	1	2
Cup Holder	2	1
Head Rest	3	6
Narrow	4	3
Firm	5	4
Lounge	21	5
Easily Stored	21	7
...	21	21

Weighted Cohen's kappa was used to calculate preference accuracy. Traditionally, weighted Cohen's kappa is used to establish inter-rater agreement for ordinal scales between raters, judges, or observers. We used weighted kappa to assess the agreement between preference ranking of the user and designer. The level of agreement between the user's ranking of their own preferences, and the designer's ranking of the user's preferences is used as the accuracy score. In this study, weighted kappa (κ_w) with linear weights [29] was run to determine the level of agreement between the user and designer preference rankings. This calculation was run for each pairing. For example, User 1A rank order was rated with Designer 1B rank order. Designer 1A rank order is rated with User 1B rank order. The level of agreement is assigned to the designer as a preference accuracy score.

The strength of agreement across the data spread from poor to good agreement according to Landis and Koch [30]. See Figure

5 for the distribution of preference accuracy for the corresponding partner. Reviewing Figure 5, we highlight that as a metric, preference accuracy did result in moderate variance across the population, with an average preference accuracy score of 0.394 and a standard deviation of preference accuracy of 0.182. This suggests that, given a larger population, preference accuracy may be used by researchers to more clearly distinguish the ability of designers to accurately identify and rank user preferences. We highlight that the accurate identification and ranking of user needs is dependent upon the ability of the designer to empathize with users.

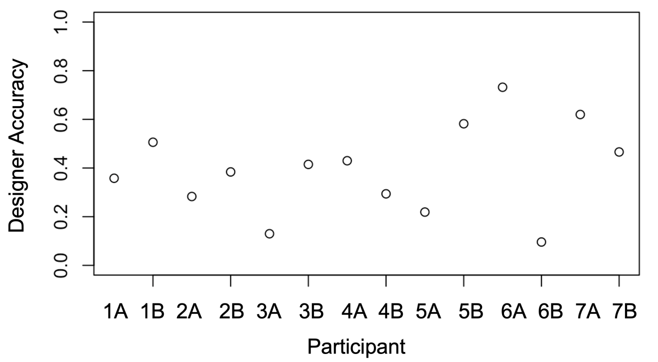

FIGURE 5: DISTRIBUTION OF PREFERENCE ACCURACY

5. RESULTS

In order to understand the viability of leveraging preference accuracy to estimate empathic ability, we sought to evaluate the predictive power of preference accuracy in determining user satisfaction. Statistical analyses were computed using SPSS software and the significance level for all analyses was 0.05.

Our hypothesis was devised to assess if preference accuracy could predict if the user would be satisfied by a design. We hypothesized that increased accuracy would lead to higher user satisfaction. User satisfaction was evaluated by a 5-point Likert scale. Presented in Figure 6 are the descriptive statistics for user satisfaction. An ordinal logistic regression was calculated to determine if preference accuracy predicts user satisfaction.

5.1 Predicting User Satisfaction

A cumulative odds ordinal logistic regression with proportional odds was run to determine if preference accuracy predicts user satisfaction. The dependent variable is user satisfaction. The independent variable is designer preference accuracy. The assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds location model to a model with varying location parameters, $\chi^2(1) = 0.682, p > 0.05$. Multi-collinearity was not tested because only one continuous independent variable was tested. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(25) = 28.785, p > 0.05$. However,the final model does not statistically significantly predict the dependent variable over and above the intercept-only

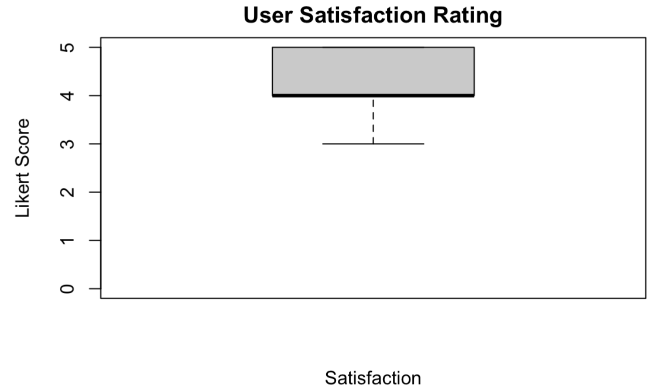

FIGURE 6: DISTRIBUTION OF USER SATISFACTION RATING

model, $\chi^2(1) = 0.992, p > 0.05$. An increase in preference accuracy was not associated with an increase in user satisfaction with an odds ratio of $0.058, 95\% CI[0, 19.132]$, Wald $\chi^2(1) = 0.924, p > 0.05$.

6. DISCUSSION

The main goal of the paper was to propose a new method of leveraging preference accuracy to measure the empathic abilities of designers. We define preference accuracy as the level of agreement between the user and designer for preference ranking. We hypothesized that preference accuracy would positively predict user satisfaction, as prior research suggests that the accurate identification of user needs is critical to the successful satisfaction of requirements. We highlight that a moderate distribution of preference accuracy scores was observed in this case study. With larger sample sizes we anticipate large distributions of preference accuracy, indicating the utility of this method to inform our understanding of designers' abilities to accurately identify and rank user preferences.

We did not find a clear statistical link between preference accuracy and user satisfaction. We argue, that two key factors could contribute to these findings, complicating the interpretation of findings. First, we observe a very small range of user satisfaction scores, indicating some level of social desirability bias. On average, the participants agreed that they were satisfied with the chair built for them. Some social desirability bias may have come into play because the participants did not want to come across as mean or ungrateful. Familiarity bias may have also been present as the participants established an emotional connection with their partner through the design thinking workshop and felt bad rating them poorly. Additionally, user satisfaction was obtained by a single Likert-type statement in the post workshop survey: *"I am satisfied with the chair that was built for me"*. The simplicity of this statement may have exaggerated the participants satisfaction. Specific statements about one-dimensional requirements [26] could result in more variation of satisfaction scores. Second, there is a moderate range of preference accuracy across the case study. According to Landis and Koch [30], the level of agreement ranged from poor to good, suggesting that preference accuracy scores could be a viable metric to delineate between designers' abilities. We highlight that this range demonstrates a difference in need finding ability among individuals.

The research efforts surrounding empathy in design have revolved around contextualizing everyday empathy into the specific nature of engineering design. The utilization of measurements from other fields and the creation of methods in engineering contexts have mixed results developing a relationship with design outcomes. Findings regarding the effect of empathy on design outcomes remains split, and our findings support this. While our work suggests that preference accuracy may be a useful method to determine designers' abilities to accurately identify and rank user preferences, we did not find a link between preference accuracy and user satisfaction. We underscore a critical gap in engineering design literature to understand when and under what conditions empathy positively or negatively affects design outcomes.

7. LIMITATIONS AND FUTURE WORK

This study attempted to devise a method to leverage preference accuracy as a measure of empathic ability. To achieve this goal, a preference assessment was completed by participants with the roles of users and designers. The agreement between user/designer ranking was used as a measure of preference accuracy. Preference accuracy was then analyzed as a predictor of user satisfaction. The findings from this study do not establish a clear link between preference accuracy and satisfaction, however the method can be used to measure designers' empathic ability.

There are several limitations in the study that could be addressed by future work. First, we recognize the small sample size of the study, with only fourteen participants, thus the results should be validated with larger samples. Further, the participants were recruited from the same program, in the same region. Validating the results with a more diverse pool of participants should be explored in future work as well.

Secondly, the evaluation of user satisfaction is limited. The participants were only asked if they were satisfied by the chair designed for them. A more critical analysis of satisfaction should be studied in future work to compare preference accuracy to a holistic view of satisfaction. Additionally, the design task of the study was to re-design the seating experience. Future work should test the bounds of preference accuracy in the context of a design task that is not as familiar to the participants.

Finally, this study is limited to establishing a relationship with user satisfaction. The success of a design is dependent on a myriad of factors such as quality, usefulness, uniqueness, and technical feasibility. Thus, future work should explore the relationship preference accuracy has on other design outcomes. The findings from this research contribute another method to measuring designer empathic ability. This research adds to the existing efforts to establish measures of empathic ability in engineering design.

8. CONCLUSION

While it is well accepted that empathy is critical to innovative design, the empirical evidence to support this claim is not so straightforward. The effects of empathy in design remain difficult to interpret. While empathy is central to Human-Centered design practices and Design Thinking, our understanding of the role empathy plays, empirically, in design remains muddied. Core to this challenge is the difficulty associated with measuring the empathic abilities of designers. Although some methods have recently emerged to assess the empathic abilities of designers within the context of design problems, often these methods are time- and resource-intensive, rendering them impractical for use in design practice. The overarching aim of this work was to investigate the utility of preference accuracy as a proxy measure of designers' empathic abilities. Empathy is a necessary prerequisite to the accurate identification and ranking of user preferences, and we therefore hypothesized that preference accuracy would positively predict user satisfaction. While findings suggest that preference accuracy may be a useful measure of designers' abilities to accurately identify user needs, we did not find a statistically significant link between preference accuracy and user satisfaction. We highlight further work is needed to more clearly understand the linkages between user satisfaction, empathy, and preference accuracy. This work builds fundamental knoweldge regarding the measurement of empathic abilities in designers, a critical first step towards the advancement of designers' empathic abilities.

REFERENCES

[1] Buchanan, Richard. "Wicked problems in design thinking." *Design issues* Vol. 8 No. 2 (1992): pp. 5–21.

[2] McLaren, Karla. *The art of empathy: A complete guide to life's most essential skill*. Sounds True (2013).

[3] Battarbee, Katja et al. *Co-experience: understanding user experiences in interaction*. Aalto University (2004).

[4] Surma-aho, Antti Oskari, Björklund, Tua A and Holtta-Otto, Katja. "Assessing the development of empathy and innovation attitudes in a project-based engineering design course." *2018 ASEE annual conference & exposition*. 2018.

[5] Konrath, Sara, Meier, Brian P and Bushman, Brad J. "Development and validation of the single item trait empathy scale (SITES)." *Journal of research in personality* Vol. 73 (2018): pp. 111–122.

[6] Rasoal, Chato, Danielsson, Henrik and Jungert, Tomas. "Empathy among students in engineering programmes." *European journal of engineering education* Vol. 37 No. 5 (2012): pp. 427–435.

[7] Hess, Justin L, Strobel, Johannes, Pan, Rui and Wachter Morris, Carrie A. "Insights from industry: a quantitative analysis of engineers' perceptions of empathy and care within their practice." *European Journal of Engineering Education* Vol. 42 No. 6 (2017): pp. 1128–1153.

[8] Fila, Nicholas D, Hess, Justin L, Dringenberg, Emily and Purzer, Senay. "Engineering students' utilization of empathy during a non-immersive conceptual design task." *The International journal of engineering education* Vol. 32 No. 3 (2016): pp. 1336–1348.

[9] Chang-Arana, Álvaro M, Piispanen, Matias, Himberg, Tommi, Surma-aho, Antti, Alho, Jussi, Sams, Mikko and Hölttä-Otto, Katja. "Empathic accuracy in design: Exploring design outcomes through empathic performance and physiology." *Design Science* Vol. 6 (2020): p. e16.

[10] Salmi, Aleksi, Li, Jie and Holtta-Otto, Katja. "Facial expression recognition as a measure of user-designer empa-

thy." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 86267: p. V006T06A022. 2022. American Society of Mechanical Engineers.
[11] Neumann, David L, Chan, Raymond CK, Boyle, Gregory J, Wang, Yi and Westbury, H Rae. "Measures of empathy: Self-report, behavioral, and neuroscientific approaches." *Measures of personality and social psychological constructs* (2015): pp. 257–289.
[12] Davis, Mark H. "Interpersonal reactivity index." (1980).
[13] Raviselvam, Sujithra, Sanaei, Roozbeh, Blessing, Lucienne, Hölttä-Otto, Katja and Wood, Kristin L. "Demographic factors and their influence on designer creativity and empathy evoked through user extreme conditions." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 58219: p. V007T06A011. 2017. American Society of Mechanical Engineers.
[14] Li, Jie, Surma-Aho, Antti and Hölttä-Otto, Katja. "Measuring designers' empathic understanding of users by a quick empathic accuracy (QEA)." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 85420: p. V006T06A027. 2021. American Society of Mechanical Engineers.
[15] Rae, Jeneanne. "Design Value Index Exemplars Outperform the Samp;P 500 Index (Again) and a New Crop of Design Leaders Emerge." (2016). URL https://www.dmi.org/page/2015DVIandOTW.
[16] Davis, Mark H et al. "A multidimensional approach to individual differences in empathy." (1980).
[17] Hoffman, Martin L. "Empathy, its development and prosocial implications." *Nebraska symposium on motivation*. 1977. University of Nebraska Press.
[18] Deutsch, Francine and Madle, Ronald A. "Empathy: Historic and current conceptualizations, measurement, and a cognitive theoretical perspective." *Human development* Vol. 18 No. 4 (1975): pp. 267–287.
[19] Davis, Mark H. "Empathic concern and the muscular dystrophy telethon: Empathy as a multidimensional construct." *Personality and Social Psychology Bulletin* Vol. 9 No. 2 (1983): pp. 223–229.
[20] Alzayed, Mohammad Alsager, Miller, Scarlett R, Menold, Jessica, Huff, Jacquelyn and McComb, Christopher. "Can design teams be empathically creative? A simulation-based investigation on the role of team empathy on concept generation and selection." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 83976: p. V008T08A024. 2020. American Society of Mechanical Engineers.
[21] Hess, Justin L, Sanders, Elizabeth and Fila, Nicholas D. "Measuring and promoting empathic formation in a multidisciplinary engineering design course." *ASEE annual conference exposition proceedings*. 2022.
[22] Hess, Justin L, Fila, Nicholas D, Kim, Eunhye and Purzer, Senay. "Measuring empathy for users in engineering design." *International Journal of Engineering Education* Vol. 37 No. 3 (2021): pp. 733–743.
[23] Li, Jie and Hölttä-Otto, Katja. "The influence of designers' cultural differences on the empathic accuracy of user understanding." *The Design Journal* Vol. 23 No. 5 (2020): pp. 779–796.
[24] Li, Jie, Surma-aho, Antti, Chang-Arana, Álvaro M and Hölttä-Otto, Katja. "Understanding customers across national cultures: the influence of national cultural differences on designers' empathic accuracy." *Journal of Engineering Design* Vol. 32 No. 10 (2021): pp. 538–558.
[25] Li, Jie and Hölttä-Otto, Katja. "Inconstant Empathy—Interpersonal Factors That Influence the Incompleteness of User Understanding." *Journal of Mechanical Design* Vol. 145 No. 2 (2023): p. 021403.
[26] Kano, Noriaki. "Attractive quality and must-be quality." *Journal of the Japanese society for quality control* Vol. 31 No. 4 (1984): pp. 147–156.
[27] Matzler, Kurt and Hinterhuber, Hans H. "How to make product development projects more successful by integrating Kano's model of customer satisfaction into quality function deployment." *Technovation* Vol. 18 No. 1 (1998): pp. 25–38.
[28] "About I-Corps." URL https://beta.nsf.gov/funding/initiatives/i-corps/about-i-corps.
[29] Cicchetti, Domenic V and Allison, Truett. "A new procedure for assessing reliability of scoring EEG sleep recordings." *American Journal of EEG Technology* Vol. 11 No. 3 (1971): pp. 101–110.
[30] Landis, J Richard and Koch, Gary G. "The measurement of observer agreement for categorical data." *biometrics* (1977): pp. 159–174.
[31] Kane, Richard M and Yuschik, Matthew. "A case example of human factors in product definition: Needs finding for a voice output workstation for the blind." *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*: pp. 69–73. 1986.
[32] Genco, Nicole, Johnson, Danny, Ho¨ ltta¨ Otto, Katja and Seepersad, Carolyn Conner. "A study of the effectiveness of empathic experience design as a creativity technique." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 54860: pp. 131–139. 2011.
[33] Bairaktarova, Diana, Bernstein, William Z, Reid, Tahira and Ramani, Karthik. "Beyond surface knowledge: An exploration of how empathic design techniques enhances engineers understanding of users' needs." *Int. J. Eng. Educ* Vol. 32 No. 1 (2016): pp. 111–122.
[34] Surma-Aho, Antti and Hölttä-Otto, Katja. "Conceptualization and operationalization of empathy in design research." *Design Studies* Vol. 78 (2022): p. 101075.
[35] Li, Jie and Hölttä-Otto, Katja. "Does Empathising With Users Contribute to Better Need Finding?" *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 86267: p. V006T06A023. 2022. American Society of Mechanical Engineers.
[36] Johnson, Daniel G, Genco, Nicole, Saunders, Matthew N, Williams, Paul, Seepersad, Carolyn Conner and Hölttä-

Otto, Katja. “An experimental investigation of the effectiveness of empathic experience design for innovative concept generation.” *Journal of Mechanical Design* Vol. 136 No. 5 (2014): p. 051009.

[37] Quinones, Pablo-Alejandro, Greene, Tammy, Yang, Rayoung and Newman, Mark. “Supporting visually impaired navigation: a needs-finding study.” *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (2011): pp. 1645–1650.

[38] Rismani, Shalaleh, der Loos, Van, Machiel, HF et al. “Improving needs-finding techniques for medical device development at low resource environments using Activity Theory.” *DS 87-1 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 1: Resource Sensitive Design, Design Research Applications and Case Studies, Vancouver, Canada, 21-25.08. 2017*: pp. 249–258. 2017.

[39] Loweth, Robert P, Daly, Shanna R, Sienko, Kathleen H and Liu, Jiangqiong. “Student reflections on needs finding in community-based design work.” *Proceedings of the Design Society: International Conference on Engineering Design*, Vol. 1. 1: pp. 569–578. 2019. Cambridge University Press.

[40] Brown, Tim and Katz, Barry. “Change by design.” *Journal of product innovation management* Vol. 28 No. 3 (2011): pp. 381–383.

[41] Kramer, Julia, Agogino, Alice M and Roschuni, Celeste. “Characterizing competencies for human-centered design.” *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 50190: p. V007T06A026. 2016. American Society of Mechanical Engineers.

[42] “The value of Design - Design Management Institute - dmi.org.” URL https://www.dmi.org/page/DesignValue/The-Value-of-Design-.htm.

[43] Mattelmäki, Tuuli, Vaajakallio, Kirsikka and Koskinen, Ilpo. “What happened to empathic design?” *Design issues* Vol. 30 No. 1 (2014): pp. 67–77.

[44] Gasparini, Andrea. “Perspective and use of empathy in design thinking.” *ACHI, the eight international conference on advances in computer-human interactions*: pp. 49–54. 2015.

[45] Tracey, Monica W and Hutchinson, Alisa. “Empathic design: Imagining the cognitive and emotional learner experience.” *Educational Technology Research and Development* Vol. 67 (2019): pp. 1259–1272.

[46] Tracey, Monica W and Baaki, John. “Empathy and empathic design for meaningful deliverables.” *Educational technology research and development* Vol. 70 No. 6 (2022): pp. 2091–2116.

[47] Ciavola, Benjamin, Ning, Yan and Gershenson, John K. “Empathic design for early-stage problem identification.” *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 44137: pp. 267–276. 2010.

[48] Walther, Joachim, Miller, Shari E and Sochacka, Nicola W. “A model of empathy in engineering as a core skill, practice orientation, and professional way of being.” *Journal of Engineering Education* Vol. 106 No. 1 (2017): pp. 123–148.

[49] Reniers, Renate LEP, Corcoran, Rhiannon, Drake, Richard, Shryane, Nick M and Völlm, Birgit A. “The QCAE: A questionnaire of cognitive and affective empathy.” *Journal of personality assessment* Vol. 93 No. 1 (2011): pp. 84–95.

