Design Science

Validation of a checklist-style intervention for mitigation of availability bias in professional designers

Anastasia Schauer [1], Habeeb Afolabi[2] and Katherine Fu [2]

[1]*School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[2]*Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI, USA*

Abstract

In this study, professional engineers and designers (n = 30) participated in a 1-hour-long design activity in which they brainstormed a list of ideas for two design problems (a smart grill and a smart laundry machine), created a sketched concept for each design problem, filled out a survey about their perceptions of the market for the concept they developed, participated in a bias mitigation intervention and then repeated the pre-intervention steps. The design problems were intended to trigger availability bias based on the participants' occupations (engineers and designers at a kitchen appliance company) as well as conflict between the gender of the participants and the gender-stereotyping of the household tasks fulfilled by the smart machines. Based on correlations in the market survey, the participants, who were mostly men, displayed availability bias toward the smart laundry machine design problem. A key marker of availability bias – an association between participants' personal enjoyment of the product and the belief that the product would be commercially successful – was eliminated after the bias mitigation intervention. Qualitative analysis of participants' reflections indicated that the intervention primarily assisted designers in making additional considerations for users, such as increasing accessibility and building awareness of excluded user groups.

Keywords: design creativity, design cognition, cognitive bias, bias mitigation, availability bias

Received 26 August 2024
Revised 21 April 2025
Accepted 24 April 2025

Corresponding author
Katherine Fu
kfu26@wisc.edu

Des. Sci., vol. 11, e23
journals.cambridge.org/dsj
DOI: 10.1017/dsj.2025.10006

1. Introduction

Historically, designers have been found to empathize more easily with users who are similar to them (Li & Hölttä-Otto 2020; Li *et al.* 2021). As a result, they may struggle to identify with users who they perceive to be different from themselves, whether this is due to differences in their bodies, demographics or culture. The lack of demographic diversity in the US engineering workforce thus presents challenges when designing for a wide variety of end users with diverse needs. For example, although many widespread products, such as cars, are intended for use by the general population, less than 20% engineers in the US are women (United States Bureau of Labor Statistics 2023), meaning that women's perspectives are greatly underrepresented in design processes. Historically, only test dummies representing the average-sized male body have been used in vehicle crash safety assessments

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

(Linder & Svensson 2019), resulting in car safety systems, such as seat belts, being optimized for men's bodies instead of women's bodies. Consequently, women are 47% more likely than men to sustain severe injuries in a comparable seatbelt-restrained car accident (Bose, Segui-Gomez, & Crandall 2011).

Availability bias is the tendency to bias judgments toward information that most readily comes to mind (Tversky & Kahneman 1973). In design, availability bias has been found to lead engineers to over- or underestimate the size of a potential market based on their preexisting understanding of customer needs and demand or based on their personal knowledge of the market (Fillingim, Shapiro, & Fu 2022). We posit that availability bias may also manifest as designers projecting their own needs, wants or characteristics onto a hypothetical user, resulting in the aforementioned lack of consideration for diverse users. This paper will present findings from a design activity developed to elicit availability bias in professional engineers and designers, allowing an investigation of how this implicit cognitive bias may impact concept development as well as perceptions of users. The second main goal of this study is to formalize a cognitive bias mitigation treatment and examine its impact on the outcomes of the design activity.

1.1. Empathy in design

Empathy-building is a key factor for successful design and should be applied throughout the design process (Chang-Arana *et al.* 2022), but it is most critical early in the design process when user involvement is high (Surma-Aho, Björklund, & Hölttä-Otto 2018). The Stanford d.school model of human-centered design features empathy as the "centerpiece" and first step of human-centered design processes, followed by problem definition, ideation, prototyping and testing (Hasso Plattner Institute of Design 2010). Surma-aho and Hölttä-Otto (2022) identified understanding, action, research, orientation and mental processes as the five core principles that form the concept of empathy in design. Empathy-building may be integrated into customer need-finding methods through user observations, interviews, focus groups, complaints or surveys (Ulrich & Eppinger 2007; ideo.org 2015). The modality through which a design problem is communicated to a design team, such as narrative storytelling or personas (fictional representations of users), can also foster engagement and empathy between a designer and user (Grudin 2006; Chen, Nivala, & Chen 2011; Carmel-Gilfilen & Portillo 2016; Marsden & Haag 2016). Despite these strategies, designers' inherent cognitive biases can inhibit the development of equitable designs, although this phenomenon is not well-studied in the field of mechanical design.

Work on wearable simulations has sought to expand designers' ability to empathize with physically disabled users (Heylighen & Dong 2019). For example, McDonagh, Woodcock, and Iqbal (2018)) found that using glasses, ear plugs and gloves to reduce hand sensation to simulate the experience of users with reduced physical senses was effective in building empathy in designers. Raviselvam, Hölttä-Otto, and Wood (2016)) found that simulating visual impairments in designers resulted in higher empathy and more creative designs developed for visually impaired users compared to designers who were simply briefed about the target users, and these benefits were enhanced when designers had a pre-existing connection to the target user population (Raviselvam *et al.* 2017). Designers in a study by Boffi *et al.* (2014) wore a hand-shaking device to simulate the experience of users with Parkinson's disease while using a product, but researchers ultimately found

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

that a greater understanding of the end-users resulted from close observation and interaction with them. The Engineering Design Center at the University of Cambridge offers an online Inclusive Design Toolkit containing resources such as a calculator allowing designers to calculate the percentage of British adults who would be unable to effectively use a product or service and the opportunity to purchase glasses and gloves that simulate vision loss or mobility limitations (Engineering Design Centre, University of Cambridge 2024).

Cook stoves make an excellent case study for failure to empathize with the end user while designing a product. In South Asia, 75% of families use biomass fuels for energy (Bailis *et al.* 2015). When used indoors, biomass stoves create toxic fumes that are responsible for 2.9 million yearly deaths worldwide (Olopade *et al.* 2017). However, efforts to develop and implement improved cookstoves have low adoption rates because they lack basic design attributes that would meet the needs of users (Khandelwal *et al.* 2017), causing an increase in cooking time and active attendance to the stove (Perez 2019). In this scenario, the problem was that stove designers, men located in the U.S., did not consult the Bangladeshi women who would be using the stove, instead prioritizing technical parameters in their design process. As a result, 98% of the population of rural Bangladesh continued to use traditional biofuel-burning stoves, despite hundreds of attempts to introduce cleaner stoves since the 1980s (Bailis *et al.* 2015; Perez 2019).

As a counterexample, the Embrace baby incubator is an award-winning incubator designed for low-birth-weight babies in countries with less industrialized economies ("Embrace", no date). While designing this product, the design team traveled to rural Nepal to conduct their need-finding, where they found that an incubator would be useless to the customers who needed it most if it was to run on electricity and be kept in a hospital. The result of this project was the Embrace incubator, which was developed to keep babies warm using a packet of phase-change material inside a type of sleeping bag that was designed for easy integration into local culture (Misra 2014). This case study reinforces the need for empathic design and successfully implements some of the techniques described above. Given its success, further study is needed to understand the factors that cause the story of the failed cookstoves to be more common than the story of the successful Embrace incubator and to explore the effect of designers' empathy with end users.

These studies highlight dramatic challenges that can arise when designing for an end user outside of one's own culture. Even within the same community and culture, a breakdown in understanding and empathy between designers and users can have catastrophic impacts. As previously discussed, the usage of male-bodied crash test dummies has resulted in higher rates of severe injuries for women in seatbelt-restrained car accidents compared to men (Bose *et al.* 2011; Linder & Svensson 2019). Even more dangerous than being a woman driver is being a pregnant woman driver – the largest cause of accidental death for pregnant women is automobile accidents. Guidance on how to properly position a seatbelt on one's body is challenging and uncomfortable to apply to a pregnant body, resulting in decreased seatbelt use among pregnant women (Acar, Edwards, & Aldah 2018). To better protect pregnant women and their fetuses, products such as seat belt extenders have been created to be used as an aftermarket modification to seat belts.

Despite the positivity of innovations such as this, the question remains: why don't cars accommodate women's pregnant bodies in the first place? With over 3 million births registered in the United States in 2020 (National Center for Health

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Statistics 2022), pregnant women comprise a not-insignificant proportion of automobile users. Even when not pregnant, women still expose themselves to significantly more risk than men do when driving a car. While this design flaw (and others) has been retroactively identified as harmful to women, the damage has been done. There has been little formal investigation into when and how the design process disadvantages women or other members of underrepresented groups in engineering. A deeper understanding of how the design process is affected by cognitive bias would allow designers to proactively prevent situations such as this, rather than remaining blind to them until statistics begin to show consequences.

1.2. Cognitive bias in design

Although heuristics (defined by Fu, Yang, and Wood (2016)) as "a context-dependent directive, based on intuition, tacit knowledge, or experiential understanding, which provides design process direction to increase the chance of reaching a satisfactory but not necessarily optimal solution" are often leveraged in engineering as valuable education tools, the erroneous application of heuristics as mental shortcuts may manifest as cognitive bias in a variety of scenarios (Tversky & Kahneman 1974). Korteling and Toet (2022) identified four "tendencies," or categories of cognitive bias, that humans exhibit when carrying out high-complexity tasks such as decision-making. In short, these tendencies are (1) to associate (unrelated) information, (2) to prioritize information in accordance with present beliefs, (3) to retain information or to falsely believe that information is relevant and (4) to ignore relevant information while focusing on more accessible information (Korteling & Toet 2022). The following paragraphs will provide examples of cognitive biases that correspond to each tendency and a brief discussion of how the biases may manifest in the design field.

The tendency to establish relationships and patterns by associating information that may not be related can often manifest as representativeness bias, stereotyping, or superstition (Korteling & Toet 2022). The representativeness heuristic leads people to evaluate the probability of an occurrence based on similarities to a "typical case" rather than by simple base rate (Kahneman & Tversky 1972). In other words, people may falsely believe that a small sample of information is representative of a larger population. Representativeness bias can also lead to drawing inaccurate conclusions about large groups from small or incomplete samples, which, in design, may manifest as the stereotyping of a target user group (Hallihan, Cheong, & Shu 2013). In some cases, people are even able to recognize that they hold erroneous stereotypes and associations, but superstition leads them to follow these beliefs regardless (Risen 2016).

The tendency to prioritize information according to current expectations, knowledge and choices may manifest as sunk cost fallacy, belief bias, cognitive dissonance, familiarity bias, status quo bias or system justification (Korteling & Toet 2022). Both status quo bias and sunk cost fallacy can lead to negative outcomes as a result of the desire to "stay the course" or maintain a current plan or decision (Samuelson & Zeckhauser 1988; Arkes & Ayton 1999). In particular, sunk cost fallacy results from a desire to save money, time or effort (Arkes & Blumer 1985). In engineering, when time and resources have been invested into developing physical prototypes, designers tend to fixate on concepts that have been prototyped as a way of avoiding perceived wastefulness (Viswanathan & Linsey

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

2013). Design fixation, sunk cost fallacy and status quo bias contribute toward less innovation, novelty and variety in ideation and in design solutions (Fillingim *et al.* 2022).

The tendency to retain information that is irrelevant or better ignored can lead to decision-making based on impertinent information and may manifest as hindsight bias, anchoring bias, outcome bias or framing bias (Korteling & Toet 2022). For example, framing a question about a design problem or concept with either positive or negative connotations can impact the way designers answer the questions (Tversky & Kahneman 1981; Fillingim *et al.* 2022). The effort heuristic, in which solution alternatives are evaluated based on the amount of effort put into developing them rather than their actual value (Kruger *et al.* 2004), can also impact designers' desire to bring a product to market based on the amount of effort they felt they had put into developing the product (Fillingim *et al.* 2022). The views and preferences of teammates may account for potential sources of bias in design decision-making. The bandwagon effect can lead to a tendency to "join the crowd" (Leibenstein 1950), resulting in a bias toward the majority opinion or preference on a design team (Fillingim *et al.* 2022). Another finding related to bias impacting design teams by Fillingim *et al.* (2022)) revealed that designers exhibited hindsight bias by believing that they should have advocated more for their ideas throughout the course of a design project, or that mistakes made should have been easily avoidable.

Finally, the tendency to focus on certain information for decision-making while neglecting relevant information may manifest as confirmation bias, availability bias, survivorship bias, ego-centric bias or the priority heuristic (Korteling & Toet 2022). Most of these biases relate to erroneous prioritization of information; for example, the priority heuristic refers to a tendency to base decisions on a single prominent piece of information (Brandstätter, Gigerenzer, & Hertwig 2006). Confirmation bias is a tendency to seek out and interpret evidence or information that supports one's pre-existing beliefs, often accompanied by a disregard for disconfirming information (Nickerson 1998). Confirmation bias has been directly studied in the design field, with researchers finding that the bias may lead to the discounting of unexpected information about a product or user group (Hallihan *et al.* 2013), or a preference toward making design decisions in line with user feedback that confirmed designers' initial beliefs (Fillingim *et al.* 2022).

Availability bias is the tendency to bias judgments toward information that most readily comes to mind (Tversky & Kahneman 1973). In design, this may lead engineers to over- or underestimate the size of a potential market based on their preexisting understanding of customer needs and demand or based on their personal knowledge of the market (Fillingim *et al.* 2022). Similarly, ownership bias may result in the inflated importance of customer needs that a designer was previously aware of Fillingim *et al.* (2022), but it may also manifest as a preference for one's own ideas compared to others' ideas on a design team (Toh *et al.* 2016). Toh, Strohmetz, and Miller (2016)) further explored ownership bias in engineering design teams and found that while men exhibited ownership bias as a preference toward their own ideas, women were more likely to show the opposite bias by selecting ideas generated by their teammates. Survivorship bias, the tendency to overlook elements that did not survive a selection process (Brown *et al.* 1992), may also manifest in concept selection in the form of discarding customer needs or design features that were not prioritized in the selected concept.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

1.3. Bias mitigation

Education or training specifically designed for raising awareness of bias, such as workshops (Carnes *et al.* 2015) or brochures (Legault, Gutsell, & Inzlicht 2011), has been tested in a diverse range of fields to varying degrees of success. In "draw a scientist" tests (DASTs), researchers were able to increase the proportion of women scientists drawn by children after an intervention involving classroom visits with women scientists (Flick 1990). This technique can be used to reduce bias even when it is not performed face to face; exposure to information about women's successes in traditionally masculine occupations has been found to be successful in reducing gender bias in hiring selection decisions (Heilman & Martell 1986). An intervention developed by Jenkins and Youngstrom (2016) was successful in reducing occurrences of base rate neglect and search-satisfying bias. The intervention consisted of education on common cognitive pitfalls when making diagnoses of pediatric bipolar disorder, as well as information on avoiding these common pitfalls in the future. However, similar interventions targeting self-serving bias (Babcock & Loewenstein 1997), availability bias and search-satisfying bias (Sherbino *et al.* 2011, 2014) did not yield a significant improvement in participants' decision-making skills.

A similar form of bias education, video interventions for diversity in STEM (VIDS), was developed by Moss-Racusin *et al.* (2018) with the goal of targeting gender biases that contribute to the underrepresentation of women in STEM. VIDS leveraged both narrative storytelling as well as facts presented by experts to successfully increase awareness of gender bias, positive attitudes toward women in STEM, anger, empathy and motivation to promote gender parity in STEM fields (Moss-Racusin *et al.* 2018). Serious games, or applied games, have been shown to be even more effective than bias training videos due to their interactive elements (Dunbar *et al.* 2017). *Missing* is one such educational game in which participants search for a missing neighbor in a first-person point-of-view game targeted at eliciting and mitigating bias blind spot, confirmation bias and fundamental attribution error (Barton *et al.* 2016). The game was not only successful in producing debiasing effects within the initial study session but also in producing positive effects that persisted for at least 2 months (Morewedge *et al.* 2015). Similarly, *Heuristica* is a modular puzzle game targeted at reducing the same three biases as *Missing* and was similarly successful in promoting knowledge of cognitive biases as well as enduring bias mitigation effects (Mullinix *et al.* 2013).

Checklists are tools that are often used in bias mitigation scenarios to facilitate metacognition, which can aid in reflection (Chew, Durning, & van Merriënboer 2016). These interventions have often been validated in medical settings, in part thanks to a well-established body of work on the impacts of bias in the medical field (Agyemang, Andreae, & McComb 2023). For example, checklists have been found to enhance electrocardiogram interpretation by reducing errors; although the decisions took longer to make with the checklist, perceived cognitive load did not increase (Sibbald, de Bruin, & Merrienboer 2013). Another checklist was found to increase the accuracy of cardiac exam diagnostics, but only if the user retained access to the checklist and was able to return freely to it (Sibbald *et al.* 2013). Yet another study of 76 doctors found that a cognitive forcing tool did not significantly impact the amount of diagnostic errors made, but qualitative results from a smaller pool of participants revealed that doctors had an overall positive experience with

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

the checklist, claiming that it helped them "slow down and avoid bias" (O'Sullivan & Schofield 2019). Formalized step-by-step procedures have also been developed and recommended for bias mitigation and error management in fields such as fire engineering (Kinsey *et al.* 2021) and the aerospace sector (Emmons *et al.* 2018).

Another method of reducing cognitive bias is by removing the opportunity for people to apply stereotypes or biases toward their interpretation of a situation or group of people. This can be done by selectively presenting data to participants in a way that is expected to minimize bias. For example, Mourad and Tewfik (2016) developed an algorithm to select and order data in a way that would aid participants in making optimal, unbiased decisions. Selection bias can also be mitigated through the use of data visualization tools such as DataPilot, developed by Narechania *et al.* (2023). "Blinding" data, or removing demographic markers that may otherwise result in the application of stereotypes or bias, is another way to mitigate cognitive biases that may arise from demographic factors. In research projects conducted at the Hubble Space Telescope, male principal investigators were more likely than women to have research proposals accepted (Reid 2014); however, women's success rate increased when proposals were anonymized (Johnson & Kirk 2020). Similarly, more women were hired to symphony orchestras when auditions were "blinded" with candidates concealed behind a screen (Goldin & Rouse 2000).

A more drastic method of reducing cognitive-biased behavior is to confront people with their biases. Many methods have been developed and validated for measuring or quantifying implicit cognitive biases. For example, the "beads task" is a method of measuring participants' likelihood of "jumping to conclusions" based on available information (Phillips & Edwards 1966), while the "Hinting Task" measures social cognition, particularly one's proficiency at "mental state reasoning" (Lindgren *et al.* 2018). In line with the previous discussions of bias mitigation in medical fields, many of these scales were developed to assess psychosis or delusion in participants with a history of mental illness (Green *et al.* 2008; van der Gaag *et al.* 2013). Enumerating individuals' bias is useful not only as a method of tracking changes in cognitive bias over time but also as a mitigation tool. Being presented with a quantification of one's own implicit gender biases using the Harvard Implicit Association Test (Nosek, Banaji, & Greenwald 2002) has been found to reduce the number of gender-biased decisions made in reporting by journalists (Kalra & Boukes 2021). In addition, the serious games previously described each contained a feedback element in order to educate participants on the amount of biased decisions they personally made (Mullinix *et al.* 2013; Morewedge *et al.* 2015; Dunbar *et al.* 2017).

1.4. Bias mitigation in design practice

The techniques previously discussed to encourage empathy in design may also be used to reduce gender bias in design. Gender bias and difficulty with empathy-building are closely tied in engineering design due to the disparity in the proportion of women engineers (United States Bureau of Labor Statistics 2023) compared to the proportion of women in the general population (U.S. Census Bureau 2023), indicating that many engineers may struggle to empathize with women users (Li & Hölttä-Otto 2020; Li *et al.* 2021). Similar to the "draw a scientist" intervention, interactions with sample users may help designers to reframe women as an

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

important user of their product. Representing women users throughout the design process builds designers' ability to empathize with them (Heylighen & Dong 2019), increasing the chances that a product will be accessible to them.

Similarly, user personas are a previously discussed tool to help designers build empathy for users' specific needs. Visser and Stappers (2007) found that personas with more visual detail stimulate more empathy and a deeper view of users' needs, and a separate study by Riggs and Knobloch–Westerwick found that audio passages are more effective at stimulating feelings of empathy compared to textual narratives. GenderMag is a tool that encourages developers of problem-solving software to adopt the mindset of sample user personas to identify any issues with their product's gender inclusivity (Burnett, Stumpf, *et al.* 2016). In testing, the tool has been effective in enabling participants to identify gender-inclusiveness issues (Burnett, Peters, *et al.* 2016), and the developers of the tool speculate that its scope could be widened and applied to a larger range of products (Burnett *et al.* 2016). One practical application of GenderMag not only closed the gender gap for successfully using the software by improving women's success rates, but men's success rates also increased when using the more gender-inclusive software (Vorvoreanu *et al.* 2019). A systematic analysis such as the Social Impact Evaluation Analysis may also be an effective tool for helping designers consider far-reaching social impacts of their design work (Armstrong *et al.* 2024).

Genco *et al.* (2012) developed the Empathic Experience Design (EED) method to engage designers in "bodystorming," which Kelley and Littman (2001) define as a process where designers "act out current behavior/usage patterns and see how they might be altered. The five steps of the EED are to (1) define the design problem, (2) define typical and empathic users and usage environments, (3) design empathic experiences, (4) simulate empathic experiences and (5) generate concepts (Genco *et al.* 2012). Participants who interacted with products while experiencing a simulated impairment in vision or dexterity developed concepts with significantly higher originality with no decrease in technical quality. The wearable simulations discussed previously are other methods of bodystorming that can build designers' creativity and ability to empathize with users with physical disabilities (Boffi *et al.* 2014; Raviselvam *et al.* 2016; Raviselvam *et al.* 2017; McDonagh *et al.* 2018).

There have been many recent attempts to develop protocols or methodologies for mitigating the effects of design fixation, a form of sunk cost bias. For example, Zahner *et al.* (2010) found that presenting problem statements in more abstract forms resulted in more original ideas, although the solutions developed did not necessarily conform to the constraints of the problem. It was found that even more novel solutions were developed by designers who conducted the re-representation of a problem statement themselves; in other words, participants who were given a concrete design problem and then rewrote it as a more generalized problem statement were more original when developing concepts to solve the problem (Zahner *et al.* 2010). Moreno *et al.* tested two design-by-analogy methods, WordTree and SCAMPER, finding that although both methods were effective in increasing the novelty of generated ideas, which addressed one potential consequence of sunk cost bias, only the WordTree method was effective in mitigating design fixation (Moreno *et al.* 2016). Chrysikou and Weisberg (2005) found that providing pictorial examples in a design problem-solving scenario can cause fixation resulting in the inclusion of unsuitable features from the examples,

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

although inclusion of specific defixation instructions resulted in fewer design features indicative of fixation.

As previously established, cognitive bias commonly manifests at the concept development or concept selection stages of the design process. Tools such as weighted decision matrices or Pugh charts provide methods of facilitating concept selection (Dieter & Schmidt 2013). Hallihan *et al.* (2013) applied a formalized decision matrix to a concept evaluation design activity and found that participants who used the decision matrix were less biased and more thorough when evaluating data. This finding validated the use of guided decision-making as a method for mitigating confirmation bias. However, engineering students often tend to rely on their own intuition over formalized methodologies for problem-solving (Taleyarkhan *et al.* 2023), and as a result, three out of seven graduate student participants in the study by Hallihan *et al.* expressed that the decision matrix was an "unnatural" way of thinking for them. This participant discomfort, in addition to the increased cognitive effort required to utilize the procedure compared to relying on intuition, may present an obstacle to the widespread adoption of formalized bias mitigation in design (Hallihan *et al.* 2013).

2. Aims

Through a design activity with 30 professional engineers and designers, the following research questions were investigated in this study:

RQ1: How do indicators of availability bias manifest through design outcomes in a brainstorming and sketching design activity?

Because participants worked on two design problems at a time with a limited amount of time to complete both problems, it was hypothesized that availability bias may lead participants to prioritize the design problems according to their level of interest in the topic or whether or not they felt that they would be a target user for the product. This was expected to result in more brainstormed ideas being developed for the problem that participants identified more strongly with (H1-A). It was also hypothesized that participants would make less considerations for end users other than themselves as a result of availability bias, leading to fewer user representations in design sketches for the problem that was more relatable to participants (H1-B), since Makhlouf *et al.* (2023) have connected sketches of users with greater consideration for users.

RQ2: How do indicators of availability bias manifest through designers' perceptions of the market for a product and the value of a product?

Based on findings from Fillingim *et al.* (2022)), it was expected that availability bias would play a role in how participants viewed the marketability and likelihood of success for the products they design. In particular, it was expected that participants would overvalue the market size and value of the product that was more relatable to them and undervalue the market size and value of the product that they found less relatable (H2-A). For both design problems, it was expected that high value and market size would be positively correlated with participants' level of interest in the product (H2-B). It was also expected that in a head-to-head comparison, participants would display preference toward the product that was more relatable to them in positively framed questions about the market (H2-C).

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

RQ3: How does a checklist-style bias mitigation intervention impact manifestations of availability bias and considerations made for a diverse collection of users?

The checklist-style bias mitigation intervention was expected to reduce the markers of availability bias discussed in RQ1 and RQ2. For example, it was expected that the number of brainstormed ideas for the two design problems would be more similar post-intervention compared to pre-intervention (H3-A). The checklist was intended to help designers develop an awareness of diverse potential users, so it was hypothesized that more representations of users would be present in the post-intervention sketches compared to the pre-intervention sketches (H3-B). Finally, it was expected that the intervention would reduce the positive correlations between participants' interest in a product and their view of the product's value and market size, as well as lead to more evenly distributed scores in the head-to-head comparison of the two products (H3-C).

3. Materials and methods

To explore the research questions, data were collected during an eight-step within-subjects design activity, as shown in Figure 1. Participants were recruited from a selection of employees at a Midwestern kitchen appliance company who had chosen to participate in a half-day design workshop. First, participants read and agreed to a consent document. The study, including the consent procedure, was conducted under the guidance of the Institutional Review Board at Georgia Institute of Technology. After agreeing to participate in the study, participants were given 7 minutes to complete Step 1 of the study, which contained two design problems for each of which participants developed a written list of brainstormed ideas. The two design problems involved the development of features that could be included in an innovative "smart grill" or "smart laundry machine," and the prompt suggested that participants could include concepts from artificial intelligence (AI), the internet of things (IoT) or robotics. Participants were given a total of 7 minutes to complete both design problems and were allowed to allocate their time however they preferred.

Next, participants moved on to Step 2, the sketching phase of the study. Participants were instructed to create one detailed concept sketch for each of the two design problems that incorporated features from their brainstormed list from Step 1. Once again, participants were given 7 minutes to allocate toward the two sketch problems as they preferred.

After completing their sketches, participants were given 5 minutes to complete Step 3, the market survey. First, participants provided a written response to the

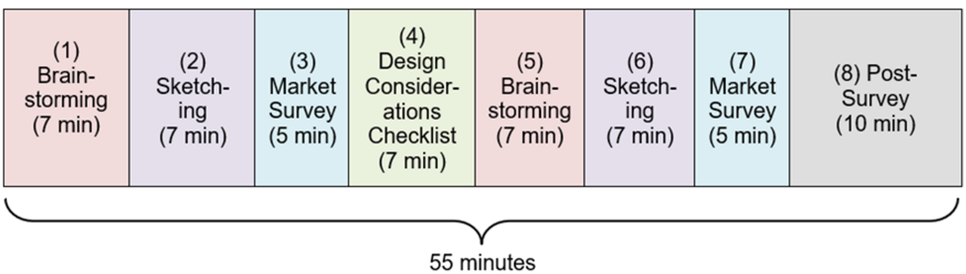

Figure 1. Timeline of study.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

open-ended questions: "Who is the smart [grill/laundry machine] for? Who do you imagine would use it?" (Makhlouf *et al.* 2023). Next, they provided an estimate to the nearest $100 of how much the smart grill/laundry machine would be sold for. They then used a 1–5 Likert scale to assess their level of agreement with a series of seven statements regarding their design, such as "The market for the smart [grill/laundry machine] is large," and "The smart [grill/laundry machine] is likely to be commercially successful." Participants answered the open-ended question, price question and Likert questions twice each: once while considering the smart grill concept they sketched, and once while considering the smart laundry machine concept that they sketched in Step 2. Finally, participants used a 1–5 Likert scale (1 – definitely the grill, 2 – probably the grill, 3 – neither the grill nor the laundry machine, 4 – probably the laundry machine, 5 – definitely the laundry machine) to answer a series of four questions directly comparing their concepts for the smart grill and smart laundry machine, such as "Which is more likely to be commercially successful?" and "Which would you use more often?"

After completion of the brainstorming, sketching and market survey (Steps 1–3), participants completed the Design Considerations Checklist (full text in the Appendix). The checklist contained 10 questions that were intended to guide participants in reflecting on the decisions and considerations that they had made during Steps 1–3. For example, some questions were "What assumptions did you make while creating this design?" and "How was your design influenced by your judgement of who the user is? Did you apply stereotypical associations to your perception of the user?" Participants were given 7 minutes to complete the Design Considerations Checklist. They were instructed to check the box next to each question when they felt that they had sufficiently considered it, and space was provided under each checklist item for participants to reflect or take notes. The development of the checklist items is discussed further in Section 3.1.

Because of the within-subjects format of Study III, participants repeated Steps 1–3 after the intervention. The instructions for each of the three parts were identical to the first time they were presented, except for that during the brainstorming phase, participants were told to feel free to repeat any ideas from their first round of brainstorming. After repeating each of the three steps, participants were given 10 minutes to complete the post-survey. In the post-survey, participants filled out demographic information, including their age, gender identity, race/ethnicity, higher education and professional experience as an engineer or designer. They also evaluated their level of familiarity with smart appliances, AI, IoT and robotics on a scale from 1 (novice) to 5 (expert). In the second section of the post-survey, participants considered their typical interactions with a grill and with a laundry machine. They provided information such as whether or not they own the appliance, who in their household typically interacts with the appliance, how many times per week they use the appliance and how much they enjoy using the appliance. In the final section of the post-survey, participants reflected on their experience with the Design Considerations Checklist intervention, evaluating how helpful they felt it was while reflecting on their designs from the first round of brainstorming and sketching, as well as while developing new concepts in the second round of brainstorming and sketching. Finally, they provided a written response to the open-ended question "How did the checklist impact the way you approached the design problems and questions in the second round?"

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

3.1. Study material development

The smart grill and smart laundry machine design problems were selected due to the likelihood that they would elicit availability bias in participants. During the design of the study, researchers hypothesized that participants would prioritize the grill problem over the laundry machine problem due to (1) the relationship between the demographics of the participant pool and the stereotypical gendered breakdown of household tasks and (2) the connection between the grill problem and the participants' roles as kitchen appliance engineers.

Because the participant demographics were skewed toward men, it was expected that differences in the way various household tasks are gender-stereotyped could impact the way participants interacted with the design problems. A 2015 survey of American households found that the primary griller was twice as likely to be a man than a woman (Applegate 2015). Conversely, a 2019 Gallup poll about household chore breakdowns in heterosexual American households found that women were more than twice as likely as men to be primarily responsible for doing the household's laundry (Brenan 2020). Because 23 out of the 29 study participants who provided their gender identified as men, it is likely that participants would feel more connected or have more prior experience with the grill problem. This was expected to manifest in various ways – for example, it was hypothesized that participants would develop more grill concepts, place a higher value on their grill concept and view the user of the smart grill as more similar to themselves, potentially resulting in an over-expanded view of the market for the smart grill.

Originally, the two design problems were a smart outdoor grill and a smart indoor stovetop due to their functional similarity; however, this iteration of the study design was unsuccessful during piloting *because* of the functional similarity – pilot participants developed near-identical feature lists for the two design problems during Step 1 and near-identical sketches during Step 2. Replacing the indoor stovetop with a laundry machine had the advantage of differentiating the design problems and also provided another opportunity for availability bias to impact participants: because all participants were presently employed at a kitchen appliance design company, it was expected that they may show a preference toward the grill problem due to their level of familiarity with the design and engineering of cooking appliances compared to cleaning appliances.

The order in which the smart grill and smart laundry machine were presented was randomized for each participant to mitigate the impact of serial position bias, as people tend to show a preference for the first item presented in a sequence (Mantonakis *et al.* 2009). When experimental packets were passed out, participants were assigned a numeric identifier so that data could be collected anonymously. Participants with odd-numbered packets received the grill problem first in Step 1, the laundry problem first in Step 2 and the grill problem first in Step 3, with the pattern repeating post-intervention. Participants with even-numbered packets received the laundry problem first in Step 1, with the order of the problems alternating thereafter.

A checklist-style bias mitigation tool was selected due to its ease of use and its ability to be customized and applied to various design scenarios. The Design Considerations Checklist was developed from a range of proposed or validated decision checklists intended to reduce bias or error in users. Because such a checklist had not previously been formalized for the field of engineering design,

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

the 10 items on the checklist were derived from protocols related to fire engineering (Kinsey *et al.* 2021), data crowdsourcing (Draws *et al.* 2021), weather (Walmsley & Gilbey 2019), energy (Cao *et al.* 2016), emergency medicine (Jin, Anaebere, & Haar 2021), behavioral economics (Kahneman, Lovallo, & Sibony 2011) and pilot training (Murata, Nakamura, & Karwowski 2015). For example, the emergency medicine checklist encouraged medical practitioners to ask themselves “Are there cultural differences in the patient’s expression of frustration and control?” when determining whether to physically restrain an agitated patient (Jin *et al.* 2021). In the Design Considerations Checklist, this concept was adapted to the design field with the checklist item, “Are there cultural differences in the way people might use this product?” Questions on the checklist were aimed at mitigating bias by encouraging designers to reflect on how their past experiences or implicit biases may have influenced their thought process. Specifically, several items on the checklist targeted availability bias by encouraging participants to think beyond their own needs or preferences as a consumer, and instead consider how a variety of users with different backgrounds and needs may interact with products.

3.2. Participants

In total, 30 workshop participants attended the workshop and provided consent for their data to be used. Of the participants who provided demographic data, 23 were men, 5 were women, and 1 identified as non-binary. The average participant was 35 years old (±10.6 SD) and had 11.8 years (±9.9) of professional experience in design or engineering. Of participants who reported their race, 23 were White; 2 were Hispanic or Latino; 2 were Asian, Native Hawaiian, or Other Pacific Islander and 1 identified as more than one race. All participants held a bachelor’s degree, with the most common being a Bachelor of Science in Electrical Engineering (11 participants), Computer Engineering (6), Mechanical Engineering (5) and Computer Science (5). In addition, three participants held a Master of Science in Electrical Engineering, and other degrees included an MBA, as well as bachelor’s degrees in Engineering, Software Engineering, English Literature, Physics, Systems Engineering, Information Systems and Software Development. When asked to evaluate their level of familiarity with various technologies as novice, beginner, proficient, advanced or expert (corresponding to a scale from 1–5), participants reported average experience levels of 3.13 (±1.15 SD) with smart appliances, 2.30 (±0.86 SD) with artificial intelligence, 3.07 (±1.03 SD) with Internet of Things and 2.43 (±0.92 SD) with robotics.

4. Results and analysis

Statistical analysis was conducted using R 4.1.2 and RStudio. Because the sample size (n = 30) was sufficient to apply the Central Limit Theorem and assume a normal distribution (Islam 2018), data were analyzed using analysis of variance (ANOVA) tests at a significance level of $\alpha = 0.05$ to consider all variables as well as their potential interactions. When a significant relationship was indicated by ANOVA, Dunn’s test was used to perform multiple pairwise comparisons (Dunn 1964). Reported p-values have been corrected for multiple pairwise comparisons. The Kendall rank correlation coefficient (τ) was selected to check for

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

statistically significant correlations between variables, due to the ordinal nature of the data and the prevalence of ties (Kendall 1938).

Cohen's Kappa was used in Section 4.1 to validate a coding rubric in which codes were mutually exclusive (only one code is applied to each data point). Because Cohen's Kappa cannot be calculated when codes are not mutually exclusive (more than one code can be applied to each data point), simple percentage agreement was used to validate the rubric in Section 4.2.

4.1. Brainstorming and sketching (RQ1)

The quantity of concepts brainstormed for each design problem at the pre- and post-intervention phases of the study was of interest because a high number of generated ideas may be an indicator of increased creativity (Shah, Smith, & Vargas-Hernandez 2003). In total, 104 concept lists were analyzed to determine the number of concepts that each participant brainstormed throughout the course of the design activity. When counting concepts, only the overarching concept was counted in cases in which sub-concepts or bulleted details were included. In addition, post-intervention cases in which participants wrote "same as Brainstorm 1" or "no new ideas" were not counted, although concepts that were repeated between the pre-intervention and post-intervention brainstorm were counted. For the grill problem, participants brainstormed an average of 6.250 concepts pre-intervention and 4.800 concepts post-intervention. For the laundry machine problem, participants developed an average of 5.250 concepts pre-intervention and 4.926 concepts post-intervention, as shown in Figure 2. However, neither the intervention (p = 0.235) nor the design problem (p = 0.629) had a statistically significant impact on the number of concepts generated, contradicting Hypotheses H1-A and H3-A.

In Step 2 of the study, participants were instructed to "include how a user would interact with the concept" when creating the sketches of their smart grill and laundry machine concepts. For analysis of the sketches, the presence of a user or person interacting with the smart grill or laundry machine was of particular interest due to findings from Makhlouf *et al.* (2023) that sketches containing drawings of people result in more consideration of social and physical use contexts. A total of 94 sketches were examined for indication of a visual depiction of a user interacting with the sketched concepts. Two independent judges categorized each sketch as either containing (example shown in Figure 3) or not containing (example shown in Figure 4) a visual representation of a user, achieving a sufficient Cohen's Kappa of 0.963. In total, 6 out of 48 (12.5%) pre-intervention sketches and

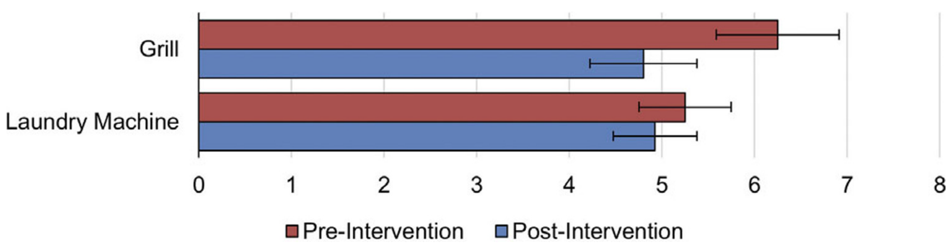

Figure 2. Average number of brainstormed concepts pre- and post-intervention for the smart grill and smart laundry machine design problems; error bars indicate ±1 SE.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

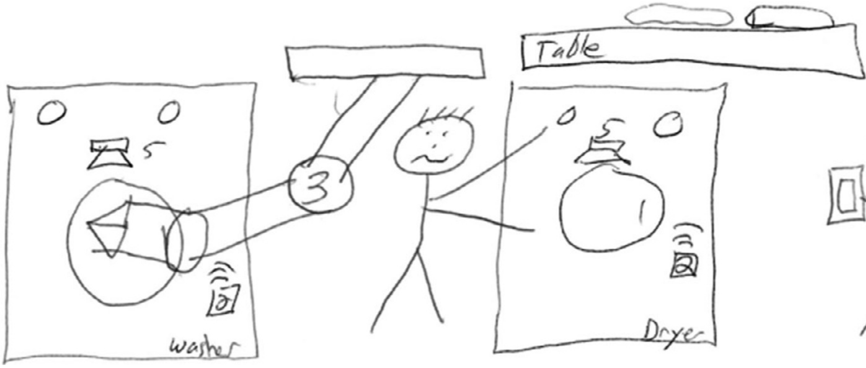

Figure 3. Post-intervention smart laundry machine sketch by Participant 11. Note the depiction of two representative users in the form of stick figures.

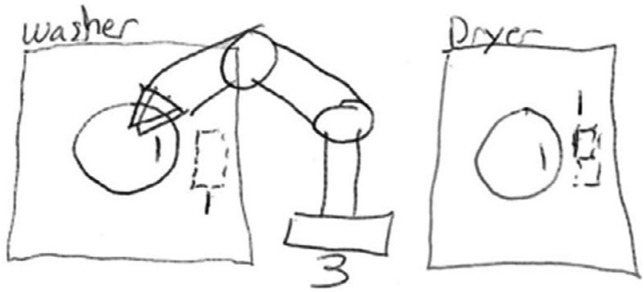

Figure 4. Pre-intervention smart laundry machine sketch by Participant 11. Note the focus of the sketch on the components of the machine and the lack of a depiction of a user.

10 out of 46 (21.7%) post-intervention sketches included some visual depiction of a user interacting with the concept, manifesting either as a stick figure or as a hand interacting with some component of the sketch. This increase in user depictions after the intervention supported Hypothesis H3-B; however, Pearson's chi-squared test found that this difference was not statistically significant ($\chi^2 = 1.420, p = 0.233$).

At the pre-intervention stage, 3 out of 26 (11.5%) grill concept sketches and 3 out of 22 (13.6%) laundry machine concept sketches featured a visual representation of a user, as shown in Figure 5. At the post-intervention stage, these numbers increased to 4 out of 22 (18.2%) and 6 out of 24 (25.0%) for the grill and laundry machine sketches, respectively. Although it was expected that there would be more user representations among the smart laundry machine sketches, these differences were insignificant, contradicting Hypothesis H1-B.

4.2. Market survey (RQ2 and RQ3)

Upon beginning the market survey, participants provided a written response to the questions "Who is the smart [grill/laundry machine] for? Who do you imagine would use it?" One judge utilized an inductive coding scheme (Boyatzis 1998) to develop a rubric of 18 recurring themes from the data, as detailed in Table 1.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

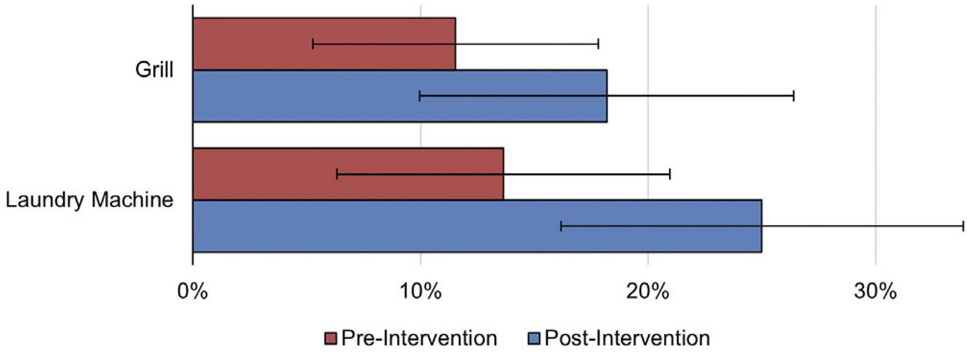

Figure 5. Percentage of sketches featuring a representation of a user pre- and post-intervention for the smart grill and smart laundry machine design problems; error bars indicate ±1 SE.

Table 1. List of themes and sample responses for the market survey question "Who is the smart [grill/laundry machine] for? Who do you imagine would use it?"

Code	Description of theme	Example response
1	Tech-savvy	"people who […] embrace technology"
2	Family	"family of 5–10 with a lot of laundry"
3	Heavy users, enthusiasts	"people who love to grill"
4	Lower class	"good for […] areas w/ limited resources"
5	Middle class, average	"regular US household, middle class family"
6	Upper class, luxury	"it's for rich people"
7	Professional, commercial use	"shared laundry spaces – laundromats, etc."
8	Men, fathers	"middle aged suburban dads"
9	Women, mothers	"slightly more female"
10	Homeowners, households	"it would mainly be used in household"
11	Everyone, anyone	"anyone who washes clothes"
12	People who are busy, looking for convenience, trying to save time	"people too busy to do laundry"
13	Novices, people looking for an easier experience	"people who […] do not grill regularly enough to internalize expertise"
14	People with disabilities	"those w/ impairments"
15	Social aspect	"people looking to entertain"
16	Me/self	"I would use it"
17	People who are environmentally conscious, looking for sustainability	"people who are energy and environmentally conscious"

Because of the number of possible codes, percent agreement was calculated for two independent judges rather than Cohen's Kappa, and the judges achieved sufficient agreement of 83.1% on the full data set. Participants wrote an average of 11.2 words (±0.7 SE) in response to the questions and appeared to be non-statistically

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

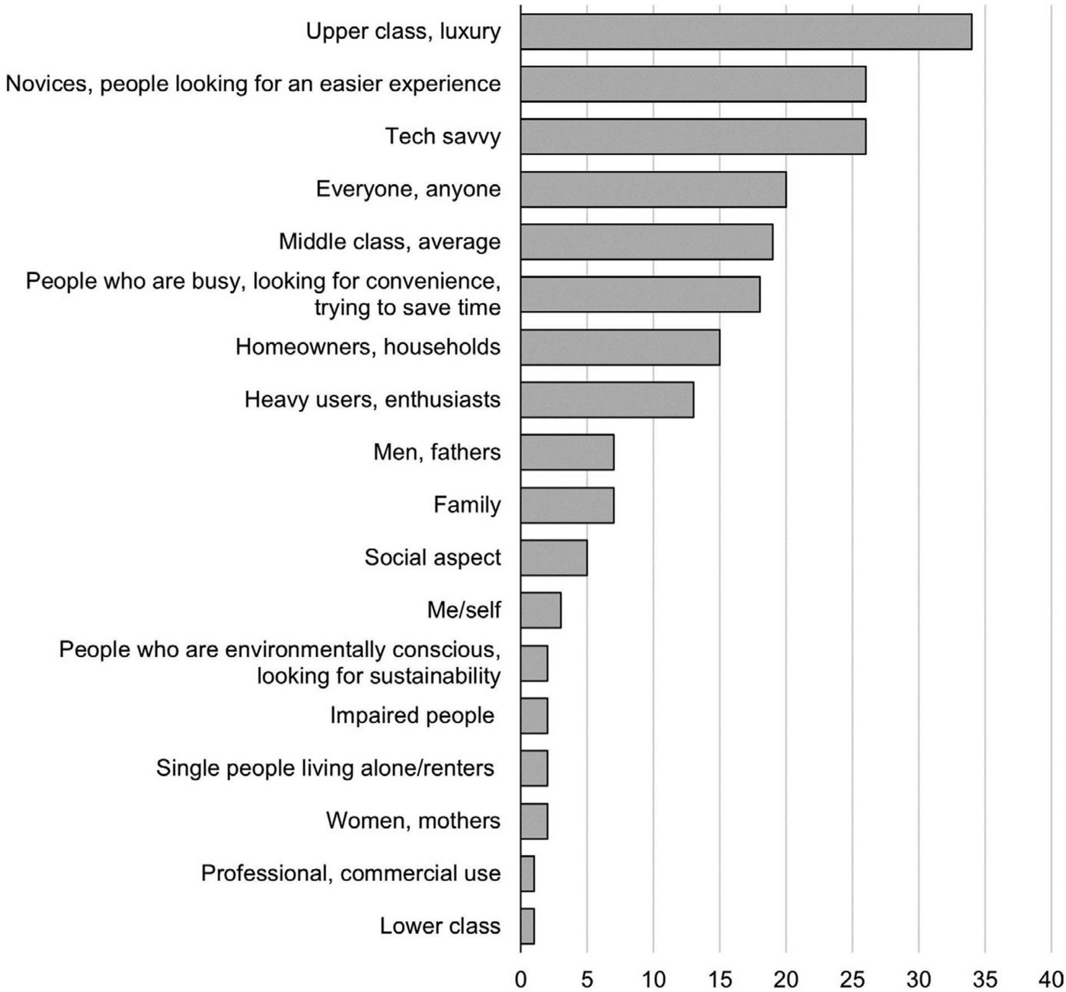

Figure 6. Number of occurrences of qualitative themes across 110 written responses to "Who is the smart [grill/laundry machine] for? Who do you imagine would use it?"

significantly less verbose post-intervention (p = 0.150), as the number of words written decreased from 12.4 (±1.1 SE) to 10.2 (±0.8 SE).

Figure 6 shows the breakdown of the number of occurrences of each theme across the responses. The most commonly referenced theme was the theme of an upper-class or "luxury" user. While smart appliances are typically associated with a higher price point, this finding may also be indicative of availability bias. All participants worked for a luxury appliance company, meaning that they may have applied the profile of the user that they typically encounter at work to this design problem. The number of appearances of this theme decreased from 20 to 14 after the bias mitigation intervention, indicating that the intervention was successful in reducing availability bias. Another theme with a reduction in appearance after the intervention was the idea of a "tech-savvy" user, which decreased from 16 to 10 appearances. Participants seemed to realize that targeting a tech-savvy user excluded many other potential users.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Another commonly occurring theme was that of a novice or inexperienced user who was looking for a smart appliance to make a task easier. While the number of participants who applied this theme to their description of the smart laundry machine user did not change, the use of this theme in relation to the smart grill user decreased from 12 to 4 instances after the intervention. It was somewhat unexpected that the Design Considerations Checklist would lead the designers to make fewer considerations for inexperienced users; however, this finding makes sense in the context of the accompanying decrease in usage of the "tech-savvy" theme, as a user who is neither tech-savvy nor skilled at using a grill would likely find the appliance undesirable or too difficult to use.

Some themes were only applied to either the grill or the laundry machine. For example, the smart grill was associated with social gatherings five times, while the smart laundry machine never was. Similarly, 12 responses associated the grill user with an enthusiast or heavy user, while only 1 response made the same connection to the laundry machine user. Rather, the smart laundry machine was more often associated with homeowners and busy people more often than the smart grill, perhaps because laundry was viewed as a more integral chore for running a household compared to grilling.

Figure 7 contains a more detailed view of how some themes were applied at the different stages of the study and to the different design problems. One notable result is that, as expected, participants appeared to apply gender-stereotyping to their belief of who would use the products. Seven responses referenced "men" or "fathers" as users of the smart grill, and two responses referenced "women" or "mothers" as users of the smart laundry machine, but each of these themes were never applied in relation to the other appliance. One participant playfully mused that she was designing a "mojo dojo casa grill," referencing the satirical description of toxic masculinity from the 2023 film Barbie (2023) and indicating a perceived lack of belonging to the target user group for the smart grill.

Of particular note is the notion that "anyone" or "everyone" could be a potential user of the product. In both the smart grill and smart laundry machine cases, four additional participants referenced "everyone/anyone" as a target user for their product after the intervention. While participants may have intended to indicate that their product was suitable for a wide range of users, Makhlouf *et al.* (2023)

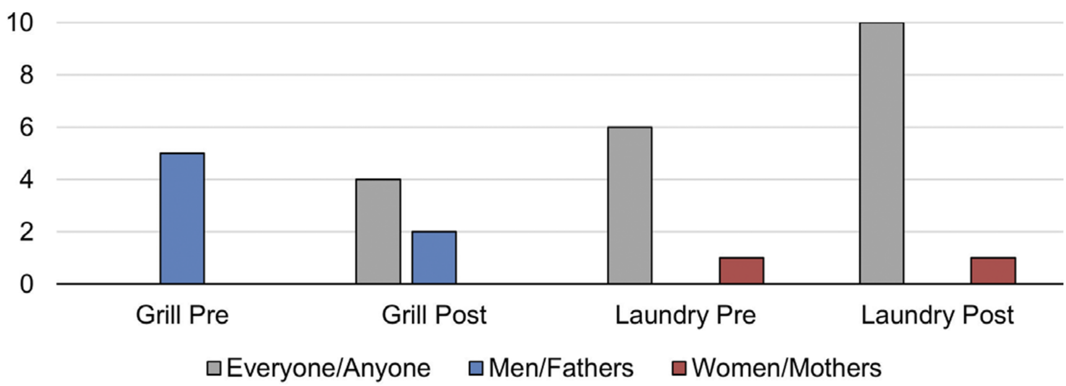

Figure 7. Number of occurrences of select themes broken out by design problem and intervention stage (pre-intervention or post-intervention).

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

suggest that designers' claims of a "general user" may indicate that they are designing for an "average" user, potentially leading to androcentrism, or a default assumption that a person of unspecified gender is a man (Schauer *et al.* 2024), in addition to a lack of consideration for diverse users' specific needs. The Design Considerations Checklist attempted to encourage reflection beyond a "general user" by asking questions such as: "Are there any people who are unlikely to benefit from this design? Are there any people or groups of people who this design would not work for?" However, richer detail in both bias mitigation and data collection may be needed to understand the intentions of participants who claim that their designs are suitable for "everyone." Accompanied by the decrease in occurrence of mentions of a "luxury" or "tech savvy" user, this result indicates that the items on the checklist were successful in helping participants design for a wider range of potential users. A checklist similar to this could be introduced as part of engineering design education alongside similar cognitive forcing/checklisting techniques, such as SCAMPER (Boonpracha 2023).

The prices generated by participants when asked to imagine how much each of their concepts would be sold for were used to understand the value that the participants placed on each of their concepts. As a result, responses to the market survey question initially ranged from \$60 to \$58,000, potentially because the participants were employed as engineers and designers and likely did not have a firm understanding of the business- and marketing-related aspects of the products. The method for omitting outliers proposed by Clark-Carter (1997), in which values greater than three sample standard deviations from the sample mean were eliminated, was applied to the data for analysis. From the pre-intervention to the post-intervention market survey, the monetary value that participants assigned to their grill designs decreased non-statistically significantly (p = 0.559) from \$2,905.96 (±\$517.75 SE) to \$2,225.00 (±\$376.26 SE). Conversely, the monetary value placed upon participants' laundry machine designs increased non-statistically significantly (p = 0.383) from \$2,511.25 (±\$369.27 SE) to \$3,232.90 (±\$483.46 SE), as shown in Figure 8. The higher value assigned to the grill compared to the laundry machine at the pre-intervention stage supported Hypothesis H2-A; however, the difference in the prices assigned to the smart grill and smart laundry machine was

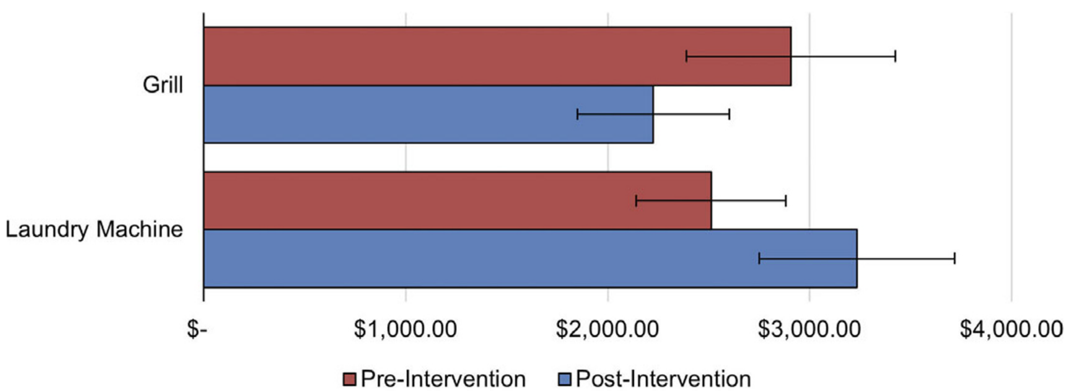

Figure 8. Average responses in USD to the market survey question "To the nearest \$100, how much do you think the smart [grill/laundry machine] would be sold for?" pre- and post-intervention for the smart grill and smart laundry machine design problems; error bars indicate ±1 SE.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

not significantly different at either the pre-intervention (p = 1.000) or post-intervention (p = 0.136) market surveys.

The next part of the market survey was the Likert scale responses indicating participants' level of agreement with various statements about their sketched concepts. For each of the statements, an ANOVA model was created to examine the impact of (1) the pre- or post-intervention stage and (2) the smart grill or smart laundry machine problem statement, as summarized in Figures 9 and 10. Results showed that participants agreed more strongly with the statement "The market for

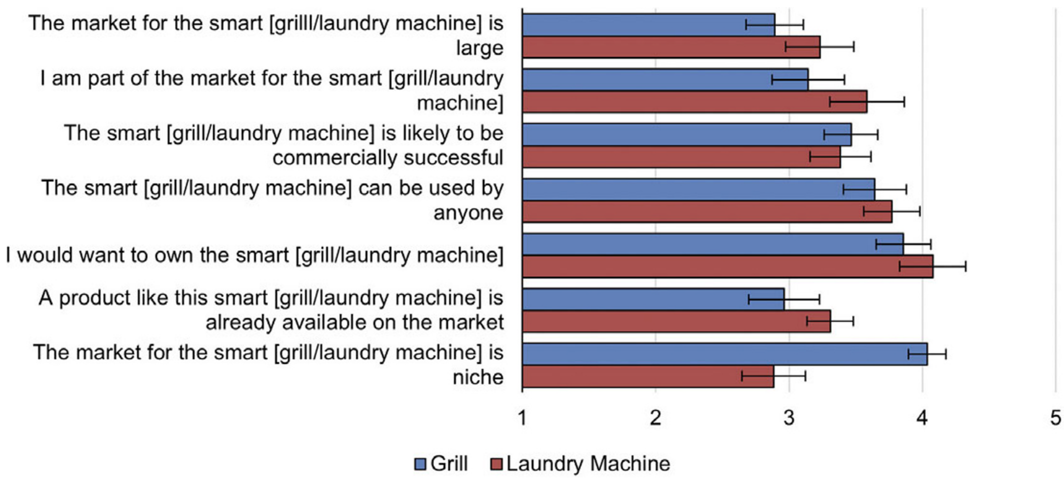

Figure 9. Pre-intervention comparison of participants' agreement (1 – Strongly disagree, 2 – Somewhat disagree, 3 – Neutral, 4 – Somewhat agree, 5 – Strongly agree) with statements about the market for sketched smart grill and smart laundry machine concepts; error bars indicate ±1 SE.

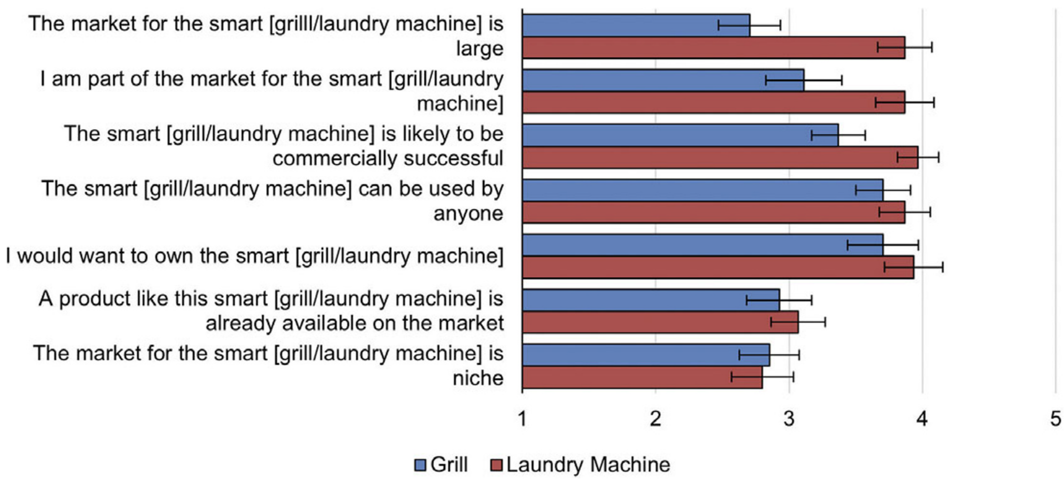

Figure 10. Post-intervention comparison of participants' agreement (1 – Strongly disagree, 2 – Somewhat disagree, 3 – Neutral, 4 – Somewhat agree, 5 – Strongly agree) with statements about the market for sketched smart grill and smart laundry machine concepts; error bars indicate ±1 SE.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

the smart grill is niche" compared to "The market for the smart laundry machine is niche" at both the pre-intervention ($p = 0.003$) and post-intervention stages ($p < 0.001$). As discussed by Fillingim *et al.* (2022)), the reference to the market as "niche" could be interpreted as a negative framing of the market, reflecting negative feelings toward the grill concepts, which contradicted Hypothesis H2-A.

However, other statements presented a more positive framing of the market: "The market for the smart [grill/laundry machine] is large," and "The smart [grill/laundry machine] is likely to be commercially successful." For each of these statements, there was no significant difference when comparing participants' agreement with the statements pertaining to the smart grill and smart laundry machine at the pre-intervention stage ($p = 0.288$ and 0.956, respectively), as shown in Figure 9. However, at the post-intervention stage, as shown in Figure 10, participants agreed more strongly with the statements when they were applied to the smart laundry machine concept compared to the smart grill concept ($p = 0.001$ and 0.026, respectively). Similarly, at the pre-intervention stage, there was no statistically significant difference in the responses to "I am part of the market for the grill" and "I am part of the market for the smart laundry machine" ($p = 0.250$); however, after the intervention, participants identified more strongly with the smart laundry machine market compared to the smart grill market ($p = 0.045$). These three statements with which participants agreed more strongly post-intervention for the smart laundry machine compared to the smart grill can all be interpreted as being positively framed, indicating that participants took on a more positive view of the smart laundry machine, or a more negative view of the smart grill, after the intervention, despite the lack of a statistically significant increase or decrease. This somewhat supported Hypothesis H3-C, as it was predicted that the intervention would lead to an increase in favorable sentiments toward the laundry machine problem. The Design Considerations Checklist focused heavily on encouraging participants to reflect on user groups that would or would not be able to access their designs, so it is possible that the intervention was successful in reducing availability bias that caused participants to initially misjudge the potential user groups for their smart grill (overestimated) and smart laundry machine (underestimated) designs.

Kendall's tau-b correlation coefficient (1938) was calculated for every combination of questions from the market survey to determine if there were any strong positive or negative correlations between participants' Likert scale responses. For example, a significant positive correlation was found for all experimental cases between "The market for the smart [grill/laundry machine] is large" and "The smart [grill/laundry machine] is likely to be commercially successful" ($\tau = 0.548$, $p < 0.001$). This finding also validates the notion that participants may have viewed the "large market" statement as positively framed, since it was positively associated with commercial success. Similarly, there was a significant positive correlation for all experimental cases between "I am part of the market for the smart [grill/laundry machine]" and "I would want to own the smart [grill/laundry machine]" ($\tau = 0.576$, $p < 0.001$).

Prior investigation into cognitive bias in design by Fillingim *et al.* (2022) proposed that correlation between believing that one is part of a market for a product and believing that the market for a product is large may be indicative of availability bias. With that in mind, correlations between responses to the market survey statements "I am part of the market for the smart [grill/laundry machine]" and “The market for the smart [grill/laundry machine] is large” were checked. While there was no significant correlation between responses to these statements pertaining to the smart grill ($\tau = 0.179$, $p = 0.113$), there was a significant positive correlation between responses to the statements regarding the smart laundry machine at both the pre-intervention ($\tau = 0.511$, $p = 0.002$) and post-intervention ($\tau = 0.391$, $p = 0.014$) stages. At the post-intervention stage, there was also a significant positive correlation between “I am part of the market for the smart laundry machine” and “The smart laundry machine is likely to be commercially successful” ($\tau = 0.445$, $p = 0.006$).

Similarly, there was a significant *negative* correlation between “I am part of the market for the smart laundry machine” and “The market for the smart laundry machine is niche” at both the pre-intervention ($\tau = -0.420$, $p = 0.012$) and post-intervention ($\tau = -0.317$, $p = 0.042$) stages, while there was no significant correlation between responses to these statements pertaining to the smart grill ($\tau = 0.037$, $p = 0.750$). In other words, participants were more likely to think that the market for the smart laundry machine was large if they personally felt that they were part of the target market for it, while having this sense of belonging to the market did not impact how large the participants felt the market for the smart grill was. This finding is suggestive of availability bias displayed by participants toward the smart laundry machine design problem, as their interpretation of the smart laundry machine market may have been driven by whether they personally identified with it, indicating that participants projected their own experience or desire for a product onto the larger group of potential users. However, the grill was viewed as marketable to a large number of people, regardless of whether designers personally felt as if they were the target market. Also, the bias mitigation intervention did not impact the availability bias exhibited by the designers toward the smart laundry machine, as this positive correlation persisted in the post-intervention market survey. These findings partially support Hypothesis H2-B, as it was expected that these correlations would be present for both the grill and laundry machine concepts based on the predicted effect of availability bias.

A variety of confounding factors may have impacted these correlations, as well as the unexpected lack of availability bias exhibited toward the grill. A laundry machine could be viewed as more “necessary” or “valuable” to a household compared to a grill, as the alternatives to a household laundry machine are offsite laundry or tediously handwashing clothes, whereas the alternative to an outdoor grill is an indoor stovetop. All 30 participants lived in a household with a laundry machine, while 24 out of 30 participants owned a grill. Participants who owned both also reported that their household used the laundry machine around four times more often than they used the grill on a weekly basis.

The feminine-stereotyping of laundry chores compared to the masculine-stereotyping of grilling chores may have impacted the observed correlations; perhaps participants’ view of the size of the grill market was not impacted by their personal level of interest in the product because they subconsciously recognized the inherent value of the product in accordance with the tendency to ascribe higher value to masculine-stereotyped tasks or traits (Eveleth 2013; Bailey, LaFrance, & Dovidio 2019). Also, participants may have felt that the “smart” aspect of the laundry machine was unnecessary, especially since most participants were not the majority user of their household’s laundry machine: the average participant estimated that of the time their household spent doing laundry, they personally

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

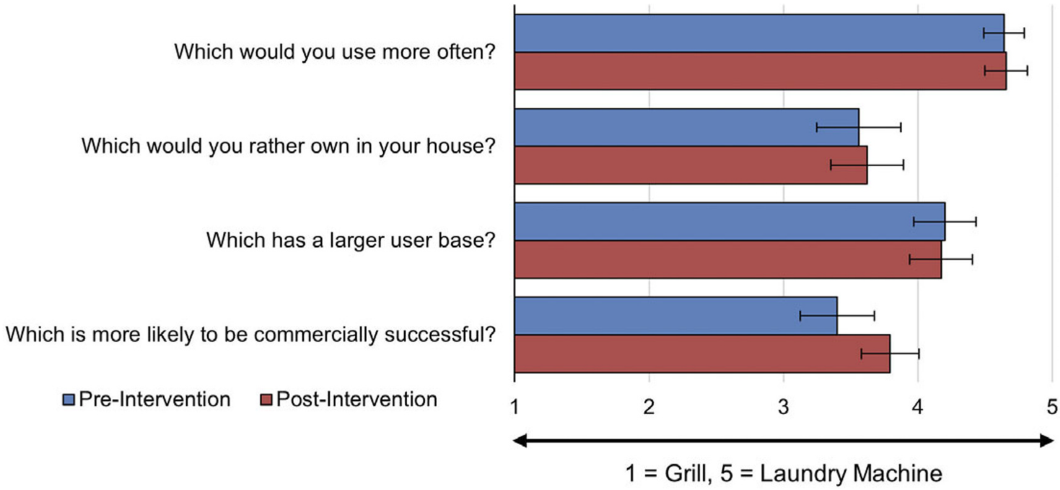

Figure 11. Participant perceptions of statements comparing their sketched concepts for the smart grill and smart laundry machine problems; error bars indicate ±1 SE.

did laundry 38.4% (±6.0% SE) of the time. Conversely, participants spent 65.0% (±7.9% SE) of household time grilling, which was significantly more than the percent of time participants spent doing laundry ($p < 0.001$). There was also a significant negative correlation between participant age and percent of household time spent doing laundry ($\tau = -0.466$, $p = 0.001$), indicating that older men participants may have been more likely to adhere to stereotypical gender roles in household chore assignment.

In the final section of the market survey, participants directly compared their perception of their smart grill and smart laundry machine concepts. The means of these responses were compared at both the pre-intervention and post-intervention stages to determine whether the intervention impacted the relative value that participants assigned to each concept. As shown in Figure 11, the intervention did not have a statistically significant impact on which design participants felt they would use more often ($p = 0.783$), which they would rather own ($p = 0.928$), which has a larger user base ($p = 0.930$) or which was more likely to be commercially successful ($p = 0.284$). It is likely that the factors described in previous paragraphs somewhat "cemented" participants' views of each item, and as the Design Considerations Checklist was mostly focused on encouraging participants to consider their own experiences and assumptions in the context of the design problem rather than the design problem itself, participants did not significantly change the way they valued or perceived either of the items.

Participants displayed an unexpected preference toward the laundry machine for all four questions, contradicting Hypothesis H2-C. As shown in Figures 12 and 13, a minimum of 60% of participants responded that they "probably" or "definitely" preferred the laundry machine for each of the four questions at both the pre-intervention and post-intervention stages. As previously discussed, it is possible that participants recognized the value of the laundry machine for their household despite not personally interacting with it as much, leading them to display a preference toward it in positively framed questions.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

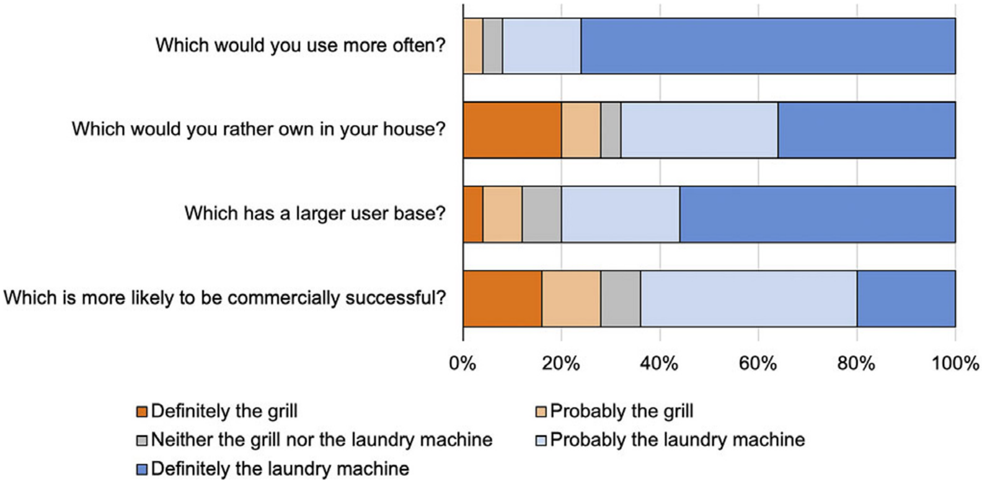

Figure 12. Direct comparison of participants' sketched smart grill and smart laundry machine concepts, pre-intervention.

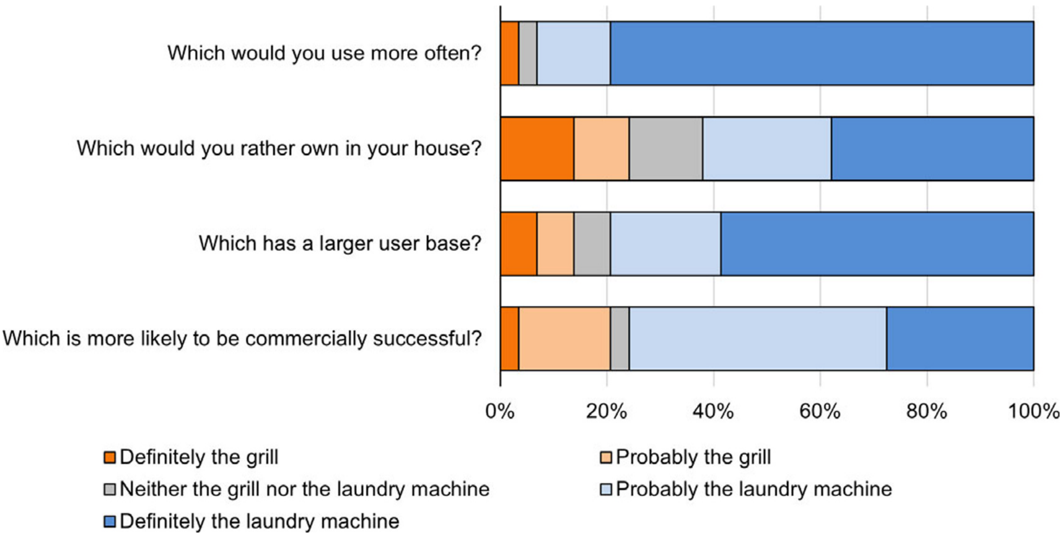

Figure 13. Direct comparison of participants' sketched smart grill and smart laundry machine concepts, post-intervention.

4.3. Reflections and post-survey (RQ3)

The post-survey investigated participants' typical interactions with grills and laundry machines. In addition to reporting how often their household used each item and whether or not they personally used each item, participants also reported how much they enjoyed using each item. On a scale from 1 to 5 (1 – strongly dislike, 2 – somewhat dislike, 3 – neither like nor dislike, 4 – somewhat like, 5 – strongly like), participants reported that they enjoy grilling (mean = 3.750±0.190 SE) significantly more than they enjoy doing laundry (mean = 3.150±0.154 SE, p = 0.009). At the pre-intervention stage, there was a significant positive correlation

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

between how much participants reported enjoying using a laundry machine and their agreement with the statement "The smart laundry machine is likely to be commercially successful" ($\tau = 0.349$, $p = 0.042$), although this correlation was no longer significant after the intervention ($\tau = 0.151$, $p = 0.353$). Although findings from the market survey did not support Hypothesis H3-C, this finding suggests that the intervention reduced availability bias in participants as it pertains to their belief that the smart laundry machine would be commercially successful, lending support to the hypothesis.

Participants used a 1–5 Likert scale (1 – extremely unhelpful, 2 – somewhat unhelpful, 3 – neither helpful nor unhelpful, 4 – somewhat helpful, 5 – extremely helpful) to describe how helpful they found the Design Considerations Checklist. As shown in Figure 14, over 70% of participants stated that the Design Considerations Checklist was "somewhat" or "extremely" helpful in reflecting on their designs from the pre-intervention stages, while over 60% stated that the checklist was "somewhat" or "extremely" helpful as they developed new designs post-intervention. This slight difference was somewhat expected, as the instructions for the Design Considerations Checklist explicitly instructed participants to "reflect on each question as it pertains to each of your two sketch concepts," and made no mention of keeping the considerations in mind throughout the remainder of the activity. Notably, less than 10% of participants indicated that the Design Considerations Checklist was "somewhat" or "extremely" *un*helpful during the design activities, highlighting its promise as a decision-making tool.

Finally, participants provided a written response to the question "How did the checklist impact the way you approached the design problems and questions in the second round?" One judge utilized an inductive coding scheme to develop a rubric of 11 recurring themes from the data (24 coherent responses), as shown in Table 2. Because of the number of possible codes, percent agreement was calculated for two independent judges rather than Cohen's Kappa, and the judges achieved sufficient agreement of 80% on the full data set.

In addition to some responses pertaining to output during the brainstorming and sketching sessions (some participants mentioned that they felt they did not

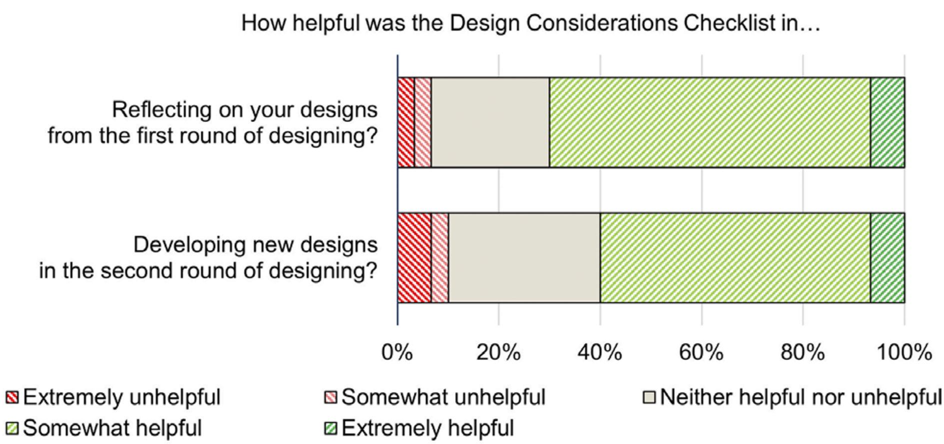

Figure 14. Participants' perceptions of the helpfulness of the Design Considerations Checklist intervention during the two phases of the study.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Table 2. List of themes and sample responses for the post-survey question "How did the checklist impact the way you approached the design problems and questions in the second round?"

Code	Description of theme	Example response
1	No change, marginal change	"couldn't form any new/different ideas"
2	Worked slower, fewer ideas	"made me slow down in my design, not get swept up in to my own thoughts"
3	Worked faster, more ideas	"the 2nd time which is very easy and ideas are coming very soon"
4	Thought more about the user or customer	"thought a bit more about the potential customer"
4.1	Considered a wider variety, diversity of users	"I tried to make both appliances usable by a larger group"
4.2	Considered users' ability to use or access	"modified design to allow people without smartphone or internet access to use them manually"
4.3	Considered users besides myself	"much more considerate of alternate needs beyond those similar to my own"
4.4	Identified a target user or target user group	"identified my target audience"
5	Thought more about the product	"impacted my design"
5.1	Considered feasibility, usefulness, realizability of product	"helped bring further things to mind regarding usefulness of the product"
5.2	Considered cost of product	"thought about cost a bit more"

change their designs much, or that they brainstormed fewer ideas than they did pre-intervention), two main overarching themes emerged: participants mostly felt that the Design Considerations Checklist led them to either consider the potential user or customer more (13 responses), or led them to consider the actual product they had designed more (6 responses). This breakdown lends additional support to earlier speculation that the Design Considerations Checklist was more effective in helping participants to consider their perception of the user rather than their perception of the design problem or the value they placed on the subject of the design problem.

Figure 15 shows a breakdown of the number of occurrences of each theme across the responses. While a plurality of participants stated that the way they approached the design problem changed very little, if at all, many responses referenced making considerations for a wider variety of users or recognized that they were targeting a particular group. Interestingly, four responses specifically referenced making considerations for users other than themselves – in other words, participants who had previously applied availability bias to their work by primarily considering their own needs or preferences made a conscious effort to think of what other users would want. For example, one participant stated that "I tried to determine what other users may want rather than what I would want" in the second

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

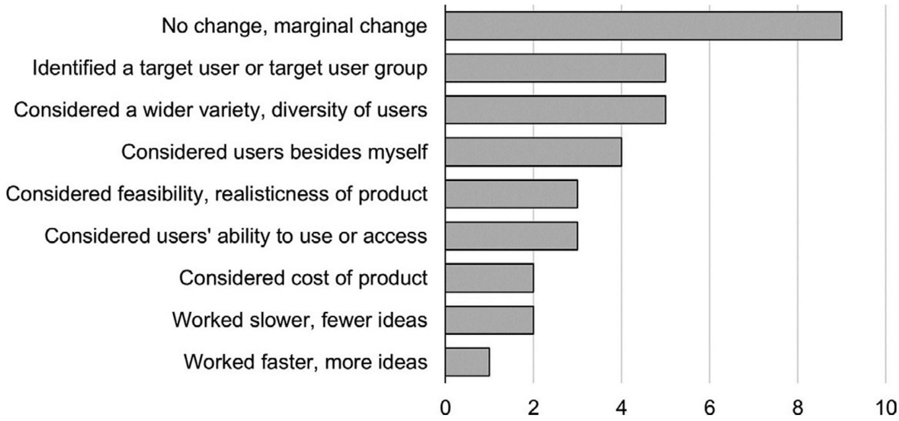

Figure 15. Number of occurrences of qualitative themes across 24 written responses to "How did the checklist impact the way you approached the design problems and questions in the second round?"

round of designing, showing that the intervention successfully led participants to reflect on their availability bias and take steps to avoid it.

5. Conclusions

5.1. Limitations and future work

One disadvantage of a within-subjects experimental design is that participants may grow tired of the data collection activities on the second time through the procedure, as evidenced by the shortening of written responses and some participants' refusal to participate in a second round of brainstorming or sketching. Although 30 participants provided a satisfactory sample size, especially considering that the participants were professional engineers and designers, this sample size limited the study design to within-subjects. A larger-scale between-subjects study would mitigate the collection of low-effort data from participants and would also enable the comparison of multiple types of bias mitigation interventions, such as checklists, user personas in various formats and simulations. Given additional time with the participants, future work should also include greater emphasis and analysis of the participants' interactions with the bias mitigation tools. For example, requiring participants to provide a written response to each checklist item, or audio-recording verbal responses, would allow the researchers to analyze common themes in responses and connect participants' reflections to design features in their sketches. These detailed data obtained from highly engaged participants have the potential to provide greater insights and findings that are more closely aligned with expectations from existing literature.

In a larger study such as the proposed between-subjects study, the efficacy of various interventions could be assessed in addition to the cognitive load imposed by each, which is a relevant concern in ensuring user comfort during decision-making tasks (Sibbald *et al.* 2013). Another limitation of the study design was that due to time and scheduling constraints, researchers' interactions with the participants were limited to a single session. A longitudinal study format with continuous access to participants would allow researchers to study the longer-term efficacy of

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

the bias mitigation intervention to understand the effects of temporal distance from the intervention.

When recruiting participants for future work, there are several additional considerations to be made. First, participants in this study, despite being employed as engineers or designers, came from a variety of educational backgrounds, as detailed in Section 3.2. As a result, some participants had little to no formal training on human-centered design principles, while others had spent several years obtaining a degree in closely related fields. While these diverse backgrounds can be representative of the makeup of engineering teams in industry, results of this study would hold more validity if these confounding factors were controlled by recruiting participants from more similar backgrounds. Comparing these results to a similar study conducted with novice engineers or designers would afford additional insight into the role of education on bias mitigation in design. In addition, the majority of participants were men, and all participants were employed at a kitchen appliance company who had chosen to participate in a half-day design cognition workshop. As a result of the lack of gender diversity, it was not possible to draw conclusions related to availability bias that may have been a result of gender-stereotyped household chores. Participants were also limited to those who had an existing interest in design cognition, as participation in the workshop was voluntary and occurred during working hours, meaning that results are not necessarily representative of the general engineering design workforce. Although participants likely had varying levels of prior familiarity with cognitive bias, all participants were given the same background information on different types of bias (excluding availability bias) in design contexts directly before the design activity.

Variations in the way participants interpreted Likert-scale statements throughout the market survey, given a lack of definition or context, may have led to confusion among participants. For example, the response to statements such as: "The smart [grill/laundry machine] is likely to be commercially successful" or "The smart [grill/laundry machine] can be used by anyone" may vary between participants depending on their interpretation of the words "initiative" or "anyone." Framing bias may have also played a role in the way participants responded to questions, as identified previously. To mitigate this, care was taken to include both positively and negatively framed statements. Similarly, one main finding from the qualitative analysis of the market survey was that participants moved toward creating designs that were for "everyone" or "anyone." While this may be a positive indication that participants applied principles of universal design in their work, Makhlouf *et al.* (2023) view this as a negative finding indicating that participants were designing for a non-existent "average" user. More in-depth data collection methods, such as focus groups or short interviews, would build a deeper understanding of designers' intentions and thought processes when developing a design "for everyone." Future directions for research may involve developing and validating guided reflection tools for encouraging designers to consider the full range of diverse users' needs, rather than a general user.

5.2. Contributions

The two main contributions of this work are (1) a mixed-methods comparison of how availability bias manifests in a design concept development activity with professional engineers and designers and (2) a novel application of a checklist-

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

based bias mitigation intervention targeted at reducing availability bias and building empathy for users. Based on the analysis discussed above, this paper explored the following research questions, filling established gaps in literature:

RQ1: How do indicators of availability bias manifest through design outcomes in a brainstorming and sketching design activity?

The number of brainstormed concepts was examined to understand how availability bias may have manifested in the flow of brainstormed ideas. In contradiction to Hypothesis H1-A, it was found that neither the intervention nor the design problem ($p = 0.629$) had a statistically significant impact on the number of concepts generated. The sketches were also examined to identify whether participants had included a representation of a user in their sketches, due to correlations between visual inclusion of a user and higher consideration for user needs and contexts (Makhlouf *et al.* 2023). Although 3 out of 26 grill sketches and 3 out of 22 laundry machine sketches at the pre-intervention stage, and 4 out of 22 grill sketches and 6 out of 24 laundry machine sketches at the post-intervention stage featured a depiction of a user, these differences were not statistically significant, contradicting Hypothesis H1-B.

RQ2: How do indicators of availability bias manifest through designers' perceptions of the market for a product and the value of a product?

After brainstorming and sketching, participants filled out a market survey in which they estimated a price that their smart grill and smart laundry concepts would be sold for, then responded to a series of statements regarding the potential market for each concept. Contradicting Hypothesis H2-A, there were no significant differences between the average prices assigned to the smart grill and smart laundry machine. Participants were also more likely to agree with negatively framed statements about the smart grill and positively framed statements about the smart laundry machine, which was unexpected and contradicted Hypothesis H2-A. Similarly, participants displayed a preference toward the smart laundry machine in a series of four head-to-head comparison questions, contradicting Hypothesis H2-C. Although many of the participants were not the primary users of their household's laundry machine, it is probable that these results are indicative of their recognition that the laundry chore is essential to their household. Finally, checking for correlations between responses to survey questions, such as "I am part of the market for the smart [grill/laundry machine]" and "The market for the smart [grill/laundry machine] is large," revealed that a positive correlation was present for the statements pertaining to the smart laundry machine, but not the smart grill. This indicated that participants exhibited availability bias toward the smart laundry machine problem but not toward the smart grill problem, partially supporting Hypothesis H2-B.

RQ3: How does a checklist-style bias mitigation intervention impact manifestations of availability bias and considerations made for a diverse collection of users?

Participants found the bias intervention especially helpful in reflecting on designs they developed during the first brainstorming and sketching phase; however, Hypothesis H3-A was not supported by the finding that the number of brainstormed concepts did not significantly change from pre-intervention to post-intervention. Similarly, although the number of sketches containing a visual

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

representation of a user increased from 12.5% to 21.7%, this difference was not significantly significant, failing to support Hypothesis H3-B. The intervention did not impact the scores of the head-to-head comparison of the product markets, and also did not impact the correlations discussed in RQ2, as significant positive correlations persisted post-intervention. These findings contradict Hypothesis H3-C; however, the intervention was successful in eliminating a positive correlation between participants' enjoyment of doing laundry machines and their response to "The smart laundry machine is likely to be commercially successful," which is indicative that availability bias was present and then mitigated. Qualitative analysis of participants' reflections indicated that the intervention primarily assisted designers in making additional considerations for users, such as increasing accessibility and building awareness of excluded user groups.

Financial support

This material is based upon the work supported by the National Science Foundation Graduate Research Fellowship under grant number DGE-2039655. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Competing interests

The authors declare none.

References

Acar, B. S., **Edwards, A. M.** & **Aldah, M.** 2018 Correct use of three-point seatbelt by pregnant occupants. *Safety* **4** (1), 1; doi:10.3390/safety4010001.

Agyemang, M., **Andreae, D. A.** & **McComb, C.** 2023 Uncovering potential bias in engineering design: A comparative review of bias research in medicine. *Design Science* **9**, e17; doi:10.1017/dsj.2023.17.

Applegate 2015 Girls get their grill on: Survey finds men think they're the masters, but women like to spice it up. *PR Newswire*. https://www.prnewswire.com/news-releases/girls-get-their-grill-on-survey-finds-men-think-theyre-the-masters-but-women-like-to-spice-it-up-300084613.html (Accessed: 16 March 2024).

Arkes, H. R. & **Ayton, P.** 1999 The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin* **125** (5), 591–600; doi:10.1037/0033-2909.125.5.591.

Arkes, H. R. & **Blumer, C.** 1985 The psychology of sunk cost. *Organizational Behavior and Human Decision Processes* **35** (1), 124–140; doi:10.1016/0749-5978(85)90049-4.

Armstrong, A. G., *et al.* 2024 Social impact evaluation analysis: A systematic qualitative method for assessing engineered products. *Design Science* **10** (45), 1–20; doi:10.1017/dsj.2024.31.

Babcock, L. & **Loewenstein, G.** 1997 Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives* **11** (1), 109–126; doi:10.1257/jep.11.1.109.

Bailey, A. H., **LaFrance, M.** & **Dovidio, J. F.** 2019 Is man the measure of all things? A social cognitive account of Androcentrism. *Personality and Social Psychology Review* **23** (4), 307–331; doi:10.1177/1088868318782848.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Bailis, R. *et al.* 2015 *Demand for Nontraditional Cookstoves in Bangladesh*. Innovations for Poverty Action. http://www.poverty-action.org/study/demand-nontraditional-cookstoves-bangladesh (Accessed: 14 October 2022).

Barbie 2023 Warner Bros. Pictures.

Barton, M. *et al.* 2016 The use of theory in designing a serious game for the reduction of cognitive biases. *Transactions of the Digital Games Research Association* **2** (3). http://todigra.org/index.php/todigra/article/view/53 (Accessed: 3 April 2024).

Boffi, L. *et al.* 2014 Supporting the designers to build empathy with people with Parkinson's disease: The role of a hand tremor simulating device and of user research with end-users. *DRS Biennial Conference Series [Preprint]*. https://dl.designresearchsociety.org/drs-conference-papers/drs2014/researchpapers/94.

Boonpracha, J. 2023 SCAMPER for creativity of students' creative idea creation in product design. *Thinking Skills and Creativity* **48**, 101282; doi:10.1016/j.tsc.2023.101282.

Bose, D., **Segui-Gomez, M.** & **Crandall, J. R.** 2011 Vulnerability of female drivers involved in motor vehicle crashes: An analysis of US population at risk. *American Journal of Public Health* **101** (12), 2368–2373; doi:10.2105/AJPH.2011.300275.

Boyatzis, R. E. 1998 *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage Publications, Inc.

Brandstätter, E., **Gigerenzer, G.** & **Hertwig, R.** 2006 The priority heuristic: Making choices without trade-offs. *Psychological Review* **113** (2), 409–432; doi:10.1037/0033-295X.113.2.409.

Brenan, M. 2020 Women still handle Main household tasks in U.S., *Gallup.com*. https://news.gallup.com/poll/283979/women-handle-main-household-tasks.aspx (Accessed: 16 March 2024).

Brown, S. J., *et al.* 1992 Survivorship bias in performance studies. *The Review of Financial Studies* **5** (4), 553–580; doi:10.1093/rfs/5.4.553.

Burnett, M., **Peters, A.**, *et al.* 2016 Finding gender-inclusiveness software issues with GenderMag: A field investigation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2586–2598. Association for Computing Machinery (CHI '16); doi:10.1145/2858036.2858274.

Burnett, M., **Stumpf, S.**, *et al.* 2016 GenderMag: A method for evaluating software's gender inclusiveness. *Interacting with Computers* **28** (6), 760–787; doi:10.1093/iwc/iwv046.

Cao, K.-K., *et al.* 2016 Raising awareness in model-based energy scenario studies—A transparency checklist. *Energy, Sustainability and Society* **6** (1), 28; doi:10.1186/s13705-016-0090-z.

Carmel-Gilfilen, C. & **Portillo, M.** 2016 Designing with empathy: Humanizing narratives for inspired healthcare experiences. *HERD: Health Environments Research & Design Journal* **9** (2), 130–146; doi:10.1177/1937586715592633.

Carnes, M., *et al.* 2015 Effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine: Journal of the Association of American Medical Colleges* **90** (2), 221–230; doi:10.1097/ACM.0000000000000552.

Chang-Arana, Á. M., *et al.* 2022 Under the umbrella: Components of empathy in psychology and design. *Design Science* **8**, e20; doi:10.1017/dsj.2022.13.

Chen, R. C. C., **Nivala, W. C.-Y.** & **Chen, C.-B.** 2011 Modeling the role of empathic design engaged personas: An emotional design approach. In *Universal Access in Human-Computer Interaction. Users Diversity* (ed. C. Stephanidis), pp. 22–31. Springer (Lecture Notes in Computer Science); doi:10.1007/978-3-642-21663-3_3.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Chew, K. S., **Durning, S. J.** & **van Merriënboer, J. J.** 2016 Teaching metacognition in clinical decision-making using a novel mnemonic checklist: An exploratory study. *Singapore Medical Journal* **57** (12), 694–700; doi:10.11622/smedj.2016015.

Chrysikou, E. G. & **Weisberg, R. W.** 2005 Following the wrong footsteps: Fixation effects of pictorial examples in a design problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **31** (5), 1134–1148; doi:10.1037/0278-7393.31.5.1134.

Clark-Carter, D. 1997 *Doing Quantitative Psychological Research: From Design to Report*, pp. xxii–666. Psychology Press/Erlbaum (UK) Taylor & Francis (Doing Quantitative Psychological Research: From Design to Report).

Dieter, G. E. & **Schmidt, L. C.** 2013 *Engineering Design*, 5th Edn. McGraw-Hill.

Draws, T. *et al.* 2021 A checklist to combat cognitive biases in crowdsourcing. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, **9**, 48–59; doi:10.1609/hcomp.v9i1.18939.

Dunbar, N. E., *et al.* 2017 Mitigation of cognitive bias with a serious game: Two experiments testing feedback timing and source. *International Journal of Game-Based Learning (IJGBL)* **7** (4), 86–100; doi:10.4018/IJGBL.2017100105.

Dunn, O. J. 1964 Multiple comparisons using rank sums. *Technometrics* **6** (3), 241–252; doi:10.1080/00401706.1964.10490181.

'**Embrace**' (n.d.) Design for Extreme Affordability. https://extreme.stanford.edu/projects/embrace/ (Accessed: 14 October 2022).

Emmons, D. L., *et al.* 2018 Mitigating cognitive biases in risk identification: practitioner checklist for the aerospace sector. *Defense Acquisition Research Journal* **25** (1), 52–93; doi:10.22594/dau.16-770.25.01.

Engineering Design Centre, University of Cambridge 2024 Inclusive design toolkit. https://www.inclusivedesigntoolkit.com/ (Accessed: 17 January 2025).

Eveleth, R. 2013 Computer programming used to be women's work. *Smithsonian Magazine*, **7** October. https://www.smithsonianmag.com/smart-news/computer-programming-used-to-be-womens-work-718061/ (Accessed: 1 August 2022).

Fillingim, K. B., **Shapiro, H.** & **Fu, K. K.** 2022 Error management bias in student design teams. *Journal of Mechanical Design* **145** (042302); doi:10.1115/1.4055899.

Flick, L. 1990 Scientist in residence program improving children's image of science and scientists. *School Science and Mathematics* **90** (3), 204–214.

Fu, K. K., **Yang, M. C.** & **Wood, K. L.** 2016 'Design principles. *Literature Review, Analysis, and Future Directions', Journal of Mechanical Design* **138** (10); doi:10.1115/1.4034105.

van der Gaag, M., *et al.* 2013 Development of the Davos assessment of cognitive biases scale (DACOBS). *Schizophrenia Research* **144** (1), 63–71; doi:10.1016/j.schres.2012.12.010.

Genco, N. *et al.* 2012 A study of the effectiveness of empathic experience design as a creativity technique. In *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 131–139. American Society of Mechanical Engineers Digital Collection; doi:10.1115/DETC2011-48256.

Goldin, C. & **Rouse, C.** 2000 Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review* **90** (4), 715–741; doi:10.1257/aer.90.4.715.

Green, C. E. L., *et al.* 2008 Measuring ideas of persecution and social reference: The Green *et al.* paranoid thought scales (GPTS). *Psychological Medicine* **38** (1), 101–111; doi:10.1017/S0033291707001638.

Grudin, J. 2006 Why personas work: The psychological evidence. In *The Persona Lifecycle*, pp. 642–663. Morgan Kaufmann Interactive Technologies; doi:10.1016/B978-012566251-2/50013-7.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Hallihan, G.M., **Cheong, H.** & **Shu, L.H.** 2013 Confirmation and cognitive bias in design cognition. In *ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 913–924, American Society of Mechanical Engineers Digital Collection; doi:10.1115/DETC2012-71258.

Hasso Plattner Institute of Design 2010 *An Introduction to Design Thinking Process Guide*. Stanford University.

Heilman, M. E. & **Martell, R. F.** 1986 Exposure to successful women: Antidote to sex discrimination in applicant screening decisions? *Organizational Behavior and Human Decision Processes* **37** (3), 376–390; doi:10.1016/0749-5978(86)90036-1.

Heylighen, A. & **Dong, A.** 2019 To empathise or not to empathise? Empathy and its limits in design. *Design Studies* **65**, 107–124; doi:10.1016/j.destud.2019.10.007.

ideo.org 2015 *The Field Guide to Human-Centered Design*, 1st Edn. Design Kit.

Islam, M. R. 2018 Sample size and its role in central limit theorem (CLT). *Computational and Applied Mathematics Journal* **4** (1), 1–7.

Jenkins, M. M. & **Youngstrom, E. A.** 2016 A randomized controlled trial of cognitive debiasing improves assessment and treatment selection for pediatric bipolar disorder. *Journal of Consulting and Clinical Psychology* **84** (4), 323–333; doi:10.1037/ccp0000070.

Jin, R. O., **Anaebere, T. C.** & **Haar, R. J.** 2021 Exploring bias in restraint use: Four strategies to mitigate bias in care of the agitated patient in the emergency department. *Academic Emergency Medicine* **28** (9), 1061–1066; doi:10.1111/acem.14277.

Johnson, S. K. & **Kirk, J. F.** 2020 Dual-anonymization yields promising results for reducing gender bias: A naturalistic field experiment of applications for Hubble space telescope time. *Publications of the Astronomical Society of the Pacific* **132** (1009), 034503; doi: 10.1088/1538-3873/ab6ce0.

Kahneman, D., **Lovallo, D.** & **Sibony, O.** 2011 Before you make that big decision. http://ir.vnulib.edu.vn/handle/123456789/4136 (Accessed: 4 October 2023).

Kahneman, D. & **Tversky, A.** 1972 Subjective probability: A judgment of representativeness. *Cognitive Psychology* **3** (3), 430–454; doi:10.1016/0010-0285(72)90016-3.

Kalra, P. & **Boukes, M.** 2021 Curbing journalistic gender bias: How activating awareness of gender bias in Indian journalists affects their reporting. *Journalism Practice* **15** (5), 651–668; doi:10.1080/17512786.2020.1755344.

Kelley, T. & **Littman, J.** 2001 *The Art of Innovation*. Doubleday.

Kendall, M. G. 1938 A new MEASURE of rank correlation. *Biometrika* **30** (1–2), 81–93; doi: 10.1093/biomet/30.1-2.81.

Khandelwal, M., *et al.* 2017 Why have improved cook-stove initiatives in India failed? *World Development* **92**, 13–27; doi:10.1016/j.worlddev.2016.11.006.

Kinsey, M. J., *et al.* 2021 Burning biases: Mitigating cognitive biases in fire engineering. *Fire and Materials* **45** (4), 543–552; doi:10.1002/fam.2824.

Korteling, J. E. H. & **Toet, A.** 2022 Cognitive biases. In *Encyclopedia of Behavioral Neuroscience* (ed. S. D. Sala), 2nd Edn. Elsevier; doi:10.1016/B978-0-12-809324-5.24105-9.

Kruger, J., *et al.* 2004 The effort heuristic. *Journal of Experimental Social Psychology* **40** (1), 91–98; doi:10.1016/S0022-1031(03)00065-9.

Legault, L., **Gutsell, J. N.** & **Inzlicht, M.** 2011 Ironic effects of Antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science* **22** (12), 1472–1477; doi:10.1177/0956797611427918.

Leibenstein, H. 1950 Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *The Quarterly Journal of Economics* **64** (2), 183–207; doi:10.2307/1882692.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Li, J., *et al.* 2021 Understanding customers across national cultures: The influence of national cultural differences on designers' empathic accuracy. *Journal of Engineering Design* **32** (10), 538–558; doi:10.1080/09544828.2021.1928022.

Li, J. & **Hölttä-Otto, K.** 2020 The influence of designers' cultural differences on the empathic accuracy of user understanding. *The Design Journal* **23** (5), 779–796; doi:10.1080/14606925.2020.1810414.

Linder, A. & **Svensson, M. Y.** 2019 Road safety: The average male as a norm in vehicle occupant crash safety assessment. *Interdisciplinary Science Reviews* **44** (2), 140–153; doi:10.1080/03080188.2019.1603870.

Lindgren, M., *et al.* 2018 Theory of mind in a first-episode psychosis population using the hinting task. *Psychiatry Research* **263**, 185–192; doi:10.1016/j.psychres.2018.03.014.

Makhlouf, T. E. *et al.* 2023 Who is this design for? Promoting consideration of people during concept generation. In *ASME 2023 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers Digital Collection; doi:10.1115/DETC2023-114673.

Mantonakis, A., *et al.* 2009 Order in choice: Effects of serial position on preferences. *Psychological Science* **20** (11), 1309–1312; doi:10.1111/j.1467-9280.2009.02453.x.

Marsden, N. & **Haag, M.** 2016 Stereotypes and politics: Reflections on personas. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 4017–4031. Association for Computing Machinery (CHI '16); doi:10.1145/2858036.2858151.

McDonagh, D., **Woodcock, A.** & **Iqbal, S.** 2018 *'Experience + Experience + Experience = Empathy: Design for Health', in*. Design for Health.

Misra, M. 2014 Warmth for Newborns: The embrace infant warmer. In *Innovations in Maternal Health: Case Studies from India*. SAGE Publications India, 147–157.

Moreno, D. P., *et al.* 2016 Overcoming design fixation: Design by analogy studies and nonintuitive findings. *AI EDAM* **30** (2), 185–199; doi:10.1017/S0890060416000068.

Morewedge, C. K., *et al.* 2015 Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights From the Behavioral and Brain Sciences* **2** (1), 129–140; doi:10.1177/2372732215600886.

Moss-Racusin, C. A., *et al.* 2018 Reducing STEM gender bias with VIDS (video interventions for diversity in STEM). *Journal of Experimental Psychology: Applied* **24** (2), 236–260; doi:10.1037/xap0000144.

Mourad, S. & **Tewfik, A.** 2016 Real-time data selection and ordering for cognitive bias mitigation. In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4393–4397; doi:10.1109/ICASSP.2016.7472507.

Mullinix, G. *et al.* 2013 Heuristica: Designing a serious game for improving decision making. In *2013 IEEE International Games Innovation Conference (IGIC). 2013 IEEE International Games Innovation Conference (IGIC)*, pp. 250–255; doi:10.1109/IGIC.2013.6659159.

Murata, A., **Nakamura, T.** & **Karwowski, W.** 2015 Influence of cognitive biases in distorting decision making and leading to critical Unfavorable incidents. *Safety* **1** (1), 44–58; doi:10.3390/safety1010044.

Narechania, A. *et al.* 2023 DataPilot: Utilizing quality and usage information for subset selection during visual data preparation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–18. Association for Computing Machinery (CHI '23); doi:10.1145/3544548.3581509.

National Center for Health Statistics 2022 Births and natality. https://www.cdc.gov/nchs/fastats/births.htm (Accessed: 31 October 2022).

Nickerson, R. S. 1998 Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* **2** (2), 175–220; doi:10.1037/1089-2680.2.2.175.

Nosek, B. A., **Banaji, M. R.** & **Greenwald, A. G.** 2002 Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice* **6**, 101–115; doi:10.1037/1089-2699.6.1.101.

Olopade, C. O., *et al.* 2017 Effect of a clean stove intervention on inflammatory biomarkers in pregnant women in Ibadan, Nigeria: A randomized controlled study. *Environment International* **98**, 181–190; doi:10.1016/j.envint.2016.11.004.

O'Sullivan, E. D. & **Schofield, S. J.** 2019 A cognitive forcing tool to mitigate cognitive bias – A randomised control trial. *BMC Medical Education* **19** (1), 12; doi:10.1186/s12909-018-1444-3.

Perez, C. C. 2019 *Invisible Women: Data Bias in a World Designed for Men*. Abrams Press.

Phillips, L. D. & **Edwards, W.** 1966 Conservatism in a simple probability inference task. *Journal of Experimental Psychology* **72** (3), 346–354; doi:10.1037/h0023653.

Raviselvam, S. *et al.* 2017 Demographic factors and their influence on designer creativity and empathy evoked through user extreme conditions. In *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers Digital Collection; doi: 10.1115/DETC2017-68380.

Raviselvam, S., **Hölttä-Otto, K.** & **Wood, K. L.** 2016 User extreme conditions to enhance designer empathy and creativity: Applications using visual impairment. In *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers Digital Collection; doi:10.1115/DETC2016-59602.

Reid, I. N. 2014 Gender-correlated systematics in HST proposal selection. *Publications of the Astronomical Society of the Pacific* **126** (944), 923–934; doi:10.1086/678964.

Risen, J. L. 2016 Believing what we do not believe: Acquiescence to superstitious beliefs and other powerful intuitions. *Psychological Review* **123** (2), 182–207; doi:10.1037/rev0000017.

Samuelson, W. & **Zeckhauser, R.** 1988 Status quo bias in decision making. *Journal of Risk and Uncertainty* **1** (1), 7–59; doi:10.1007/BF00055564.

Schauer, A. M. K., *et al.* 2024 Thinking beyond the default user: The impact of gender, stereotypes, and modality on interpretation of user needs. *Journal of Mechanical Design* **146** (5), 051403; doi:10.1115/1.4064263.

Shah, J. J., **Smith, S. M.** & **Vargas-Hernandez, N.** 2003 Metrics for measuring ideation effectiveness. *Design Studies* **24** (2), 111–134; doi:10.1016/S0142-694X(02)00034-0.

Sherbino, J., *et al.* 2011 The effectiveness of cognitive forcing strategies to decrease diagnostic error: An exploratory study. *Teaching and Learning in Medicine* **23** (1), 78–84; doi:10.1080/10401334.2011.536897.

Sherbino, J., *et al.* 2014 Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: A controlled trial. *Canadian Journal of Emergency Medicine* **16** (1), 34–40; doi:10.2310/8000.2013.130860.

Sibbald, M., *et al.* 2013 Do you have to re-examine to reconsider your diagnosis? Checklists and cardiac exam. *BMJ Quality & Safety* **22** (4), 333–338; doi:10.1136/bmjqs-2012-001537.

Sibbald, M., **de Bruin, A. B. H.** & **Merrienboer, J. J. G. v.** 2013 Checklists improve experts' diagnostic decisions. *Medical Education* **47** (3), 301–308; doi:10.1111/medu.12080.

Surma-Aho, A., **Björklund, T.** & **Hölttä-Otto, K.** 2018 An analysis of designer empathy in the early phases of design projects. In *Proceedings of NordDesign 2018. NordDesign 2018*.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Linköping, Sweden. https://www.designsociety.org/publication/40875/An+analysis+of+designer+empathy+in+the+early+phases+of+design+projects (Accessed: 28 October 2022).

Surma-aho, A. & **Hölttä-Otto, K.** 2022 Conceptualization and operationalization of empathy in design research. *Design Studies* **78**, 101075; doi:10.1016/j.destud.2021.101075.

Taleyarkhan, M., *et al.* 2023 Approach to problem solving and use of intuition by engineering technology students. *Journal of Global Education and Research* **7** (1), 81–98; doi:10.5038/2577-509X.7.1.1174.

Toh, C.A. *et al.* 2016 My idea is best! Ownership bias and its influence on engineering concept selection. In *ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers Digital Collection; doi:10.1115/DETC2015-46478.

Toh, C. A., **Strohmetz, A. A.** & **Miller, S. R.** 2016 The effects of gender and idea goodness on ownership bias in engineering design education. *Journal of Mechanical Design* **138** (10); doi:10.1115/1.4034107.

Tversky, A. & **Kahneman, D.** 1973 Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* **5** (2), 207–232; doi:10.1016/0010-0285(73)90033-9.

Tversky, A. & **Kahneman, D.** 1974 Judgment under uncertainty: Heuristics and biases. *Science* **185** (4157), 1124–1131; doi:10.1126/science.185.4157.1124.

Tversky, A. & **Kahneman, D.** 1981 The framing of decisions and the psychology of choice. *Science* **211** (4481), 453–458; doi:10.1126/science.7455683.

Ulrich, K. T. & **Eppinger, S. D.** 2007 *Product Design and Development*, 4th Edn. McGraw-Hill.

United States Bureau of Labor Statistics 2023 Labor force statistics from the current population survey. https://www.bls.gov/cps/cpsaat11.htm (Accessed: 5 June 2024).

U.S. Census Bureau 2023 QuickFacts: United States. https://www.census.gov/quickfacts/fact/table/US/SEX255221 (Accessed: 28 May 2024).

Visser, F. S. & **Stappers, P. J.** 2007 Mind the face. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces*, pp. 119–134. Association for Computing Machinery (DPPI '07); doi:10.1145/1314161.1314172.

Viswanathan, V. K. & **Linsey, J. S.** 2013 Role of sunk cost in engineering idea generation: An experimental investigation. *Journal of Mechanical Design* **135** (121002); doi:10.1115/1.4025290.

Vorvoreanu, M. *et al.* 2019 From gender biases to gender-inclusive design: An empirical investigation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19: CHI Conference on Human Factors in Computing Systems*, pp. 1–14. ACM; doi:10.1145/3290605.3300283.

Walmsley, S. & **Gilbey, A.** 2019 Understanding the past: Investigating the role of availability, outcome, and hindsight bias and close calls in visual pilots' weather-related decision making. *Applied Cognitive Psychology* **33** (6), 1124–1136; doi:10.1002/acp.3557.

Zahner, D., *et al.* 2010 A fix for fixation? Rerepresenting and abstracting as creative processes in the design of information systems. *AI EDAM* **24** (2), 231–244; doi:10.1017/S0890060410000077.

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press

Appendix

Design Considerations Checklist:

- ☐ Are your design decisions being influenced by the degree to which you are interested in the topic?
- ☐ How does your design offer room or account for user error?
- ☐ How was your design influenced by your judgment of who the user is? Did you apply stereotypical associations to your perception of the user?
- ☐ Are there any people who are unlikely to benefit from this design? Are there any people or groups of people who this design would not work for?
- ☐ Did you consider alternate use cases for this category of product?
- ☐ What assumptions did you make while creating this design?
- ☐ Are there cultural differences in the way people might use this product?
- ☐ Are there any events or occurrences in your life that influenced any parts of your design?
- ☐ What additional information would you want while tackling this design problem?
- ☐ Did you make any considerations while creating your design to make it accessible to a wide range of people?

https://doi.org/10.1017/dsj.2025.10006 Published online by Cambridge University Press