ASME Journal of Mechanical Design
Online journal at:
https://asmedigitalcollection.asme.org/mechanicaldesign

Examining the Design Actions and Reasoning Factors That Impact Design Performance

Yakira Mirabito
Mem. ASME
Department of Mechanical Engineering,
University of California, Berkeley,
Berkeley, CA 94720
e-mail: yakira.mirabito@berkeley.edu

Kosa Goucher-Lambert[1]
Mem. ASME
Department of Mechanical Engineering,
University of California, Berkeley,
6179 Etcheverry Hall,
Berkeley, CA 94720
e-mail: kosa@berkeley.edu

Engineers often do not select the best designs available to them. This research investigates whether specific design actions impact performance in a design exploration task and whether the reasoning factors underpinning these actions can be inferred directly. This study uses objective performance metrics to quantitatively evaluate multiple dimensions of design behavior and cognition within concept selection. Fifty-six participants were tasked with identifying an optimal design for the gripping contact of a dishwashing robot. Results identified that specific design actions correlated with improved design performance, including exploring fewer design alternatives and isolating parameters. We found that reasoning factors stated by participants did not accurately map onto their observed actions and did not correlate with task performance. Implications related to future computational design support tools are discussed. [DOI: 10.1115/1.4064414]

Keywords: decision theory, design evaluation, design process, design theory and methodology

1 Introduction

Improving final design outcomes is a shared goal within the design research community. While efforts aim to increase creativity and innovation in early-stage concept generation [1,2], recent findings from the design community have indicated that engineers do not always select the best available, instead opting for more feasible solutions [3–6]. This challenge serves as the initial motivation for this research, which aims to advance knowledge regarding the decision-making processes involved in concept selection in engineering design. Concept selection is a crucial phase that significantly impacts later stages, such as testing, development, and final deliverables [7]. After creating a set of design alternatives, engineers systematically narrow the options by evaluating and deciding which one(s) to further develop and implement. Although the selected concept may not become the final design, its features or functions of the concept may appear in the final solution [8]. Therefore, selecting an appropriate concept early on is vital to avoid negatively impacting project resources and stakeholders.

We investigate concept selection dynamics (design actions and reasoning factors) to explain what design behaviors and cognitive processes might contribute to selecting less optimal concepts (design performance). Design actions and reasoning factors are the focus variables due to their significance in design process documentation. In engineering education and in industry, engineers need to explain their design decisions often in a written report or oral presentation. In these mediums, what an engineer writes or says is believed to reflect the justifications behind their final design solution and its features. Note that the terms designers and engineers are used interchangeably throughout this paper.

We define *design actions* as the observable and quantifiable steps a designer takes within the task. For example, design actions might include changing between two distinct concepts, a bean bag to a chair with four legs, or modifying features like armrests or no armrests. We define *reasoning factors* as the rationale, motivation, and preferences designers use to guide their design actions. For the chair example, reasoning factors include each option's aesthetic characteristics or financial considerations. We define *performance* as the measure of the product's performance, not the performance of the design process or team (e.g., speed of decisions). For the chair example, design performance might include comfort ratings or max weight capacity. An underlying assumption guiding this work is that the design concepts considered vary in design performance—making some better relative to others.

The experimental design used in this study quantitatively explores how design behavior (actions) and design cognition (reasoning) impact design outcomes (performance) within a resource-constrained concept selection task. Quantitative studies on design behavior usually code design actions that are unique to the designer's solutions. Design actions can be infinite (e.g., material change, color, weight, and size). As used in this study, a simplified number of actions increases the power of quantitative and statistical methods. McComb et al. and Neroni et al. used truss design tasks and tracked design actions over time [9–12]. They also tracked system attributes (e.g., weight and costs) as their participants solved a design task. We tracked progress over time and asked our participants to select one concept to move forward to production.

Previous design research studies on reasoning often use qualitative analyses [13,14]. Design reasoning factors can be infinite, often

[1]Corresponding author.

Contributed by the Design Theory and Methodology Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 4, 2023; final manuscript received December 19, 2023; published online January 29, 2024. Assoc. Editor: Christine Toh.

captured via think-aloud protocols [15,16]. However, think-aloud methods are primarily qualitative, with sample sizes that are not statistically representative and create challenges in investigating causality. In order to quantitatively evaluate think-aloud methods, such an approach would be rather time-consuming and labor-intensive for the large sample size necessary. Additionally, objective performance measures are used rather than human raters to evaluate concept performance. We acknowledge that this research design presents limitations since other actions, reasoning factors, and performance measures could be used. However, this quantitative research design that uses constrained design actions and reasoning factors enables larger sample sizes and reproducibility that a qualitative approach lacks.

A two-part human subject study, a design task and a post-task survey, gathered information on design actions, reasoning factors, and performance from over 50 participants. The two research questions are:

> RQ1: What design actions and reasoning factors lead to improved design performance?

> RQ2: How might reasoning factors be inferred directly from design actions?

By evaluating multiple dimensions of concept selection, we expect to identify trends between each element. If clear trends are identified, extracting design reasoning information from design action data could be possible. The ability to infer reasoning from actions or actions from reasoning would reduce the amount of data necessary and enable researchers to extract this knowledge from incomplete data sources (i.e., missing data for actions or factors).

These expected outcomes are based on the assumption that design performance differences (dependent variable) can be explained by specific design actions or reasoning factors (independent variables). For example, designers who design a chair using free-body diagrams and equations might navigate a design challenge differently than engineers relying on their intuition from years of industry experience. Different approaches reveal insights into learning style or topic familiarity [17,18]. Several designers who solve a task in a similar manner may indicate a heuristic or bias. Implications for this work include identifying which design actions or reasoning factors to suggest designers use or avoid when engaging in design. Moreover, design researchers may consider using variations of this research design to quantitatively collect multiple dimensions of design behavior—in concept selection or other design stages.

2 Background

The following sections define concept selection, standard methods and tools, and related research that informed the specific design actions and reasoning factors explored in this paper. The term concept selection, as used in this paper, can be defined as a convergent phase in the design process in which design alternatives are evaluated and prioritized to determine which one(s) will be further developed and implemented. Selecting a concept is an executive function that influences design outcomes [19]. This systematic narrowing of design alternatives can be done individually or as a group. Research notes that all designers use some method to choose among concepts regardless of whether the method is explicitly stated [7]. By examining actions, reasoning, and performance in a single design study, understanding the relationships between elements could be realized and leveraged to inform design theory, practice, and education.

2.1 Concept Selection Methods and Tools. Due to the importance of concept selection in engineering design, many methods and tools are available to assist decision-makers. Tools include decision matrices, analytic hierarchy processes, uncertainty models, economic models, optimization concepts, and heuristics that could be leveraged individually or as a group [20]. Designers might select a concept using a range of formal (e.g., decision matrices and mechanical design principles) [21–23] or informal methods (e.g., intuition and gut feeling) [24]. A designer's final concept assessment could be captured as scores or weighting in Pugh matrices. However, this tool does not capture the entire evaluation process or the reasoning factors behind the decisions leading to the final design submission.

Moreover, concept selection tools such as decision matrices vary in effectiveness [21,25]. For example, a Pugh matrix is a decision-making tool where a design team (individually or as a whole) scores each concept alternative on a series of dimensions. These dimensions are usually what the team deems most important [22,24]. Although the tool focuses on objectivity, literature has identified cases where team members have selected criteria to rate to help confirm and build support for concepts of their preferences (i.e., confirmation bias) [26]. To reduce bias from concept selection methods and tools, none were provided to participants in this study. To better understand how these convergent decisions impact product performance, this work breaks down the concept selection phase into design behavior (actions) and design cognition (reasoning factors).

2.2 Design Actions. We define design actions as the observable and quantifiable steps in design. The participant's constrained set of design actions in this study include viewing, testing, or submitting a design. A design strategy consists of a string of design actions. The three strategies of interest include design space exploration, parameter isolation, and parameter prioritization. The design strategies of interest are based on optimization literature [20,27,28] and early observations from design actions and reasoning factors [29]. In previous observations, participants mentioned identifying the parameters that can be tuned (isolating variables) and focusing on one parameter at a time (prioritizing variables). These constrained design actions and strategies aim to generalize the findings beyond this specific design task.

Characterizing design behavior via design actions and strategies enables insights regarding heuristics or behavior comparisons across groups of designers. For example, work from Atman et al. [30] characterized the different stages in design and compared student and expert designers. Research has shown that designers can learn sequences based on the information provided or design actions possible, and this learning may or may not be conscious [31]. Due to the parameter tradeoffs and ability to test designs in this research study, participants may use optimization techniques [27,28,32] and interdependencies (coupled decisions) [33]. Design strategies in parameter tradeoff problems seek to minimize or maximize a function. In this study, participants can maximize design outcomes by selecting a design with a high success rate. Design actions and strategies serve as concrete evidence to explain how designers engage in a design task.

Moreover, the design research community seeks to capture sequential decision behavior from human designers to transfer to computational design agents to improve design performance. Research from McComb et al. identified beneficial decision sequences from designer behavior via Markov chains [9], aligning with prior sequence learning work [31]. In another paper, McComb et al. mined process heuristics via hidden Markov models that differed by design performance [10]. The first paper identified that participants used operation sequences but did not describe the operational sequences in detail or how they relate to engineering knowledge [9]. In the second paper, a gap remains in understanding why designers used the heuristics [10]. Similar work from Raina et al. focused on extracting design heuristics and transferring them to computational agents [34–36]. The computational agent created designs of similar performance or better, but the reasoning factors behind those actions remained unclear.

2.3 Reasoning Factors. A design reasoning factor, in this paper, is the explanation or consideration behind a design action.

Design reasoning factors are not necessarily observable. Instead, an engineer must explicitly communicate their reasoning factor(s), usually verbally or in written form (i.e., self-reported). The breadth of reasons varies from concrete, such as prototype feedback, to more abstract means, such as personal preferences or values [7]. Time and costs, independently and in combination, affect the design process and outcomes [37]. Design literature and protocol studies have also found that testing and previous experiences influence design solutions [24,38]. Designers with more years of experience tend to rely more heavily on these hands-on experiences than those with less experience who rely more on external design knowledge [13,39].

While reasoning factors in design are often captured by think-aloud protocols or interviews, research in psychology and management has shown that specific cognitive dimensions can be studied without explicitly collecting this information from participants. For example, research from Taylor et al. showed that innovative behavior (tool usage) in crows could also be explained by complex cognitive processes separate from learning mechanisms [40]. Robert Mitchell et al. found that differences in erratic decision-making processes could be explained by managers' self-reported perceptions of environmental factors (e.g., hostility) [41]. Within design research, Mao et al. were able to map specific cognitive processes to free-hand sketching actions in the design ideation phase [42]. In line with the approaches observed in previous research examining reasoning, our study combines open-ended and closed-ended responses to balance capturing rich, qualitative insights into participants' reasoning and ensuring a systematic analysis of specific reasoning factors.

Participants provided open-ended responses to explain their reasoning. They were also asked to rate the importance that a constrained set of reasoning factors (tests, guesses, principles, previous experience, datasheet, time, finances, and aesthetics) had on their decision-making. The factors asked in this study were inspired by prior findings from think-aloud protocols or interviews. The initial draft contained over 50 possible factors that were then reduced by grouping similar factors until the eight most relevant factors remained. Participants also had an opportunity to expand on the following three factors: principles, datasheet information, and previous experience. By linking multiple elements within concept selection, we expect to find specific reasoning factors to predict design performance. Thus, by identifying a subset of beneficial reasoning factors, suggestions regarding their usage could be recommended and integrated into design support systems.

3 Materials and Methods

This study examines design actions (e.g., viewing, testing, and selecting concepts) and reasoning factors (e.g., engineering principles and previous experiences) to understand their influence on one another and a participant's performance within the concept selection stage of engineering design. Data from a human subject study were analyzed to explore these patterns for insights into design behavior and cognition. Participants were tasked with submitting a design for a gripper surface for a dishwashing robot. This section outlines the participants, materials, and procedures used.

3.1 Participants. Sixty-eight participants were recruited for the design study using a call for participation at a University in California. Participants were compensated \$10 for their participation for 30 min. Participants were offered a bonus of up to \$20 contingent on task outcomes. Participants were screened and required to be 18 years or older with engineering or design experience to participate in the research study. No explicit domain expertise was required since the task could leverage different dimensions of engineering knowledge. For example, knowledge from mechanics, materials science, or design intuition could be used to solve the problem. Experience ranged from completing a single design class upwards to over ten years as an engineer in industry. Participant demographics included undergraduate, graduate, and working professionals with engineering and science backgrounds. All participants read, agreed, and signed a consent form. Data from 56 participants (29 men, 26 women, and one non-binary person) were used for data analysis. Such a diverse sample population should increase generalizability beyond one type of designer. Statistical tests were used to determine that experience, domain knowledge, or demographic factors did not confound performance. Data from 11 participants were removed due to a lack of following instructions (e.g., not completing the post-task survey) and equipment errors. One participant was removed since their total design actions exceeded three standard deviations above the mean.

3.2 Materials. For this study, participants were instructed to submit one design to move forward to production, the next step in the fictional robotics team's design process. Participants were tasked with designing a gripper surface for a robotic arm, as shown in Fig. 1. The dishwashing robot uses the grasper in a wet and slippery environment due to dish soap. The participant was

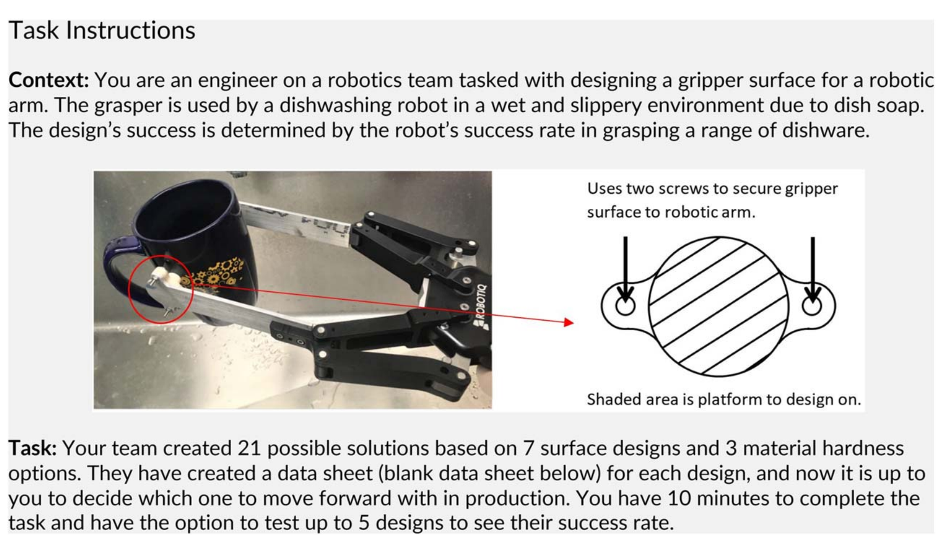

Fig. 1 Screenshot of the design instructions shown to participants. The screen includes information on the context, task, and a diagram of the gripper surface and its interaction mechanism with the grasper on the dishwashing robot [43].

presented with 21 solutions, a combination of seven surface geometries and three material hardness options. After selecting a design, a screen displayed its datasheet where participants could test that design to reveal its success rate, as noted in Fig. 2. A design's success rate was determined by the robot's ability to successfully grasp a range of dishware. Each participant had 10 min to complete the task. There was no time limit on reading instructions and understanding the design task. Participants could view all 21 designs, test up to five designs to see corresponding success rates (design outcomes), and submit one design. A test can be done at any point in the process for any of the 21 possible designs. Participants could submit any design, not necessarily a design that was tested. No specific concept selection method or tool was provided.

The task leverages empirical testing data and design alternatives based on a real haptic robotics design challenge. Our study focused on concept selection; thus, no concept generation phase was used. The decision to exclude concept generation reduced the possibility of ownership bias and design fixation that might have occurred had participants created their own solutions. In concept selection, participants use design information to evaluate and compare the alternatives. Information about each concept was conveyed via the datasheets, including the number of features, material, hardness, radius, contact area, weight, cost, and manufacturing time (exact values are noted in Table 6). Each item included would be known in an actual design process. Specific attributes distinguished the concepts from one another (e.g., geometry or hardness), while others were simply a matter of fact (e.g., materials and costs). Each participant may interpret the importance of each attribute differently. Hence, we explicitly instructed participants to select the design with the best success rate for this task.

3.2.1 Design Concepts and Success Rates Based on Empirical Data. The 21 concepts were provided by colleagues [43]. This section details the process those colleagues underwent and how the information they provided was used in this concept selection study. The 21 concepts were designed, prototyped, and tested by researchers investigating soft skins for robotic grasping applications in real-world environments such as kitchens where objects are covered in viscous fluids (i.e., soap, oil, and water). Increased friction helps improve grasping behavior. The researchers characterized the friction of soft skins of various circular features with the same total nominal contact area. Friction for each concept was measured in dry and lubricated environments (oil). Only the lubricated values are relevant to our fictional design task. Their experimental design positioned the soft skin with a weight on top and used a string and force gauge to measure friction for ten trials for each design [43]. The researchers also carried out a robotic gripper grabbing a plate in an oil bath, verifying that the lubricated friction coefficients correspond to a practical application.

Based on ten trials, the experimental kinetic friction coefficients (μ_k) ranged from 0.1743 to 3.8239. For our study, these values were translated from zero to four, where 0 represented 0%, and four

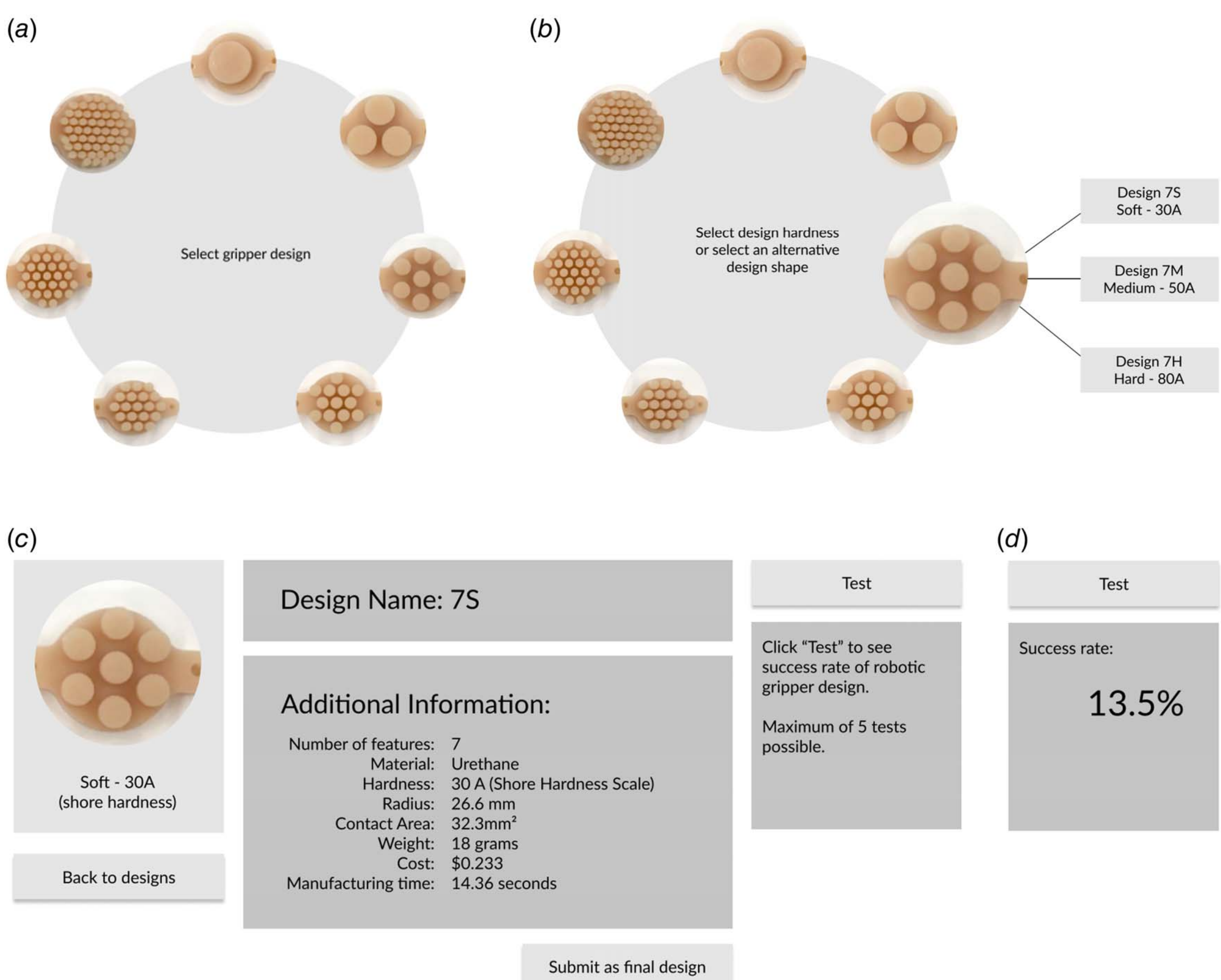

Fig. 2 The process flow of one design viewed and tested. (*a*) Main screen showing the seven geometry options. (*b*) After selecting geometry, a soft, medium, and hard material option is shown. Clicking one of the branched-out designs opens (*c*) a datasheet for that design. Clicking "test" displays (*d*) a design's success rate.

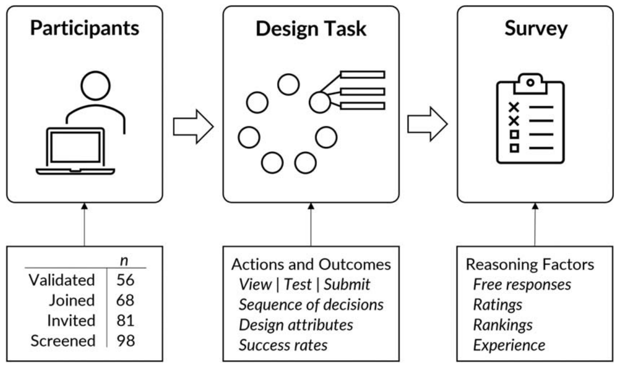

Fig. 3 Overview of explanatory research design. Participants first completed the design task in which their actions and outcomes were recorded. Afterward, a survey collected the reasoning behind their actions via open- and closed-ended questions.

represented 100%. The resulting fictional success rates ranged from 4.7% to 95.6%. Figure 6 shows a visual representation of the design space, characterized by geometry as the *x*-axis, hardness as the *y*-axis, and the success rate (as noted by the degree of shading). We opted against having values range from true zero to 100% because other designs are likely to outperform or underperform relative to those in the limited dataset provided. The experimental values were theoretically explained by bending stiffness (using the radius and height of each cylindrical feature) [43]. While designing for lubricated environments was not intuitive like dry environments, clear trends were present for friction/success rates across the geometry and hardness parameters. Our colleague's research inspired the decision to use a robotic gripper task in a dishwashing environment. While they used a grasping task in oil, the fictional task involved a slippery environment for a dishwashing robot grabbing a range of dishware rather than a single plate. The fictional test we described resembles commonly used tests in robotics dexterous grasping applications.

3.3 Procedure. The experiment took approximately 30 min and consisted of two parts: the design task and a post-task survey, as noted in Fig. 3. Colleagues provided the 21 gripper surface designs and corresponding success rates based on experimental friction data in a slippery environment [43]. The five-test limit was set to mimic real-life constraints in the design process, where a limited number of designs can be tested due to time or financial constraints. Preliminary experiments found that few participants converged on a "good" design with less than five tests. The 10-min time limit was used since time is limited in real-world scenarios. To account for the fact that they are stepping into the design after concept generation, participants had unlimited time to familiarize themselves with the task instructions and datasheet information before starting the timer. In the experimental data, the average task completion time was four and a half minutes.

Participants interacted with a graphical user interface for the design task portion, which displayed the consent form, task instructions, and possible design options (Fig. 2). The interface collected the time and number of tests but left it to the participant to monitor due to interface constraints. However, this decision to self-monitor was aligned with what engineers and designers experience outside of controlled studies, where they are expected to meet deadlines and stay within budget. Although the robotic gripper design could have been optimized using a computer program, this predefined solution space removed researcher subjectivity in classifying a participant's design actions and performance that alternative experimental setups may have introduced (e.g., having participants sketch their designs followed by researchers rating designs using rubrics). The likelihood of designer fixation and ownership bias was also reduced since the participants were not attached to their own ideas or pre-existing designs.

The post-task survey (Supplemental Material available on the ASME Digital Collection), hosted on Qualtrics, had closed-ended and open-ended questions about their design actions used and reasoning factors considered. The first three questions were open-ended, asking participants to explain how they arrived at their final design, their initial approach, and how that approach changed. Afterward, participants were asked to rate and rank the eight high-level reasoning factors (tests, guesses, principles, previous experience, datasheet, time, financial, and aesthetics) inspired by the prior literature that influenced their concept selection process. After listing and grouping similar reasoning factors from engineering design literature, the authors selected the eight most relevant factors. Any factors rated as any degree of influence were then carried forward and provided in a list for participants to rank. An additional question was asked to participants who rated the datasheet, engineering principles, and/or previous experience as having any degree of influence. These follow-up questions were asked to clarify the principles, datasheet information, or experiences participants leveraged. Since the other factors were less ambiguous, no follow-up questions were asked.

3.4 Data Analysis. Design performance was evaluated by success rate, an objective metric based on experimental friction data. Participant groupings were determined using design performance (i.e., success rate) for the final design submitted. Participant groupings enabled comparisons across design behavior and cognition between low- and high-performing designers. Participants in the high-performing category ($n = 19$) achieved the best possible design (95% success rate). The average-performing category achieved success rates above 50%, excluding the optimal design ($n = 19$), while those in the low-performing category achieved success rates under 50% ($n = 18$). The boundaries for high, average, and low categories are based on maintaining a similar sample size for each category. Additionally, the design success rates are based on empirically collected data and were not evenly distributed from 0 to 100.

Design actions defined in this study were steps traveled within the solution space. Design action data were collected from the Figma website using Maze.co, a clickstream collection platform. Each screen a participant visited was recorded, and each participant's duration, screenId, and sequential path were exported. By the nature of the experimental design, each participant's decision was linked to a corresponding success rate and moment in time. Design space explored was defined as the number of designs visited over the 21 total options (i.e., total design space). The geometries ranged from 1 to 55, while the urethane material ranged from 30 to 80 A shore hardness. Each design represented an equal proportion of the design space.

Decision strategies of isolating and prioritizing parameters were coded using the sequential path per participant, as described in Table 1. An increase, decrease, or hold was determined for each parameter, hardness, and geometry. A single parameter (SP) action meant one parameter was held constant while the other

Table 1 Concept selection strategies of interest and descriptions of the design behavior [29,44]

Strategy	Explanation
Isolating variables	Adjusting one variable at a time while holding all other parameters constant such as single parameter tuning [29]. Measured as the number of single parameter moves over the total number of single and multiparameter moves
Prioritizing variables	Focusing on a given variable throughout a portion of the decision-making process [30]. Measured by the number of sequential single parameter moves per one variable over the total number of moves

was adjusted. A participant's percentage of single parameter actions quantified the isolating variables strategy, as measured using the number of single parameter actions over the sum of single and multiparameter (MP) actions. Prioritizing parameters highlighted a participant's focus on a given variable throughout the task duration. Percent prioritization was computed as the number of sequential steps where one variable was the focus over the number of single parameter design actions.

Reasoning factors were collected with a post-task survey (Supplemental Material available on the ASME Digital Collection). The survey asked a mix of multiple-choice and open-ended questions. Ratings and rankings of the reasoning factors were computed, and statistical tests were used to determine the findings' significance. Spearman's correlations were run to assess relationships between factors considered and their influence on a participant's performance. Mann–Whitney statistical tests were carried out when comparing the performance of groups to one another. Four open-ended questions, three regarding design processes and one regarding prior experience, were qualitatively analyzed. The three questions about how they solved the problem, their initial approach, and how that approach changed were primarily used to describe an individual's observed behavior. Those qualitative insights also inspired the design strategies of interest (isolating and prioritizing variables). Meanwhile, the question on previous experiences was analyzed using thematic analysis (resulting in the reduced categories of hands-on and knowledge-based experience). The experience responses were reviewed and coded by a single researcher with prior experience in engineering design research and design practice. Afterward, each open-ended response and theme were visualized using an online whiteboard to help identify new links between the coded themes.

4 Results

This research aims to determine what design actions and reasoning factors designers might consider using to improve their design performance. The results identified that design actions such as exploring less of the design space and utilizing two design strategies correlated with improved design performance. These strategies were extracted from 2290 total design actions. Participants took as few as 20 actions up to 96 actions ($M = 42$, $SD = 18$) and viewed as few as four designs up to 14 unique designs ($M = 7$). The relative importance of specific reasoning factors guiding participants' design behavior is presented. The factors were not correlated with design performance. The impacts on design outcomes are presented as correlations (based on success rate as a percentage) and comparisons (between low- and high-performing groups). Results show that high- and low-performing designers differed in their actions and strategies.

4.1 Influence of Design Actions on Design Outcomes. Exploring a larger percentage of the design space did not correlate with improved design outcomes. Design space exploration was computed as a percentage of the total designs assessed over 21 design alternatives. Each design represents about 5% of the design space. Design space exploration ranged from 19% to 67%, as noted in Fig. 4. A statistically significant and moderate negative correlation was observed between the design space explored and the final design's success rate ($r_s(54) = -0.35$, $p < 0.01$). High- and low-performing designers differed in the percentage of design space explored ($M_{\text{high}} = 27.3\%$ and $M_{\text{low}} = 33.7\%$; Mann–Whitney $U = 101.5$, $n_1 = 19$, $n_2 = 18$, $p < 0.05$ two-tailed).

Isolating variables or single parameter tuning was the most identifiable strategy likely to improve design outcomes. Single parameter tuning means adjusting one variable at a time while holding the other parameter constant (e.g., changing geometry while holding material hardness constant). Spearman's rank correlation tests were carried out between the percent of single parameter actions and the success rate achieved. The percentage of single parameter

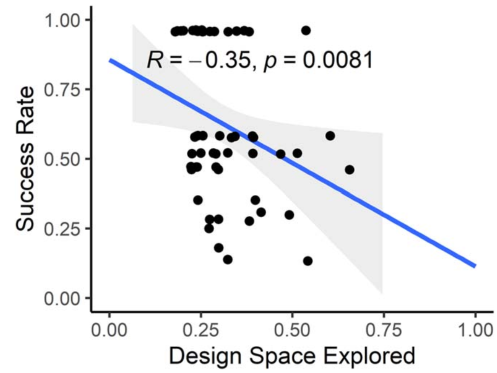

Fig. 4 Scatterplot of the percentage of design space explored (number of datasheets explored of the 21 possible datasheets) and the success rate of the final design submitted. The shaded region represents a 95% confidence interval for the regression line. Each participant is represented as a single data point.

actions was coded as the number of single parameter classifications over the sum of single and multiparameter classifications. The results show a moderate correlation, Fig. 5(*a*), which is statistically significant ($r_s(54) = 0.33$, $p < 0.05$). Therefore, participants who engaged in more single parameter actions were more likely to have improved performance. When participants did not use single parameter actions, they performed multiparameter design actions, simultaneously adjusting both parameters. Multiparameter actions highlight how participants jumped around the design space. As a result, participants could be less likely to understand each variable's influence on design performance. For the isolating variables strategy, there was a 15% statistically significant difference across usage for high- and low-performers ($M_{\text{high}} = 81\%$ and $M_{\text{low}} = 66\%$; Mann–Whitney $U = 92.5$, $n_1 = 19$, $n_2 = 18$, $p < 0.05$ two-tailed).

Prioritizing variables was the second design strategy shown to improve outcomes. Prioritizing variables means focusing on a given variable throughout the decision-making process as measured by the number of sequential steps per variable. For one variable to be held constant, isolating variables or single parameter tuning must occur. Spearman's rank correlation tests were carried out between the percent of variable prioritization and the success rate achieved. Percent prioritization is the sum of actions that are single parameter and hold parameters constant between a series of sequential steps over the total number of design actions. A participant with high prioritization carried out primarily single parameter actions and, of those actions, held hardness constant while adjusting geometry. The results show a moderate correlation, Fig. 5(*b*), which is statistically significant ($r_s(54) = 0.43$, $p < 0.01$). For the prioritizing variables strategy, there was a 23% statistically significant difference across usage between the groups ($M_{\text{high}} = 68\%$ and $M_{\text{low}} = 44\%$; Mann–Whitney $U = 72.5$, $n_1 = 19$, $n_2 = 18$, $p < 0.05$ two-tailed).

4.2 Influence of Design Reasoning Factors on Design Outcomes. The post-task survey collected design reasoning factors via closed and open-ended questions (questions included as Supplemental Materials available in the Supplemental Materials on the ASME Digital Collection). The ratings in Tables 2–4 use a 5-point Likert scale that ranges from not at all important (1-rating) to a great deal (5-rating). Table 2 shows the average ratings for the eight primary factors of interest in response to the question, "Please rate the impact that the following factors had on your design decision." Open-ended responses were coded using thematic analysis. The strategies of interest (isolating and prioritizing variables) were extracted from three questions about design approaches. Hands-on versus knowledge-based experiences were identified from one question about previous experience. Spearman's correlational tests were conducted between participants' ratings of

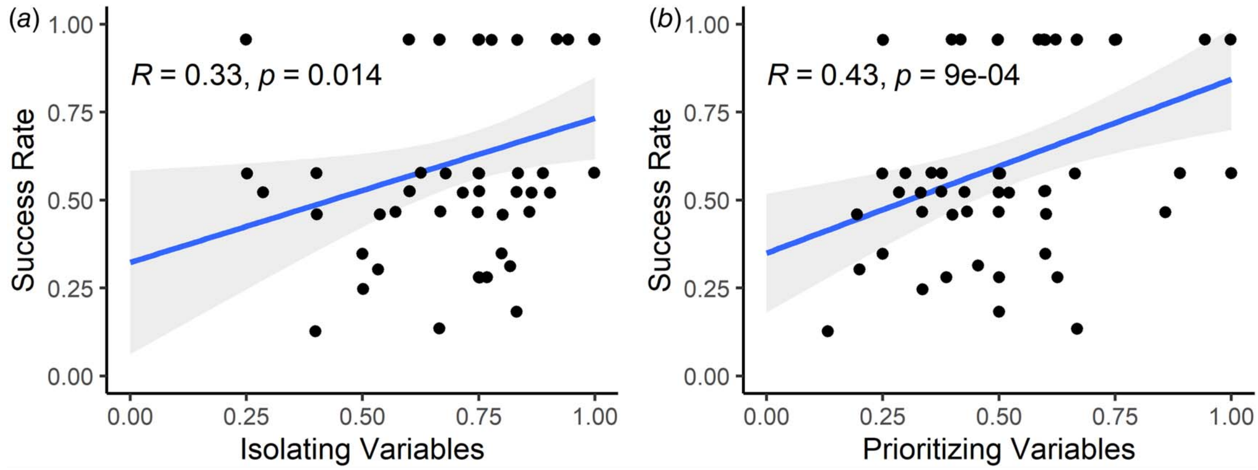

Fig. 5 (*a*) **The percentage of isolating variables and corresponding success rate percentage. (*b*) The percentage of prioritizing variables and corresponding success rate percentage. The shaded region represents a 95% confidence interval for the regression line. Each participant is represented by one data point.**

factors (e.g., tests) and their final success rate (total of eight tests). None of the findings were statistically significant. We rejected the hypothesis that design reasoning factors correlate with design performance. Additional analyses regarding task information, engineering/design principles, and previous experience are presented in the following sections.

Table 2 Factors rated by importance

Factor	Mean rating	Standard error
Tests	4.80	0.074
Guess	3.75	0.133
Principles	3.23	0.125
Previous experience	2.84	0.176
Datasheet	2.66	0.166
Time	2.30	0.171
Financial	1.43	0.088
Aesthetics	1.43	0.101

Table 3 Datasheet attributes rated by importance

Attribute	Mean rating	Standard error
Success rate	4.87	0.080
Hardness	4.26	0.130
Geometry	3.59	0.198
Contact area	3.17	0.197
Radius	2.67	0.204
Image	2.63	0.207
Material	2.41	0.198
Manufacturing	1.85	0.139
Cost	1.80	0.138
Weight	1.50	0.107

Table 4 Engineering/design principles rated by importance

Principle	Mean rating	Standard error
Friction	4.07	0.151
Materials	3.04	0.174
Mechanics	3.04	0.182
Robotics	2.15	0.163
Manufacturing	2.04	0.142

Task information relates to knowledge acquired via the design task: tests and datasheets. Test importance was rated highly across all participants. Only two participants rated tests as having moderate or little influence on their decision process. Forty-six participants rated the datasheet as having any degree of influence and were asked a follow-up question, "Please rate the impact that the following information from the datasheet had on your design decision." Table 3 shows the mean results of the datasheet attributes. None of the datasheet attributes were correlated with design outcomes.

Engineering knowledge means foundational theories or knowledge one might acquire by attending a lecture or reading a textbook. A follow-up survey question was asked to the 55 participants who rated engineering/design principles as having any degree of influence. Participants were asked to rate the influence of five topics in engineering/design that guided their design-making. The question shown read, "Please rate the impact that the following principles had on your design decision." Table 4 shows the mean results for the group. Spearman's correlation tests were run between the importance ratings and the corresponding success rates of the respondents. None of the correlations were statistically significant. Two additional questions asked participants to rate their knowledge of materials and robotics. Spearman's correlation tests between participant rating of their knowledge and design outcomes were also not statistically significant, and neither were the Mann–Whitney tests run between the high- and low-performing designers.

Participants rated the previous experience factors a 2.84 on average importance, as noted in Table 2. An additional open-ended question was asked to the 44 participants who rated previous experience as having any degree of influence in their decision-making process. Two aspects of participants' responses stood out during the thematic analysis. First, the relevance of the experience they referenced (hands-on versus not hands-on), and second, the rhetoric participants used when communicating experience. Initial codes used to group participant experiences included *robotics*, *friction*, *course knowledge*, *vague language*, and *dishwashing experience*. Afterward, focused coding categorized each participant's previous experience response as hands-on, knowledge-based, or vague, as shown in Table 5. Hands-on design experiences included creating grippers for medical or robotics applications, and others mentioned using friction devices for various applications (e.g., clamps and doorstops). Other hands-on previous experiences alluded to memories of themselves dishwashing or analyses of dishwashing gloves. The theme of knowledge-based experiences referenced information from courses or research (e.g., contact area and biomimicry). The information was similar to the engineering and design principles factors from Table 4. The last theme of vague statements

Table 5 Themes from previous experience coded segments

Previous experience code	Representative coded segment	Frequency
Hands-on	I used to be a dishwasher and having multiple gripping locations made it less likely to slip, which motivated the assumption of using more appendages in the design	28
Knowledge-based	From previous experience in my manufacturing class, I learned that a minimum of 3 contact points is important to establish a datum so I applied that to here as a datum needs to hold an object in place and so did the product	11
Vague	I feel that it should work better with a larger number of features	5

categorized experiences that did not contain enough information regarding a specific experience and thus were not compared to the other groupings. When comparing the hands-on and knowledge-based groups, no statistically significant difference exists between their success rates ($M_{\text{hands-on}} = 59\%$ and $M_{\text{knowledge}} = 73\%$; Mann–Whitney $U = 105$, $n_1 = 11$, $n_2 = 28$, $p = 0.13$ two-tailed).

5 Discussion

Concept selection dynamics were defined in this study as design actions, reasoning factors, and outcomes. Using a design task followed by a post-task survey, three main findings were identified. First, this research showed that specific design actions were correlated with improved performance and were statistically significant. Next, no trend between design reasoning factors and performance was found. Lastly, linking design actions, reasoning factors, and outcomes showed that multiple reasoning factors motivated the same design action. The following sections outline the implications of the findings, future work to fill in gaps in reasoning, and the influence of design rationale's explainability on human–artificial intelligence (AI) collaboration.

5.1 What Design Actions and Reasoning Factors Lead to Improved Design Performance?

5.1.1 Design Actions. Design space exploration, parameter isolation, and prioritization were three design behaviors that correlated with design performance. Isolating and prioritizing variables were positively correlated with performance, and both correlations were statistically significant ($r_s(54) = 0.33$, $p < 0.05$ and $r_s(54) = 0.43$, $p < 0.01$, respectively). Design space exploration was negatively correlated with performance ($r_s(54) = -0.35$, $p < 0.01$). Additional statistical tests (i.e., Mann–Whitney U test) evaluated whether each design strategy usage differed across low- and high-performing designers. All three tests were statistically significant, as shown in Sec. 4.1, meaning design behavior was unlikely due to chance. Our research adds to the existing literature that identifies and extracts beneficial design behavior [34,35]. For example, the extracted behaviors can help human designers and computational design agents select higher-performing designs in the concept selection phase. Future work may consider longer or more complex concept selection tasks. Increasing length and complexity would enrich the dataset for new insights regarding actions and strategies.

One might assume that a focused search strategy could be attributed to luck. While this may be true for a few participants, those who tested the best design early on (either strategically or by chance) still conducted all five tests. This additional testing could be attributed to participants believing a design with a 100% success rate could have existed. Even more surprising, a few participants who tested the best design submitted a design for which they had not tested. Strategic navigation of the solution space should have quickly revealed the trend that the hard material option for any given geometry was the best (see Table 6 or Fig. 6 for the exact values), yet about 25% of participants submitted designs with a soft or medium hardness. Open-ended explanations for selecting a soft option noted an individual bias that softer items in their personal experience should have better friction.

Nonetheless, based on the findings presented, human designers could benefit from understanding what they are doing well and not so well. This paper identified a few systematic design actions that correlated with improved performance. However, some participants underperformed or overperformed relative to the line of best fit. For example, in Fig. 5(*b*), a participant underperformed with a success rate of under 25% despite a high strategy usage of over 75% for isolating variables. These varying results highlight the human element of decision-making, which randomness, designer preferences, or biases might help explain. In a separate paper, data visualizations of individual design behavior (i.e., the two parameters, geometry and hardness, and over task duration) can help designers understand their concept selection dynamics [44]. Visualizations can show when participants engage with useful heuristics or less helpful biases such as anchoring or design fixation.

5.1.2 Reasoning Factors. Reasoning factors explored in this study were not correlated with task performance. Instead, the results provide insight into and quantify the relative influence that each reasoning factor had on participants' design processes. Table 2 shows the average ratings for the eight main factors. Tables 3 and 4 show the average ratings for the subfactors associated with datasheet and engineering principles, respectively. Table 5 shows the main categories (hands-on, knowledge-based, and vague) extracted from the open-ended responses of the previous experience factor. While no trend between reasoning factors and performance was found, some possible reasons this may have occurred are discussed as well as the relevance of this finding in engineering.

Differences in design performance despite having the same rating for a given factor can be explained by the differences in how participants perceive and use the information. For example, two participants who each rated 'tests' as having a great deal of influence on their design process could be explained both by participants who used testing strategically (isolating or prioritizing variables) or relied purely on a guess-and-check approach. These findings align with behavioral economics and psychology research, which finds that people's actions are inconsistent with what they say or think they are doing [45,46]. This incongruence may have occurred due to biases (e.g., social desirability, hindsight, and confirmation). While other domains have found human behavior to misalign with their explanations (e.g., purchase behavior), this finding in an engineering-specific domain is notable since stakeholders assume design reasoning in reports and presentations represent the truth.

Therefore, if self-reported reasoning does not explain a designer's actions, this raises concerns about the validity of a designer's documented reasoning. Their reasoning could also indicate other problems regarding communication ability or credibility as designers (e.g., they know they need to test designs strategically but cannot realize this through their actions). The lack of detail in open-ended questions (i.e., previous experience and approaches) was surprising since engineers and designers are expected to state their rationales and processes in design reports or industry documentation. The responses were not particularly short. Instead, the responses were vague and non-specific, resembling linguistic patterns coded (e.g., affective decisions, hedges, and boosters) in research by Krishnakumar et al. [47]. Effective communication of the design rationales behind design decisions is crucial to clearly understanding a designer's process of evaluating and selecting a design amongst multiple alternatives [48]. Several design support systems and tools have been developed to aid designers in this process [49,50]. Perhaps one of those frameworks could be used in future studies (e.g., feature, process, mechanistic, and systematic).

Future work that simultaneously investigates design actions and reasoning factors could explore alternatives to how reasoning factors are captured and represented. Conducting a more detailed investigation into one of the reasoning factors could split participants into conditions and manipulate the influence a given factor should have on their design processes. Iterating on this study design might include periodic check-ins during the design task in which participants document their design reasoning factors—explicitly asking the same survey questions, perhaps at three points, early, middle, and end of the task duration. A downfall of this method might bias participants to consider the factors shown even if they had not considered them previously. Alternatively, a semi-structured interview with participants could help fill in the gaps when participants did not communicate reasoning factors clearly in the free responses or when follow-up questions would be beneficial in understanding their reasoning processes. Furthermore, using a knowledge graph for analyses could provide further insight into how the reasoning factors designers use are interconnected [51]. Knowledge graphs are networks of data that store information and illustrate relationships within the system, such as TechNet [52]. Research shows the importance of knowledge graphs in transferring and organizing knowledge within organizations and can offer insight into designer intent and innovation [53–55].

5.2 How Might Reasoning Factors Be Inferred Directly From Design Actions?. This research sought to provide a more holistic understanding of the dynamic cognitive processes designers engage in during the concept selection phase. Design actions, reasoning factors, and outcomes were examined using a pre-defined solution space. The research design enabled direct comparisons across participants. Despite having these three directly comparable elements, design reasoning factors remain the least understood and least connected to design outcomes. Design reasoning factors and design actions were linked, but multiple reasoning factors were linked to a single design action. Similarly, a single reasoning factor was linked to multiple design actions. Gaining a shared understanding of the types of reasoning factors that participants use when carrying out a sequence of design actions could be beneficial for designers and their teams.

Thematic analysis was used to code the first three open-ended survey questions. The themes that emerged in conjunction with literature on engineering optimization approaches inspired the design strategies of isolating and prioritizing design parameters. On average, participants rated geometry and hardness as the most influential datasheet pieces of information in their decision process (Table 2). Design reasoning explicitly stating isolating and prioritizing variables was quickly extracted from design actions. There were nuances in the way participants executed such design actions. Several participants stated they wanted to explore the possible solutions, but how they explored the design space differed. Some started at the extremes, while others only explored a focused area. The designs considered could be characterized by their search or exploration behaviors and influenced by the interface's capabilities [38,56,57]. Despite grouping designers based on similar reasoning considerations, the breadth of design actions was vast.

Once design actions were concretely identified, hypotheses regarding their reasoning factors were made. The results found some reasoning factors to align with what design actions participants executed—for example, exploring less of the design space correlated with improved outcomes, which could be explained by increased domain knowledge in robotics or materials science [58]. Moreover, exploring less of the design space could also be explained by luck in selecting a closer starting position to the optimal design. We understand that design reasoning factors in this paper were captured after completing the design task. Future studies should explore sequential design reasoning factors similar to what this paper has done with design actions. Breaking down the design process into smaller segments might enable directly linking salient reasoning factors with design actions and showcase how they influence one another over task duration.

5.3 Implications in Engineering Design. The design problem used in this study aligns with scenarios in late-stage conceptual design or early embodiment design. The pre-defined solution space is mostly smooth, although the larger solution space (i.e., alternative materials or geometries) may be characterized as "rugged" since best practices for designing robotic graspers in lubricated environments are not well understood. Therefore, the researchers who provided this empirical dataset already made a series of decisions before prototyping and testing these designs [43]. Fundamentally, deciding which designs to test and select are complex decisions, such that these are decisions that involve multiple factors, uncertainties, and tradeoffs. The scope of the task captures these decision-making characteristics while also considering human capabilities (attention, perception, and memory). Increasing the number of parameters causes much higher cognitive loads, and problem-solving at that scale may be beyond that of human capabilities (e.g., requires computational approaches) [59]. Thus, the problem structure used aligns with the scale of decision-making that humans make in late-stage conceptual design or early embodiment design.

The main implications of these findings for engineering design lie in improving design processes for human designers. This study shows that design actions could be extracted, and some actions positively impact design outcomes. Once the impact on design performance for each strategy is validated, we could suggest that designers use them in practice. Each strategy can be taught in design education or practice as a heuristic for concept selection. Design heuristics have previously been successfully extracted and taught for concept generation [60,61]. Using a given strategy can be visualized to help designers understand when they are and are not engaging with the strategies. The visualization could be an assessment tool to check whether the strategy was correctly learned and applied, similar to how gradient maps help show optimization approaches in finding global minima [28]. One of the future applications of these findings relates to collaborations between humans and AI in engineering design tasks.

As outlined in prior work, the two strategies of isolating and prioritizing parameters can be incorporated into computational models that describe, explain, or predict engineering design behavior [44]. The results show that design rationales were not always clear, and instead, multiple design rationales motivated similar design actions. By incorporating human designers' rationale, a deeper understanding that combines what the agents are doing, design actions, and why the agents are doing so, design rationale, can help improve decision support tools within engineering design [62–64]. Using design reasoning factors provides richer design recommendations than currently possible (simply imitating design actions).

Previous research has shown several advantages to introducing AI design agents to assist human designers at various stages of the engineering design process [65–68]. Research from Raina et al. used deep learning to imitate human designers, where the system performed just as well or outperformed human designers. Current design support tools in development often mimic human design decisions. The agent can learn and imitate behavior but cannot articulate the motivation or evidence supporting the designer's actions [34,35,69,70]. Das and Chernova showed that clear rationale improved users' ability to understand and accept an agent's recommendation, increasing task performance [71].

Designer confidence in these decision support tools depends on the tool's explainability—the ability to explain how the tool functions [70]. Our study showed that various reasoning factors motivate various design actions. Thus, we could expect a wide range of human trust levels in an agent's design recommendation. Previous research in autonomous vehicles showed that human trust in the system was influenced by communication style and the level of information provided [72]. Dong et al. showed that the logical framing structure (i.e., abductive and deductive) significantly influenced design decisions within human-to-human interactions [73]. Deductive reasoning was more likely to cause human participants to reject proposed designs. In contrast, abductive reasoning, commonly associated with creative processes, was more likely to cause human participants to accept a design product or feature. Considering that the structure of design reasoning was shown to influence design decisions within human-to-human interactions [73], the structure used by computational agents should influence design decisions within human–AI interactions. While the findings in this paper showcase a limited breadth of reasoning factors used, additional research might want to better define the level of detail necessary to communicate rationale or focus on select reasoning factors.

5.4 Limitations. While we leveraged an experimental design that sought to capture observed design behavior and design cognition through reasoning factors, future studies should consider the following adjustments to the task and data collection procedures. Currently, the generalizability of the results is limited due to the use of a constrained solution space based on empirically collected data for one realistic design task [45]. Increasing sample sizes or using multiple design tasks instead of one increases the generalizability of future studies. The design space could be more evenly distributed (e.g., equal intervals between geometry) or more comprehensive (e.g., more than 21 concepts). When the space is evenly distributed, it reduces the chances that specific designs are underrepresented, and outliers are less likely to influence results.

We must acknowledge the inherent tradeoff between exploring more of the design space and using more resources to create high-fidelity prototypes. A solutions space that can be generated computationally may address this concern. Concerning the design space explored, each design was assumed to be the same percentage of the design space. Alternatively, had this not been equal, designs of larger geometries would have covered a more significant percentage of the design space than smaller geometries. Future work should carefully consider the creation and distribution of the design space.

Moreover, while self-reporting design reasoning is the standard in industry and education (via oral presentations and technical reports), improvements can be made to better study the impact of reasoning factors on design performance. For example, creating different conditions in which participants are primed (subconsciously) to consider one factor could enable more direct conclusions on the influence that each factor has on design performance. Alternatively, researchers could design a task where the degree of influence for a given factor can be easily extracted from actions (e.g., quantifying financial influence by the total cost of a submitted design).

6 Conclusion

This paper presents an empirical study examining multiple dimensions of design activity (i.e., design actions, reasoning factors, and performance) for a concept selection task. The experimental design addressed a gap in prior research by quantitatively comparing design actions, reasoning factors, and performance simultaneously for over 50 participants. Using a robotic gripper design problem as a test scenario, detailed design process data from 56 participants were collected with a graphical user interface and post-task survey. Results highlight specific design actions correlated with improved design performance, such as a focused search strategy and isolating and prioritizing design parameters. Design reasoning factors (the decision-making heuristics underpinning observed behavior) were segmented into task-specific knowledge, engineering knowledge, and previous experience. Although design reasoning factors motivated design actions, participants with similar reasoning considerations did not have similar design actions or achieve similar outcomes. Why designers select lower-performing designs can be partially explained by design actions but not reasoning factors. Future work will consider alternative methods to capture reasoning factors and explore the rhetoric strategies designers use to rationalize their decisions.

Acknowledgment

This work is based upon preliminary research presented at the 2021 ASME International Design Engineering Technical Conferences (DETC2021-71830) and the 2022 Design Computing and Cognition Conference (Paper 61). This work has been supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2146752, the University of California Regents, and partially supported by the National Science Foundation under Grant No. DGE-1633740. The findings presented in this work represent the views of the authors and not necessarily those of the sponsors.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

Appendix

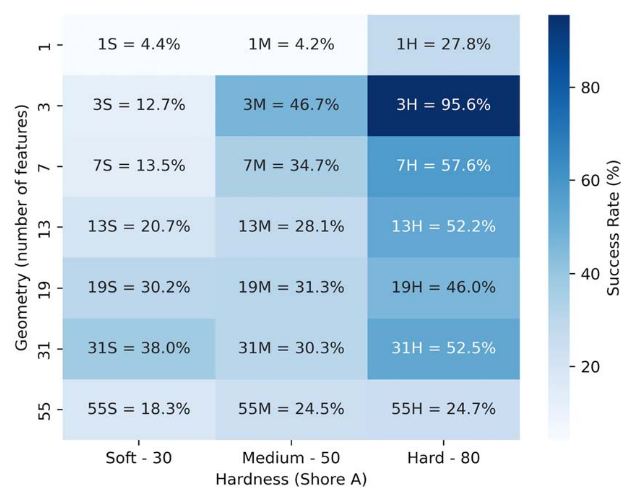

Fig. 6 Design alternatives characterized by geometry, hardness, and success rate

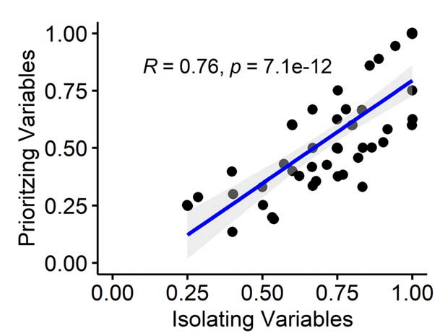

Fig. 7 Strategies of interest are correlated with each other

Table 6 Design alternatives with datasheet information

Design option	Number of features	Material	Hardness (shore A)	Radius (mm)	Nominal contact area (mm^2)	Weight (g)	Cost ($)	Time (s)	Success rate (%)
1S	1	Urethane	30	70.8	32.3	18	0.234	7.12	4.4
1M	1	Urethane	50	70.8	32.3	20	0.272	7.12	4.2
1H	1	Urethane	80	70.8	32.3	22	0.274	7.12	27.8
3S	3	Urethane	30	41.7	32.3	18	0.234	10.54	12.7
3M	3	Urethane	50	41.7	32.3	20	0.272	10.54	46.7
3H	3	Urethane	80	41.7	32.3	22	0.274	10.54	95.6
7S	7	Urethane	30	26.6	32.3	18	0.234	14.36	13.5
7M	7	Urethane	50	26.6	32.3	20	0.272	14.36	34.7
7H	7	Urethane	80	26.6	32.3	22	0.274	14.36	57.6
13S	13	Urethane	30	19.4	32.3	18	0.234	18.55	20.7
13M	13	Urethane	50	19.4	32.3	20	0.272	18.55	28.1
13H	13	Urethane	80	19.4	32.3	22	0.274	18.55	52.2
19S	19	Urethane	30	16.0	32.3	18	0.234	21.72	30.2
19M	19	Urethane	50	16.0	32.3	20	0.272	21.72	31.3
19H	19	Urethane	80	16.0	32.3	22	0.274	21.72	46.0
31S	31	Urethane	30	12.4	32.3	18	0.234	26.85	38.0
31M	31	Urethane	50	12.4	32.3	20	0.272	26.85	30.3
31H	31	Urethane	80	12.4	32.3	22	0.274	26.85	52.5
55S	55	Urethane	30	9.2	32.3	18	0.234	34.59	18.3
55M	55	Urethane	50	9.2	32.3	20	0.272	34.59	24.5
55H	55	Urethane	80	9.2	32.3	22	0.274	34.59	24.7

Table 7 Cleaned behavior vector for one high-performing participant

Action	Screen	Geometry	Hardness	Single or multiparameter
1	19S	19	30	–
2	19M	19	50	SP
3	31M	31	50	SP
4	19S	19	30	MP
5	19S_test	19	30	TEST
6	55S	55	30	SP
7	55S_test	55	30	TEST
8	7H	7	80	MP
9	7H_test	7	80	TEST
10	3H	3	80	SP
11	3H_test	3	80	TEST
12	1H	1	80	SP
13	1H_test	1	80	TEST
14	3H	3	80	SP

References

[1] Goucher-Lambert, K., and Cagan, J., 2019, "Crowdsourcing Inspiration: Using Crowd Generated Inspirational Stimuli to Support Designer Ideation," Des. Stud., **61**, pp. 1–29.

[2] Mirabito, Y., and Goucher-Lambert, K., 2022, "Factors Impacting Highly Innovative Designs: Idea Fluency, Timing, and Order," ASME J. Mech. Des., **144**(1), p. 011401.

[3] Zheng, X., Ritter, S. C., and Miller, S. R., 2018, "How Concept Selection Tools Impact the Development of Creative Ideas in Engineering Design Education," ASME J. Mech. Des., **140**(5), p. 052002.

[4] Toh, C. A., and Miller, S. R., 2015, "How Engineering Teams Select Design Concepts: A View Through the Lens of Creativity," Des. Stud., **38**, pp. 111–138.

[5] Kazerounian, K., and Foley, S., 2007, "Barriers to Creativity in Engineering Education: A Study of Instructors and Students Perceptions," ASME J. Mech. Des., **129**(7), pp. 761–768.

[6] Rietzschel, E. F., Nijstad, B. A., and Stroebe, W., 2010, "The Selection of Creative Ideas After Individual Idea Generation: Choosing Between Creativity and Impact," Br. J. Psychol., **101**(1), pp. 47–68.

[7] Ulrich, K. T., Eppinger, S. D., and Yang, M. C., 2020, *Product Design and Development*, 7th ed., McGraw-Hill Education, New York.

[8] He, Y., Camburn, B., Liu, H., Luo, J., Yang, M., and Wood, K., 2019, "Mining and Representing the Concept Space of Existing Ideas for Directed Ideation," ASME J. Mech. Des., **141**(12), p. 121101.

[9] McComb, C., Cagan, J., and Kotovsky, K., 2017, "Capturing Human Sequence-Learning Abilities in Configuration Design Tasks Through Markov Chains," ASME J. Mech. Des., **139**(9), p. 091101.

[10] McComb, C., Cagan, J., and Kotovsky, K., 2017, "Mining Process Heuristics From Designer Action Data Via Hidden Markov Models," ASME J. Mech. Des., **139**(11), p. 111412.

[11] Neroni, M. A., Vasconcelos, L. A., and Crilly, N., 2017, "Computer-Based 'Mental Set' Tasks: An Alternative Approach to Studying Design Fixation," ASME J. Mech. Des., **139**(7), p. 071102.

[12] Neroni, M. A., Oti, A., and Crilly, N., 2021, "Virtual Reality Design-Build-Test Games With Physics Simulation: Opportunities for Researching Design Cognition," Int. J. Des. Creat. Innov., **9**(3), pp. 139–173.

[13] Galle, P., 1996, "Design Rationalization and the Logic of Design: A Case Study," Des. Stud., **17**(3), pp. 253–275.

[14] Summers, J., 2005, "Reasoning in Engineering Design," International Conference on Design Theory and Methodology (DTM), Long Beach, CA, Sept. 24–28, pp. 329–340.

[15] Dorst, K., and Cross, N., 2001, "Creativity in the Design Process: Co-evolution of Problem–Solution," Des. Stud., **22**(5), pp. 425–437.

[16] van Someren, M. W., Barnard, Y. F., and Sandberg, J. A. C., 1994, *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*, Academic Press, London.

[17] Riskowski, J., Todd, C., Wee, B., Dark, M., and Harbor, J., 2009, "Exploring the Effectiveness of an Interdisciplinary Water Resources Engineering Module in an Eighth Grade Science Course," Int. J. Eng. Educ., **25**(1), pp. 181–195.

[18] Felder, R. M., and Silverman, L. K., 1988, "Learning and Teaching Styles in Engineering Education," Eng. Educ., **78**(7), pp. 674–681.

[19] Hay, L., Duffy, A. H. B., McTeague, C., Pidgeon, L. M., Vuletic, T., and Grealy, M., 2017, "Towards a Shared Ontology: A Generic Classification of Cognitive Processes in Conceptual Design," Des. Sci., **3**, p. e7.

[20] Okudan, G. E., and Tauhid, S., 2008, "Concept Selection Methods—A Literature Review From 1980 to 2008," Int. J. Des. Eng., **1**(3), pp. 243–277.

[21] Keeney, R., and Raiffa, H., 1993, *Decisions With Multiple Objectives: Preferences and Value Trade-Offs*, Cambridge University Press, New York.

[22] Pugh, S., 1991, *Total Design: Integrated Methods for Successful Product Engineering*, Addison-Wesley, Reading, MA.

[23] Budynas, R. G., and Nisbett, J. K., 2011, Shigley's Mechanical Engineering Design, McGraw-Hill, New York.

[24] López-Mesa, B., and Bylund, N., 2011, "A Study of the Use of Concept Selection Methods From Inside a Company," Res. Eng. Des., **22**(1), pp. 7–27.

[25] Pahl, G., Beitz, W., Feldhusen, J., and Grote, K.-H., 2007, *Engineering Design: A Systematic Approach*, Springer-Verlag, London.

[26] Hallihan, G. M., and Shu, L. H., 2013, "Considering Confirmation Bias in Design and Design Research," J. Integr. Des. Process Sci., **17**(4), pp. 19–35.

[27] Mattson, C., and Messac, A., 2003, "Concept Selection Using S-Pareto Frontiers," AIAA J., **41**(6), pp. 1190–1198.

[28] Martins, J. R. R. A., and Ning, A., 2020, *Engineering Design Optimization*, Cambridge University Press, Cambridge, UK.

[29] Mirabito, Y., and Goucher-Lambert, K., 2021, "Connecting Design Actions, Reasoning, and Outcomes in Concept-Selection," 33rd International Conference on Design Theory and Methodology (DTM), American Society of Mechanical Engineers Digital Collection, Virtual, Online, Aug. 17–19, p. V006T06A037, ASME Paper No. DETC2021-71830.

[30] Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., and Saleem, J., 2007, "Engineering Design Processes: A Comparison of Students and Expert Practitioners," J. Eng. Educ., **96**(4), pp. 359–379.

[31] Clegg, B. A., DiGirolamo, G. J., and Keele, S. W., 1998, "Sequence Learning," Trends Cognit. Sci., **2**(8), pp. 275–281.

[32] Otto, K. N., and Antonsson, E. K., 1993, "Tuning Parameters in Engineering Design," ASME J. Mech. Des., **115**(1), pp. 14–19.

[33] Weas, A., and Campbell, M., 2004, "Rediscovering the Analysis of Interconnected Decision Areas," AI EDAM, **18**(3), pp. 227–243.
[34] Raina, A., McComb, C., and Cagan, J., 2019, "Learning to Design From Humans: Imitating Human Designers Through Deep Learning," ASME J. Mech. Des., **141**(11), p. 111102.
[35] Raina, A., Cagan, J., and McComb, C., 2019, "Transferring Design Strategies From Human to Computer and Across Design Problems," ASME J. Mech. Des., **141**(11), p. 114501.
[36] Raina, A., Cagan, J., and McComb, C., 2022, "Design Strategy Network: A Deep Hierarchical Framework to Represent Generative Design Strategies in Complex Action Spaces," ASME J. Mech. Des., **144**(2), p. 021404.
[37] Savage, J. C. D., Moore, C. J., Miles, J. C., and Miles, C., 1998, "The Interaction of Time and Cost Constraints on the Design Process," Des. Stud., **19**(2), pp. 217–233.
[38] Hay, L., Duffy, A. H. B., McTeague, C., Pidgeon, L. M., Vuletic, T., and Grealy, M., 2017, "A Systematic Review of Protocol Studies on Conceptual Design Cognition: Design as Search and Exploration," Des. Sci., **3**, p. e10.
[39] Galle, P., and Béla Kovács, L., 1996, "Replication Protocol Analysis: A Method for the Study of Real-World Design Thinking," Des. Stud., **17**(2), pp. 181–200.
[40] Taylor, A. H., Elliffe, D., Hunt, G. R., and Gray, R. D., 2010, "Complex Cognition and Behavioural Innovation in New Caledonian Crows," Proc. R. Soc. B Biol. Sci., **277**(1694), pp. 2637–2643.
[41] Robert Mitchell, J., Shepherd, D. A., and Sharfman, M. P., 2011, "Erratic Strategic Decisions: When and Why Managers Are Inconsistent in Strategic Decision Making," Strateg. Manage. J., **32**(7), pp. 683–704.
[42] Mao, X., Galil, O., Parrish, Q., and Sen, C., 2020, "Evidence of Cognitive Chunking in Freehand Sketching During Design Ideation," Des. Stud., **67**, pp. 1–26.
[43] Li, M. S., Melville, D., Chung, E., and Stuart, H. S., 2020, "Milliscale Features Increase Friction of Soft Skin in Lubricated Contact," IEEE Robot. Autom. Lett., **5**(3), pp. 4781–4787.
[44] Mirabito, Y., and Goucher-Lambert, K., 2022, "Design Strategies That Work: How Engineers Use Sequential Decision Making to Improve Design Performance in Concept Selection," Design Computing and Cognition'22, Glasgow, UK, July 4–6, pp. 509–526.
[45] French, S., Maule, J., and Papamichail, N., 2009, *Decision Behaviour, Analysis and Support*, Cambridge University Press, Cambridge.
[46] Tversky, A., 1967, "Utility Theory and Additivity Analysis of Risky Choices," J. Exp. Psychol., **75**(1), pp. 27–36.
[47] Krishnakumar, S., Berdanier, C., McComb, C., and Menold, J., 2021, "Lost in Translation: Examining the Complex Relationship Between Prototyping and Communication," ASME J. Mech. Des., **143**(9), p. 091402.
[48] Lee, J., 1997, "Design Rationale Systems: Understanding the Issues," IEEE Expert, **12**(3), pp. 78–85.
[49] Regli, W. C., Hu, X., Atwood, M., and Sun, W., 2000, "A Survey of Design Rationale Systems: Approaches, Representation, Capture and Retrieval," Eng. Comput., **16**(3), pp. 209–235.
[50] Maier, A. M., Eckert, C. M., and Clarkson, P. J., 2005, "A Meta-Model for Communication in Engineering Design," CoDesign, **1**(4), pp. 243–254.
[51] Chen, X., Jia, S., and Xiang, Y., 2020, "A Review: Knowledge Reasoning Over Knowledge Graph," Expert Syst. Appl., **141**, p. 112948.
[52] Sarica, S., Luo, J., and Wood, K. L., 2020, "TechNet: Technology Semantic Network Based on Patent Data," Expert Syst. Appl., **142**, p. 112995.
[53] Goridkov, N., Rao, V., Cui, D., Grandi, D., Wang, Y., and Goucher-Lambert, K., 2022, "Capturing Designers' Experiential Knowledge in Scalable Representation Systems: A Case Study of Knowledge Graphs for Product Teardowns," 34th International Conference on Design Theory and Methodology, St. Louis, MO, Aug. 14–17, p. V006T06A032.
[54] Hao, J., Zhao, L., Milisavljevic-Syed, J., and Ming, Z., 2021, "Integrating and Navigating Engineering Design Decision-Related Knowledge Using Decision Knowledge Graph," Adv. Eng. Inform., **50**, p. 101366.
[55] Li, X., Chen, C.-H., Zheng, P., Wang, Z., Jiang, Z., and Jiang, Z., 2020, "A Knowledge Graph-Aided Concept–Knowledge Approach for Evolutionary Smart Product–Service System Development," ASME J. Mech. Des., **142**(10), p. 101403.
[56] Kwon, E., Huang, F., and Goucher-Lambert, K., 2021, "Multi-modal Search for Inspirational Examples in Design," 33rd International Conference on Design Theory and Methodology (DTM), Virtual, Online, Aug. 17–19, p. V006T06A020.
[57] Goel, V., 1995, "Cognitive Processes Involved in Design Problem Solving," *Sketches of Thought*, MIT Press, Cambridge, MA, pp. 95–126.
[58] Shergadwala, M., Bilionis, I., Kannan, K. N., and Panchal, J. H., 2018, "Quantifying the Impact of Domain Knowledge and Problem Framing on Sequential Decisions in Engineering Design," ASME J. Mech. Des., **140**(10), p. 101402.
[59] Sweller, J., 1988, "Cognitive Load During Problem Solving: Effects on Learning," Cogn. Sci., **12**(2), pp. 257–285.
[60] Daly, S. R., Yilmaz, S., Christian, J. L., Seifert, C. M., and Gonzalez, R., 2012, "Design Heuristics in Engineering Concept Generation," J. Eng. Educ., **101**(4), pp. 601–629.
[61] Kramer, J., Daly, S., Yilmaz, S., Seifert, C., and Gonzalez, R., 2015, "Investigating the Impacts of Design Heuristics on Idea Initiation and Development," Adv. Eng. Edu., **4**(4), pp. 1–26.
[62] Glikson, E., and Woolley, A. W., 2020, "Human Trust in Artificial Intelligence: Review of Empirical Research," Acad. Manage. Ann., **14**(2), pp. 627–660.
[63] Ezer, N., Bruni, S., Cai, Y., Hepenstal, S. J., Miller, C. A., and Schmorrow, D. D., 2019, "Trust Engineering for Human-AI Teams," Proc. Hum. Factors Ergon. Soc. Ann. Meet., **63**(1), pp. 322–326.
[64] Hu, W.-L., Akash, K., Jain, N., and Reid, T., 2016, "Real-Time Sensing of Trust in Human–Machine Interactions," IFAC-PapersOnLine, **49**(32), pp. 48–53.
[65] Song, B., Soria Zurita, N., Nolte, H., Singh, H., Cagan, J., and McComb, C., 2022, "When Faced With Increasing Complexity: The Effectiveness of AI Assistance for Drone Design," ASME J. Mech. Des., **144**(2), p. 021701.
[66] Camburn, B., Arlitt, R., Anderson, D., Sanaei, R., Raviselam, S., Jensen, D., and Wood, K. L., 2020, "Computer-Aided Mind Map Generation Via Crowdsourcing and Machine Learning," Res. Eng. Des., **31**(4), pp. 383–409.
[67] Camburn, B., He, Y., Raviselvam, S., Luo, J., and Wood, K., 2020, "Machine Learning-Based Design Concept Evaluation," ASME J. Mech. Des., **142**(3), p. 031113.
[68] Williams, G., Meisel, N. A., Simpson, T. W., and McComb, C., 2019, "Design Repository Effectiveness for 3D Convolutional Neural Networks: Application to Additive Manufacturing," ASME J. Mech. Des., **141**(11), p. 111701.
[69] Puentes, L., Raina, A., Cagan, J., and McComb, C., 2020, "Modeling a Strategies Human Design Process: Human-Inspired Heuristic Guidance Through Learned Visual Design Agents," Proceedings of the Design Society: DESIGN Conference, **1**, pp. 355–364.
[70] Zhang, G., Raina, A., Cagan, J., and McComb, C., 2021, "A Cautionary Tale About the Impact of AI on Human Design Teams," Des. Stud., **72**, p. 100990.
[71] Das, D., and Chernova, S., 2020, "Leveraging Rationales to Improve Human Task Performance," Proceedings of the 25th International Conference on Intelligent User Interfaces, Cagliari, Italy, Mar. 17–20, pp. 510–518.
[72] She, J., Neuhoff, J., and Yuan, Q., 2021, "Shaping Pedestrians' Trust in Autonomous Vehicles: An Effect of Communication Style, Speed Information, and Adaptive Strategy," ASME J. Mech. Des., **143**(9), p. 091401.
[73] Dong, A., Lovallo, D., and Mounarath, R., 2015, "The Effect of Abductive Reasoning on Concept Selection Decisions," Des. Stud., **37**, pp. 37–58.