The tools we use:
a study of user preferences for sketches, prototypes, and CAD models and the
influence on design outcome

Geoff Tsai

S.B. Mechanical Engineering, Massachusetts Institute of Technology, 2009
S.M. Mechanical Engineering, Massachusetts Institute of Technology, 2011

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Mechanical Engineering
at the
Massachusetts Institute of Technology
September 2016

© Massachusetts Institute of Technology 2016. All rights reserved.

Author Geoff Tsai
Department of Mechanical Engineering
August 22, 2016

Certified by Maria Č. Yang
Associate Professor of Mechanical Engineering
Thesis Supervisor

Accepted by Rohan Abeyaratne
Quentin Berg Professor of Mechanics
Chairman, Department Committee on Graduate Studies
The tools we use:
a study of user preferences for sketches, prototypes, and CAD models and the influence on design outcome

Geoff Tsai
Submitted to the Department of Mechanical Engineering on August 22, 2016 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mechanical Engineering

ABSTRACT
During a product design and development process, design teams use a variety of tools to generate and represent multiple design options before they eventually arrive at a singular design solution. Studying how these tools can influence the design outcome has the potential to enable designers to become more aware of the choices they make when they choose to use a tool. Because so much of the cost of a product is determined during the early stages of product development, reviewing these design alternatives with stakeholders is a valuable part of the process. Understanding how stakeholders respond to a design representation is vital to interpreting potential user feedback accurately.

In this dissertation, I investigate these questions about early-stage design process:

I. How does the use of a specific design tool influence the design outcome?
II. What design attributes do designers associate with these tools?

In Part I of this research, designers made sketches, foam model prototypes, or CAD models in a controlled experiment. Followed by a survey of the designs with potential users, Part I demonstrated how different design tools affect the quantity and quality of ideas, including that designs created as prototypes were recognized as more novel, more aesthetically pleasing, and more comfortable to use. In Part II, in a survey of designers with experience in sketching, foam models, and CAD, designs from Part I were re-represented as sketches. Experienced designers exhibited a better-than-random likelihood to identify the original tool used to create the design, despite viewing only the re-sketch. This suggests artifacts of a design tool persist in a design representation despite the design being translated from one medium to another.

Thesis Supervisor: Maria Yang
Title: Associate Professor of Mechanical Engineering
ACKNOWLEDGEMENTS

My doctoral work and experience as a graduate student are largely a product of the many stimulating and generous people I have had the good fortune to know as mentors and colleagues, friends and family, partners in work and play.

Thank you to my advisor, Maria Yang. Over the years you’ve been a perceptive guide for my research, career, and life. I could not have asked for a more positive, inspiring, and, frankly, super advisor. I wish all grad students could be as lucky.

Thank you to the members of my PhD committee. Warren Seering, our often philosophical conversations pushed me to think more deeply, as well as served as a reminder of why we do what we do. Barry Kudrowitz, your research and teaching have greatly influenced me. Your teaching of 2.00b, Toy Product Design, was a model of what was possible to do as a grad student. David Wallace, your commitment to teaching and continuous improvement, as well as your philosophy of working hard and having fun, are a lasting inspiration.

Thank you to my research collaborators, Catherine Elsen, Tomo Honda, and Anders Häggsman. Without you this research quite literally could not have been accomplished. Anders, I’ve so thoroughly enjoyed working with you over the past few years. I like to think some of your strong visual talents have rubbed off on me. Thank you as well to those who have helped out with the research at various stages, including Alison Olechowski, Catherine Fox, and Lena Yang.

Thank you to Justin Lai, for starting me on this path with 2.97 and DPD. As collaborator, instigator, sounding-board, and friend, you, most of all, have shaped my perspective on the graduate and undergraduate experience and how we can create the community at MIT we want to be in.

Thank you to Tiffany Tseng. We started our undergrad together and finished our PhD’s in the same year. In that time you’ve been a fascinating collaborator and a wonderful friend.

Thank you to Lindy Liggett, Steve Keating, and Jane Kokernak. Teaching 2.00b with you has been the highlight of my time in grad school, and I am fortunate to have shared that incredible experience with you and to consider you my friends.

Thank you to everyone, past and present, in the Ideation Lab and CADlab — including additional members of the cadlab-party mailing list. You’ve all created such an amazing, fun space that’s a joy to return to everyday.

Thank you to my family, and to Suzanne and Connie who have been like a second family, for your support and love. Lauren, my best friend and sister, you’ve provided encouragement for so many things in my life. You never got to see me begin grad school, but I know you’d have seen me finish.
For Lauren

What’s so amazing
that keeps us stargazing
and what do we think we might see?

Paul Williams & Kenneth Ascher
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Design Tools Comparisons</td>
<td>1</td>
</tr>
<tr>
<td>Product Perception</td>
<td>2</td>
</tr>
<tr>
<td>Representation Fidelity</td>
<td>3</td>
</tr>
<tr>
<td>Representation Mode</td>
<td>3</td>
</tr>
<tr>
<td>Evaluation of Design Concepts</td>
<td>4</td>
</tr>
<tr>
<td>Context for New Research</td>
<td>4</td>
</tr>
<tr>
<td>Methodology Part I</td>
<td>6</td>
</tr>
<tr>
<td>Experiment Design Overview</td>
<td>6</td>
</tr>
<tr>
<td>Design Participants</td>
<td>6</td>
</tr>
<tr>
<td>Design Task</td>
<td>7</td>
</tr>
<tr>
<td>Re-Sketching the Designs</td>
<td>8</td>
</tr>
<tr>
<td>Design Attributes</td>
<td>9</td>
</tr>
<tr>
<td>Survey of User Preferences</td>
<td>10</td>
</tr>
<tr>
<td>Survey Distribution</td>
<td>10</td>
</tr>
<tr>
<td>Survey Design</td>
<td>11</td>
</tr>
<tr>
<td>Survey Quality Control</td>
<td>14</td>
</tr>
<tr>
<td>Methodology Part II</td>
<td>16</td>
</tr>
<tr>
<td>Survey Design</td>
<td>11</td>
</tr>
<tr>
<td>Survey Design Part 1: Introduction</td>
<td>17</td>
</tr>
<tr>
<td>Survey Design Part 2: Review</td>
<td>19</td>
</tr>
<tr>
<td>Survey Design Part 3: Exit Survey and Score</td>
<td>21</td>
</tr>
<tr>
<td>Adaptive Choice-Based Conjoint</td>
<td>22</td>
</tr>
<tr>
<td>Design Attributes Extended</td>
<td>24</td>
</tr>
<tr>
<td>Survey Flow</td>
<td>25</td>
</tr>
<tr>
<td>Survey Logistics</td>
<td>26</td>
</tr>
<tr>
<td>Rewarding Participants</td>
<td>27</td>
</tr>
<tr>
<td>Demographics</td>
<td>27</td>
</tr>
<tr>
<td>Results Part I</td>
<td>30</td>
</tr>
<tr>
<td>Design Tools and Design Attributes</td>
<td>30</td>
</tr>
<tr>
<td>Design Tools and Top Designs</td>
<td>31</td>
</tr>
<tr>
<td>Design Tools, Design Attributes, and Qualities</td>
<td>35</td>
</tr>
<tr>
<td>Results Part II</td>
<td>39</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

1. Experiment structure showing interview and working sessions ............ 7

2. Examples of original designs paired with their re-sketched representations. Sketch is on the left, foam in the middle, and CAD on the right................................................................. 8

3. Logitech Harmony 650 as the baseline, "best-seller" remote ............... 9

4. Baseline remote as a sketch ........................................................................ 9

5. Sample survey page showing both re-sketch images and the user response area ......................................................................................... 13

6. Survey conceptual design with sample re-sketch image and question 16

7. Original CAD model with annotations, before being re-sketched ......... 16

8. Survey first page and introduction.............................................................. 17

9. Survey instructions first page, three different tools............................... 18

10. Survey instructions second page, additional contextual elements ...... 19

11. Survey re-sketch review ........................................................................ 20

12. Survey re-sketch review with answers filled........................................... 20

13. Survey exit with background questions ................................................... 21

14. Final survey page with accuracy score ...................................................... 22

15. Example re-sketch image with hidden attribute values.......................... 24

16. Example re-sketch image with different hidden attribute values........... 24

17. Sample page from attribute survey ............................................................. 25

18. Examples of “obvious” re-sketches............................................................. 26

19. Examples of “non-obvious” re-sketches ..................................................... 26

20. Sample pages from survey on a mobile device........................................ 27

21. Female / Male ratio of survey participants ............................................ 28

22. Survey participants’ jobs ........................................................................ 28

23. Survey participants’ ages ........................................................................ 29

24. Top 20 most creative designs, normalized by number of participants 32

25. Top 20 most comfortable looking designs, normalized by number of participants .......................................................... 33

26. Top 20 most aesthetically pleasing designs, normalized by number of participants .......................................................... 33

27. Average number of concepts per designer, error bars indicate +/- standard error ...................................................................................... 34

28. A participant’s top-ranked creative idea plotted against the number of ideas generated .......................................................... 35
29. Design with “novelty / other” form, high “creative / novel” quality and low “clarity” ................................................................. 36
30. Design with “buttons or touchpad” as input and low “creative / novel” quality ................................................................. 37
31. Design with “standard remote” form, perceived highly for “useful” and “comfortable” ................................................................. 37
32. Design with “smartphone / tablet” form and high “aesthetically pleasing” quality ................................................................. 38
33. Accuracy by students and professionals ................................................. 39
34. Accuracy by tool ........................................................................... 40
35. Accuracy by tool with yes/no response ................................................ 41
36. Accuracy by re-sketch image .............................................................. 42
37. Re-sketch image #74, originally a sketch ................................................ 42
38. Re-sketch image #29, originally a foam model ........................................ 43
39. Design #29, as a foam model .............................................................. 43
40. Example of re-sketch translation process, from original to re-sketch. Sketch is on the left, foam in the middle, and CAD on the right......... 44
41. Words used to describe designs created with sketching, foam, or CAD
LIST OF TABLES

1. Design attributes by category ........................................................................... 10
2. Spearman correlations between attributes and design tools. 
   Correlations are in Bold and p-values are in (). p<0.05 are shown in red. ................................................................. 30
3. Weighted rank accuracy of qualities ................................................................. 32
4. Variable importance of design tools and design attributes to design quality. Darker backgrounds are shown for greater magnitude .......... 35
5. Quality and tool correlations ............................................................................. 45
6. Attributes and tool correlations ........................................................................ 45
INTRODUCTION
During a product design and development process, design teams use a variety of tools to generate and represent multiple design options before they eventually arrive at a singular design solution. In this process they must often make choices between different design alternatives — weighing engineering, marketing, and user considerations — in order to select which design direction to pursue [43]. Review of these design alternatives may occur internally, with the design team, or externally, with a sample set of users or other stakeholders. Responses from this review can be fed back to the design process, affecting the design direction of future iterations. Because so much of the cost of a product is determined during the early stages of product development [5], reviewing design alternatives with stakeholders is a valuable part of the process.

In order to develop and represent these design alternatives, a design team might use a multitude of design tools ranging from paper-based 2D sketches, to digital models, to 3D physical prototypes. Use of these tools, and the representations they produce, have several implications involving, (1) the connection between the design representation and user preference, and (2) the connection between design attributes and the design tool.

Design Tools Comparisons
There is current research into understanding and comparing the use of different design tools, although there is less research into comparing physical prototypes to other tools. Acuna and Sosa [1], in measuring sketching and physical prototyping during a design exercise, observed tradeoffs with time spent during each activity, with greater prototyping time associated with higher-quality models, but greater sketching time associated with higher creativity. Viswanathan and Linsey [45] used different prototyping materials, including sketching, and found that the creation of physical models was linked to a higher percentage of functional ideas. In focusing on the psychological experience of design activity from the perspective of the designer, Gerber and Carroll [15] observed designers creating low-fidelity prototypes. In creating many quick, low-fidelity prototypes, designers were able to view failed ideas as examples of learning, and the set of ideas prototyped as an affirmation of their own creativity.

Much of the focus of design tool comparison research is on paper sketching versus digital tools. Black [7], in a survey of graphic design students, noted that compared to users of paper tools, users of digital tools admitted to compromising some of their initial ideas because of workflows in the software. Bilda and Demirkan [6] studied how interior
designers approached a design problem using traditional sketching tools versus 3D CAD tools and noted that users of traditional tools more frequently modified the goals of their solutions, thus CAD tools may be artificially limiting during the conceptual design phase. Stones and Cassidy [39] studied graphic designers working from paper sketches and digital drawings and observed that participants were less capable of reinterpreting the digital drawings into new ideas. Vasantha et al. [44] conducted a controlled study comparing designers using pen-on-paper (with digital recording), drawing tablet, or 3D CAD to solve design problems. Their results found use of different tools had an impact on how long each designer spent solving each problem, as well as the proportion of detailed design activity.

Product Perception

Product perception deals with how a product’s appearance and meaning are interpreted by the user. Many different and inter-related factors can influence product perception, and it is an area of interest for research. In the marketing research domain, Bloch [8] describes a conceptual model for how product form, as well as other factors, informs a consumer’s psychological and behavioral responses. Bloch also describes the importance of form and how form can be used to call attention to a product, how form can communicate a message, and how form can influence emotion. In addition to form, color may also have an influence on product perception. Mugge and Schoormans [27] varied the colors of 3D models and observed a positive association between the perceived novelty of a product and its level of performance. Crilly et al. [11] present a framework for consumer response to product appearance and that a product’s visual form can be interpreted, or misinterpreted, by the consumer. Petiot and Yannou [30] present a method to assess product semantics, measuring how well products suit the needs and tastes of consumers by mapping the product to a user’s perceptual space. Ahmed and Boelskifte [2] investigate what users perceive from a product about the designers’ intentions — through the use of form and color, what mood and associations are conveyed. Pérez Mata et al. [32] surveyed how perceptions and aesthetics of product relate to a consumer’s desire to own that product, including develop a set of guidelines to create a product for specific effect.

Perception can be divided into two components: the representation fidelity and the representation mode. Where fidelity is concerned with the level of detail and realism of the representation, and mode is concerned with the medium, e.g. sketches, physical models, photographs, 3D software models, etc. Different levels of fidelity can be achieved, for example, with sketches containing more or less detail, or from sketches that range from concrete to abstract. An effect of this range of concrete to abstract is the level of
interpretation required by the viewer: from a message “received” (zero interpretation) to a message that is “perceived” [26].

**Representation Fidelity**

Prior research has investigated aspects of the connection between the design representation and user preferences. Focusing on the quality of representation within a given representation mode, Kudrowitz, et al. [23] found a link between sketch quality and the perceived creativity of a product idea. Sauer and Sonderegger [36] tested the impact of prototype fidelity — including paper prototype, computer prototype, and fully-working interface — on a usability test, and found the computer prototype (medium-level fidelity) over-estimated the task completion time. Macomber and Yang [25], using sketches with varying levels of fidelity, found that crowdsourced-responses preferred high-fidelity sketches over low-fidelity sketches. Viswanathan and Linsey [46] observed that quick prototypes are associated with less premature fixation, potentially helping designers avoid the sunk cost fallacy by lowering the barrier to switch design directions. Yang [47], studying the performance of mechanisms in an undergraduate course, found that simpler prototypes were associated with a better course grade and design outcome. Hannah et al. [16] surveyed different modes of representation, although focused on the levels of fidelity of those representations, and found designers can extract more information and are more confident that a design meets requirements when viewing a high-fidelity representation. In addition, low-fidelity representations were faster to create and were effective at representing functional requirements, they did not provide enough information about geometric or manufacturing qualities.

**Representation Mode**

While fidelity can often be considered an independent variable of representation, different modes are often coupled with varying levels of fidelity due to the different mediums involved. Research into different representation modes often acknowledges these differences in fidelity, yet focuses on the different representations. Artacho-Ramirez et al. [4] observed in a case study that the representation mode — including, photography, infographic, 3D model, and stereographic 3D model — affected the ability to communicate the product’s value to a consumer. Söderman [40] compared representations of cars using sketches, virtual reality, and fully-functional models and noted participant’s level of understanding increased slightly with greater realism. Reid et al. [34] found participants provided inconsistent evaluations and preferences when viewing computer sketches, silhouettes, simple 3D renderings, and realistic 3D renderings. Tovares et al. [41] surveyed participants with sketches, 3D virtual reality models, and real samples of mugs, finding differences in preference from the representation modes. Macomber and
Yang [25], in the same study cited previously, found that crowdsourced-responses preferred high-fidelity sketches over CAD renderings. Additionally, their research indicates the importance of considering the amount of time involved in creating different types of representations, suggesting that minimizing the working time for maximal user preference is achieved with quick sketching.

**Evaluation of Design Concepts**

Conducting evaluation of design concepts is a critical part of the process for assessing design methods and design tools, as well as relevant to concept selection in a product design process. Evaluation can be performed individually, by a panel of judges, or by aggregating responses, possibly through crowdsourcing. Additionally, those evaluating may be novices or experts in the specific area of design.

Amabile [3] describes creativity as the combination of novelty and appropriateness, and that an idea is creative if independent observers can agree it is creative. Kudrowitz and Wallace [22] developed and tested a metric for evaluating large numbers of ideas at the early stage. Gathering crowd-sourced responses about idea creativity, novelty, and usefulness, they found a positive correlation between creativity and novelty, but insignificant correlation between creativity and usefulness. Tsai [42] studied idea clarity expressed in sketches and found less clearly expressed ideas scored neither highly nor poorly, suggesting participants were unwilling to score ideas they could not fully understand.

**Context for New Research**

Prior research has shown how the fidelity and mode of representation can influence perception, and in turn, the evaluation of ideas. Intuitively, we understand that a vast array of design tools is helpful in the design process, and so it is necessary we use these tools because of or despite their varying influence on perception. So far, less attention has been paid to studying how physical prototyping relates to sketching and 3D CAD in the early-stage design process, as well as how all three of these tools perform with regard to the qualities that matter to stakeholders.

Questions remain about how the mode of representation and fidelity of the representation affect the perception of the concept. There is more to learn about the tools used to create these representations: what role they play in the creative process, including exploration and development of a design space. Studying how these tools can influence the design outcome has the potential to enable designers to become more aware of the choices they make when they choose to use a tool. Understanding how stakeholders respond to a design representation is vital to interpreting potential user feedback accurately.
In this dissertation, I investigate these questions about early-stage design process:

I. How does the use of a specific design tool influence the design outcome?

II. What design attributes do designers associate with these tools?

This first question focuses on the important interaction between design tool and idea. One might conceptualize an idea or mental model as flowing from the mind, through the design tool, and manifesting in the real world as a design artifact. However, this question seeks to understand how the design tool, rather than merely serving as an instrument, has an influence on the idea itself. It is expected that choice of design tools will influence the number of ideas and exploration of the design space, possibly with some tools influencing what aspects a designer chooses to apply focus.

The second question builds upon the first by recognizing that designers are already likely aware of certain aspects of how the design tool can influence the design outcome. It seeks to understand where the separation exists, if any, between what design attributes can be demonstrated to be associated with a design tool, and what design attributes are assumed to be associated.

This dissertation documents my research in collaboration with the research team [18], and it extends it to include my additional analysis of the creativity of ideas, as well as a new study I conducted with designers and design tools. In each section of this dissertation, “Part I” refers to this first published study and largely addresses the first research question; “Part II” refers to the subsequent study addressing the second research question.
METHODOLOGY PART I

Experiment Design Overview
The experiment was designed to test for differences in quality and quantity of designs produced from sketching, prototyping with blue foam, and modeling in CAD using SolidWorks. Eighteen experienced designers were recruited to individually work on a 2.5-hour design task. In order to create a fair test to evaluate the designs created from different tools, all of the designs were re-sketched by a single artist in order to create a more uniform presentation of the designs. These re-sketch images were then presented to non-design-experts on Amazon’s Mechanical Turk platform for feedback about various qualities such as “creativity”, “aesthetics”, and “intent to purchase”.

Design Participants
Together with the research team, Prof. Maria Yang, Prof. Catherine Elsen, Anders Haggman, and Dr. Tomonori Honda, we recruited 18 experienced designers in the Boston and Belgium areas. Recruitment was conducted via email to design-related professional lists and design graduate student lists at the Massachusetts Institute of Technology. Participant age ranged 25-50 years old, and their professional-level design work ranged 2-25 years. In order to encourage participation and motivate each participant to produce best-effort work, all participants received $20 for their involvement in the study with a potential $75 awarded to the creator of the judged best design.

The experiment was structured around an initial interview and three working sessions each followed by an interview section focusing on questions about the designer’s work in the previous session. The experiment began with a brief, introductory interview where the designer’s prior experience was noted. Participants were assigned to use one of either sketching, foam modeling, or CAD based on their self-described experience during this first interview — ensuring that each participant was assigned a design tool they would be comfortable using. The final interview, while covering questions about the previous work session, also contained questions about the designer’s overall experience during the study. A pilot study for each of sketch, foam, and CAD was conducted with each work session lasting one hour, for a total of 180 minutes working time. Based on feedback from the designers, these working sessions were

1 www.mturk.com/mturk/
reduced in length to 40 minutes each, for a total working time of 120 minutes. This experiment structure is illustrated in Figure 1.

Design Task
In the description of the design prompt, designers were instructed to design a remote control for use with a living room home entertainment center. Users of the design would be a middle-class family of four, comprising two adults, one teenager, and one child. Designers could choose what device or devices the remote should control, and suggestions for those devices included a television, DVD player, digital video recorder (DVR), streaming console, gaming console, and computer.

Designers were instructed to produce any number of designs during the total working time. The fidelity of their designs — regardless of tools used — was left to the discretion of the designer. In communicating their designs, designers explained uses and features to the interviewer between work sessions.

For consistency, the same tools and materials were made available to each designer using sketching, foam, or CAD.

- Sketch participants were provided with paper sheets (letter-size for U.S., A4 for Belgium participants), pencils (2H, 2B, 4B, 6B, 8B), a pencil sharpener, an eraser, fine liner markers (0.1 mm, 0.3 mm, 0.5 mm, 0.7 mm), markers (1.0 mm and 2.0 mm), and a chisel-tip marker (10 mm).
- Prototype participants were provided with several blue foam blocks (precut in a range of sizes from 20 cm x 20 cm to 100 cm x 150 cm, and thickness ranging 3-10 cm), a collection of shaping tools (four rasps of different levels of coarseness), sandpaper (with grits of 50, 100, 150, and 220), 18-inch metal ruler, wooden toothpicks for joining foam pieces together mechanically, glue, a tabletop hot wire cutter with 5-inch span, and a chisel-tip marker.²
- CAD participants were provided with a desktop computer running the latest installation of SolidWorks modeling software.³

² Prototype participants were provided the marker solely for marking on the foam and were instructed they could not use the marker for idea generation or sketching concepts.
³ CAD participants were instructed to only use SolidWorks on the computer and not use other design programs.
Re-Sketching the Designs

In total, 83 designs were produced from the 18 sketch, prototype, and CAD participants. Sketches were scanned, prototypes were photographed, and CAD models were rendered out to still images. From previous studies, it has been observed how an idea is represented can influence the user’s perception of that idea and its quality [23, 25, 16]. In this experiment, we were interested in assessing the quality of ideas generated using sketching, foam prototypes, and CAD, but without the influence of the way those ideas would be represented. In order to measure the idea quality independent of the representation, we chose to have a single professional industrial designer re-sketch all of the designs. Re-sketches were created as sketches in Adobe Photoshop with the use of a drawing tablet.

If the design was originally created as a sketch, it was re-sketchered with all of the same information but in the style of the industrial designer. If the design was originally created as a foam prototype, the industrial designer created a new sketch in his style based from a photograph of the prototype. CAD models were sketched based on their still-image renders.

Because the original sketches often included annotations about the features or uses of the design, a similar level of detail was added to the re-sketches of foam prototypes and CAD models. The re-sketching process, as well as examples of annotations is shown in Figure 2. These added annotations were based on discussions during the interview sections of the design exercise. Rather than remove all annotations from all sketches, we chose to add these annotations to prototype sketches and CAD sketches in order to provide viewers of the sketches an aid in understanding the design and function, while at the same time providing a similar level of detail for all the re-sketches regardless of which tool was used to originally create the design.

Figure 2. Examples of original designs paired with their re-sketch representations. Sketch is on the left, foam in the middle, and CAD on the right
In addition to the 83 designs produced during the design exercise, we added one additional design to serve as a baseline reference for comparison. This particular remote, shown in Figure 3, was chosen as the baseline because of its “best-seller” status on Amazon.com at the time. Figure 4 shows this baseline remote re-created as a sketch.

Figure 3. Logitech Harmony 650 as the baseline, “best-seller” remote

Figure 4. Baseline remote as a sketch

Design Attributes

Reviewing all of the designs produced from the exercise, we noticed several recurring design attributes. For example, many of the designs featured user input via a touchscreen, or other designs looked similar to existing game controllers. Members of our research team independently generated a list of recurring design attributes. After generating separate lists, we worked together to merge the attributes to form a common set with three categories: Interaction, Form Factor, and Input Method.

Input Method describes the physical hardware of the remote that is intended as an affordance for the user to interact with the device. Form Factor describes the physical form of the remote control, what pre-existing form it visually resembles most. Interaction is distinct from Input Method in that it describes the primary type of interaction the user is required to use.
While a typical remote may only require hands to press buttons on the remote, another design may require certain body motion to trigger a specific action. This list of attributes and categories is shown in Table 1.

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Form Factor</th>
<th>Input Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>hands</td>
<td>standard remote</td>
<td>buttons</td>
</tr>
<tr>
<td>eyes</td>
<td>game controller</td>
<td>touchscreen</td>
</tr>
<tr>
<td>body</td>
<td>smartphone</td>
<td>joystick</td>
</tr>
<tr>
<td>novelty</td>
<td>mouse</td>
<td>scroll wheel</td>
</tr>
<tr>
<td></td>
<td>novelty</td>
<td>gestural</td>
</tr>
</tbody>
</table>

Table 1. Design attributes by category

Based on this set of attributes, we created a survey and administered it to a set of six expert design reviewers, who each completed the survey twice — the second time after waiting several months after the first. In the survey, the reviewers individually viewed an image of each design and were asked to note which attributes from the list most appropriately described the design. The design reviewers in this survey each had several years of experience in design practice, research, or both.

Responses from the reviewers were averaged per each image, providing the basis for a numerical score of each attribute. For example, a design could have anywhere from 0% to 100% form of a game controller. In testing the consistency of the reviewer responses, we chose Fleiss' Kappa as the inter-rater reliability metric because we surveyed six reviewers [13]. Landis Koch's criteria was used to assess that the reviewers exhibited significant inconsistencies in some of the attribute ratings. It was assumed this inconsistency was due to reviewers holding different definitions for the attributes. For this reason, we combined some of the attributes into attribute groups in order to improve consistency. The list of combined attribute categories is shown in Table 2 in [Results Part I].

Survey of User Preferences

Survey Distribution
The goal of this survey was to gather a relatively large set of responses from possible users about the quality and nature of these remote control design concepts. We chose Amazon Mechanical Turk as our survey distribution tool because of its large user base, with hundreds of thousands of users in over 100 countries [31]. On Mechanical Turk, users registered as a task provider can post a job — called a Human Intelligence Task, or HIT. Users registered as workers can view lists of HITs, organized
by categories and keywords, and choose which tasks to do. Once a worker completes a HIT and the provider reviews it, the worker receives compensation. Because of the large user base of Mechanical Turk, it is often an effective way of getting many responses to survey in a short amount of time.

One of the other advantages of using Mechanical Turk for a survey is the diversity of the subject pool and the controls for who is eligible to start a HIT. As Mechanical Turk is growing in popularity for its use in research fields, there have already been studies conducted about workers registered for the service [19, 35]. In particular, Paolacci et al. [28] compared surveys conducted on Mechanical Turk, to a typical online survey, to a university pool, and found comparable levels of reliable and unreliable responses. In addition, when the Mechanical Turk survey was restricted to the U.S., they found the pool to be at least as representative of the U.S. population as a traditional subject pool, and more representative than a typical university pool. Buhrmester et al. [9] found similar results, as well as that the amount of compensation for workers did not affect the quality of the data collected.

Survey Design

Using Mechanical Turk's administration controls we restricted the survey to U.S.-based workers with a 99% approval rating. We used online survey software from Qualtrics to design and administer our survey. This made it easy to design question templates and have the software substitute in re-sketch images from the 83 designs and one baseline remote design. One of the challenges of combining Qualtrics with Mechanical Turk is that Mechanical Turk workers leave the site when they participate in the survey hosted by Qualtrics. This might not be a concern for some survey administrators, but we wanted a way to confirm that workers completing a HIT in Mechanical Turk had in fact completed the survey on Qualtrics.

I designed the Mechanical Turk HIT page to append the unique Mechanical Turk worker ID to the link to the Qualtrics survey. Upon clicking the survey link for the Qualtrics survey, Javascript was used to extract the worker ID, append it to the link. Once at the Qualtrics survey, this worker ID was stored — invisible to the participant — as an embedded data field along with the other Qualtrics survey data for each respondent. In this way, when I reviewed completed HITs on Mechanical Turk's HIT management page, I could cross-reference the worker IDs with the worker IDs stored in the completed Qualtrics surveys.

As a fallback for when this would fail — a good precaution for a variety of reasons including different browsers, whether Javascript was enabled, and

---

4 www.qualtrics.com
different user behaviors — respondents were also asked to enter a confirmation number in the Mechanical Turk HIT. I designed the Qualtrics survey to create a randomly-generated number unique to each visitor and stored that number as an embedded data field alongside their other response data. This number was presented to the respondent at the end of the Qualtrics survey with instructions to paste it into the text-entry field of the Mechanical Turk HIT as a “confirmation number”.

With the Mechanical Turk worker ID and the confirmation number, I was able to determine which Mechanical Turk workers should or should not receive compensation based on the completeness of their Qualtrics survey. This approach for combining Mechanical Turk with Qualtrics surveys was based on work from Schubert [37] and Peer, et al. [29].

Rather than having survey respondents rate the designs on an absolute scale, we wanted to have all of the concepts ranked relative to each other. This avoids potential issues of priming where previously viewed designs affect the review of subsequent designs. However, with over 80 designs to review, we realized it would be an unreasonable and impractical task to have survey respondents complete this ranking. As a solution to this problem, we designed the survey to present a random subset of re-sketch images in randomly-generated pairs. Participants of the survey would view a pair and were able to respond with their level of preference for Concept A or Concept B using a 5-point scale, including: “strong preference for A”, “no preference either way, neutral”, and “strong preference for B”. An example page from the survey is shown in Figure 5.
We designed the survey to take about 15 minutes to complete, and initially each survey respondent would only view six pairs of images. However, from considering feedback from respondents and examining the average survey completion time, and after 204 responses were collected, we increased the survey length to include eight pairs for images. In total, 759 respondents started the survey, 506 completed it, and we accepted 406 after we checked these responses for quality. From the randomly created pairs of images that respondents viewed, each re-sketch image was rated from 58 to 78 times.

For any pair of re-sketch images, respondents used a 5-point scale to rate which of the two concepts:

- looks more useful
- looks more original/creative/novel
- looks more comfortable to use
- you would be more likely to buy (assuming they are similarly priced)
• looks aesthetically more pleasing (looks better)
• is presented more clearly (you understand how the device is meant to work)
• is a better idea (try to give an overall rating, all things considered)

This list of design quality measures was inspired, in part, by Garvin’s [14] eight dimensions of product quality: performance, features, reliability, conformance (consistency), durability, serviceability, aesthetics, and perceived quality. Because ideas would be rated by non-experts viewing a 2D sketch, the four metrics: reliability, conformance, durability, and serviceability were eliminated from the list. Performance was interpreted to include both “usefulness” and “comfortable to use”. “Presented more clearly” was included as a metric to record differences in idea interpretability and because of its use in prior research [23, 42]. “Original, novel, creative” was included in the list because of its importance to early-stage design thinking, as well as its presence in research measuring idea quality [3, 22].

Survey Quality Control
When collecting user-submitted data from anonymous human subjects, it is often necessary to check for quality responses and filter out illegitimate data. While we restricted the survey to Mechanical Turk workers located in the U.S. and with a 99% approval rating, we also designed-in four additional quality control measures:

• time recording
• instructions test
• preference test
• description test

One of the most basic methods we used to check for quality responses was how long each respondent took to complete the survey, as well as how long he spent on each individual page. Focusing on survey completion times that seemed implausibly short, we determined which participants were more carefully considering the questions and which were clicking through each page as fast as possible.

In the instructions test, participants were presented with information about the computer requirements of the survey and a description of what kind of designs they would be reviewing. On the following page, we required them to respond to three multiple-choice questions about the previous instructions. This first quality control measure served two

5 The survey was designed to not allow participants to go back to previous pages.
purposes. It provided data to us about the legitimacy of the participant's responses, and it also signaled to the participant that paying close attention to the survey would be required in order to perform well.

In the preference test, we added an eighth question to the set of seven questions asked at each pair comparison: “please click on the 'strong preference for B' option for this question”. An example of this can be seen in Figure 5. While this question appears in the third column from the right in this example, its position in the order of questions was randomized for each page of pair comparisons. The purpose of this question was to prevent participants from mindlessly clicking through the matrix of radio buttons and encourage them to read the question titles first.

In the description test, after a participant had completed responding to one of the pair comparisons, the following page would present a free-response area. The page would ask the participant to use the free-response area to describe the two concepts shown on the previous page. This description test page was presented twice during the survey to check that participants were paying attention to the images presented.
Survey Design
The survey was designed to test if people with design experience could identify the original tool by viewing the re-sketch. In its simplest form, this would consist of showing one of the re-sketch images and a multiple-choice question with options for "sketch", "foam", or "CAD", see Figure 6. For reference, Figure 7 shows the original CAD model with annotations before it was re-sketched.

Figure 6. Survey conceptual design with sample re-sketch image and question

Figure 7. Original CAD model with annotations, before being re-sketched

The survey was constructed as an online survey with three segments. In part one, participants would be given an explanation of the design and re-sketching process. In part two, participants would review a set of...
re-sketch images and make a judgement for each about which design tool was used to create the original. In part three, participants would be asked some questions about their experience and general observations.

Survey Design Part 1: Introduction

Figure 8 shows the first page presented to a participant when they enter the survey. At the top of the survey it explains approximately how much time will be required to complete the survey. It describes the survey completion reward: $20 amazon.com credit for every 10th participant. And it indicates that the participant will receive a score of their performance at the end of the survey. There are two text entry fields: one for name and the other for email address. The email address is validated for proper structure and will be used as a unique identifier in the response database (see Survey Part 2: Review). The final section of this page explains that participation is voluntary and that data will be anonymized prior to publication. It was a design goal of this first page to be as welcoming as possible — to be upfront and clear about the duration and reward system, to explain how their data would be used in a complete yet unobtrusive presentation. Additionally, it was intended that participants would be motivated to do well because they would receive their score at the end.

If participants chose to continue with the survey, they would be presented with two pages of instructions, Figure 9 and Figure 10. In the first page of instructions (Figure 9), the text and picture explained that new designs for remote controls had been created and each design could have come from either sketch, blue foam, or CAD. After the original design was created, it was then re-sketched.
In the second page of instructions (Figure 10), the text and picture explained that, in some instances, additional information was incorporated in the re-sketch to provide context. The purpose of this page was to explain to the participant that not everything in the re-sketch had to have been created with the original tool. So, for example, if the participant views a re-sketch and sees a hand using a product, they should not assume that, in addition to the product, the hand must have been created either in sketch, foam, or CAD.
Survey Design Part 2: Review

Figure 11 shows the page design for participants to review a re-sketch. During the survey, different re-sketch images would be substituted in place using this same page layout. The first question asked participants to identify the tool that was used to create the original. In the second question, participants were asked to respond "yes" or "no" if something about the design indicated to them a certain tool was used to create the original. Or, in other words, "was this an informed decision?". The third question is a followup to the second, and it asked participants to either identify what about the design indicated a certain tool or to suggest what kind of design detail would have helped to identify the tool.
While the third question was optional for participants, the first two questions were required. Only after a participant had completed both questions would the “Continue” button appear and allow them to proceed. This feature is shown in Figure 12.
Survey Design Part 3: Exit Survey and Score

Once a participant had completed reviewing 18 re-sketched images, she would be brought to the exit survey section, shown in Figure 13. This page would ask participants to respond and reflect on the designs displayed in the survey, "What about the designs in this survey seemed to indicate a certain tool was used?"; to indicate their current status as an undergraduate student, graduate student, practitioner / professional, or other; to describe their experience in professional fields such as product designer, mechanical engineer, industrial designer, etc.; and to specify their age range and gender.

Figure 13. Survey exit with background questions

Figure 14 shows the final page of the survey. On this page, participants were shown the score of their performance as a percentage. This percentage was calculated from the number of re-sketch images they identified correctly divided by the total. The first paragraph on the page explained what this percentage score meant. Later paragraphs provided links to related research being conducted by the MIT Ideation Lab as well as a way to contact the researchers. It was hoped that providing this context and a performance score at the end of the survey would provide a sense of reward for the participants.
Adaptive Choice-Based Conjoint

With 83 total re-sketch images needing to be reviewed, it would not be feasible to have individual participants review all 83 within a single survey. The more images to review, the more data would be collected per participant. Too many images to review, and the participant might feel fatigued or decide not to complete the survey. Several informal pilot studies were conducted in order to determine an ideal number of re-sketch images. From this testing, a set of 18 re-sketch images was chosen as an appropriate number, resulting in a survey that would take about 15–20 minutes to complete.

In initial designs of the survey, the set of 18 images would be chosen randomly from the total set of 83 images. However, it was expected that recruiting designers and engineers experienced with sketch, blue foam, and CAD would be difficult, and that we would possibly get around 75–200 participants. With 83 total images and a set of only 18 images being reviewed by a single participant, we were interested in a different method that would provide good quality data from a relatively small number of participants.

We chose to use Adaptive Choice-Based Conjoint (ACBC) as the method for choosing which 18 images to display to a participant. While typically used in market research studies, this method is useful for soliciting feedback on a large set of options with a minimal set of questions. This was of interest to us because of the relatively large set of images compared to the number of user responses we were expecting to receive.
When used in a market research context, ACBC varies what is presented to a participant based on their preference. Based on prior choices in the survey, ACBC will adapt to present the participant's most preferred attributes, resulting in a conjoint section that is more efficient with fewer "wasted" options that are unappealing. Furthermore, ACBC has the potential to craft a more engaging survey because it adapts to the preferences of the participant [33]. Sawtooth Software,6 makers of software for online surveys and analysis, state that Choice-Based Conjoint has been the most popular choice for conjoint-type surveys from the customers since 2000 [38].

In a typical adaptive conjoint study in a market research context, there are at least two phases. Phase one is the self-explicated section. During this phase, the participant indicates her preferences for a series of product attributes. This training section informs the adaptive algorithm of what to present in the next phase.

Phase two is the conjoint section. Here, the participant is shown a series of products, each with different sets of attributes chosen by the adaptive algorithm based on her responses from phase one. Each time she makes a selection of preference, the new data is incorporated into the model in order to improve the efficiency of the conjoint.

Typically, participants of an ACBC survey would choose what they prefer from a set of explicit options. In our survey, the goal is not what participants prefer, but whether they think a re-sketch image was originally created with sketch, foam, or CAD. In this way, "sketch," "foam", or "CAD" is the preference. However, without more information about the re-sketch image, we would not know what set of options this shows a preference for.

In order to adapt ACBC to this type of survey, re-sketch images would need to be tagged with a set of attributes describing the design and its representation. Only the image would be presented to participants of the survey. When a participant made a choice of "sketch," "foam", or "CAD" about a particular image, implicitly she would be asserting a preference for that tool associated with a particular set of attributes. Figure 15 and Figure 16 illustrate this idea with the participant-facing survey on the left and the attributes, normally hidden to the participant, on the right.

---

6 www.sawtoothsoftware.com
Design Attributes Extended

From Part I of this research, we had a set of 15 attributes that described aspects of the design including interaction, form, and input (see [Methodology Part I], Table 1). We reviewed the set of 83 images and took note of any possible detail someone might use to try to determine what tool created the original design. From this review, we extended the set of attributes to include, as examples: blockiness, “Does the general form take after a block-like shape?”, complexity, “Does the design contain complex geometry or high amounts of detail?”, people & hands, “Does the sketch show people or hands?”, wide geometric tolerances, “Does the sketch show wide geometric tolerances, with respect to button sizes/shapes, imperfect geometric forms?”.

In total, we identified 21 additional attributes relating to the idea and its presentation. We created a survey for all 83 images, and as a research team used a 3-point scale of “Yes”, “Neutral”, “No” to assign attribute values to each image. The images were presented in a randomized order. The survey was designed and administered using tools from Qualtrics.
Participating in this survey as a panel, we discussed and achieved consensus before assigning a value. An example page from this internal survey is shown in Figure 17. These 21 additional attributes were combined with the 15 attributes from Part I to create a 36-attribute description of the design and representation featured in a re-sketch image.

![Attributes Table]

**Figure 17. Sample page from attribute survey**

**Survey Flow**
As mentioned previously, ACBC requires a self-explicated phase to act as training for the adaptive conjoint. To accomplish this, six re-sketch images were chosen with the goal of spanning a large portion of the attribute space. The six images comprised two examples from each of sketch, foam, and CAD. And for each tool, one image was chosen as the more obvious example of that tool, and one image was chosen as the more non-obvious example of that tool. Figure 18 shows the three “obvious” examples of sketch, foam, and CAD. Figure 19 shows the three “non-obvious” examples.
For each participant, these same six re-sketch images would be displayed in a random sequence at the start of the survey in order to train the adaptive conjoint. As a consequence of these six images always being displayed to participants, they were reviewed many more times than other images in the set. During the survey, after the six training images were displayed, any 12 of the remaining 77 images would be displayed to the participant. Data from the participant’s preferences during the self-explicated section would inform what to display during the conjoint section. This model would continue to be updated during the conjoint section based on the participant’s choices.

Survey Logistics
The online survey was co-developed with Dr. Tomonori Honda who created the adaptive conjoint algorithm in Python and managed the survey and participant data in MySQL and MongoDB databases. Tornado was chosen as the web server so that pages with re-sketch images could be created dynamically. The web server ran on a Ubuntu machine hosted by Amazon EC2. In order to make it as convenient as possible for participants, responsive CSS layouts based on Bootstrap were designed for multiple screen sizes including desktop and mobile. Previous figures showed examples of the desktop survey layout; Figure 20 shows the survey layout on a mobile device (iPhone 6 with 1334 × 750 pixels).
Rewarding Participants
The survey introduction told participants that every 10th person would receive a $20 Amazon.com gift card. This was chosen as a dollar amount that would be appealing enough to respondents, without requiring the budget to pay every individual. In order to maintain a fast return time to respondents, every day new participant entries from the database were assigned a 10% chance to win. Winners were notified via their supplied email address and received a link to claim their reward.

Demographics
Invitations to participate in the survey were typically conducted via email. This includes mechanical engineering product design labs at MIT and undergraduate and graduate courses. We emailed out to our contacts in industry at mechanical engineering, product design, and industrial design positions. We also emailed out to the IDSA Boston network of professional members and posted to their chapter page on social media.

Over a period of about two months, this approach brought in 108 participants with complete responses to the survey. The next three figures will show the demographic breakdowns of those participants. Figure 21 shows that of 101 participants who chose to respond, 47% identified as female and 53% as male.
Figure 21. Female / Male ratio of survey participants

Figure 22 shows about one quarter were undergraduates and one quarter were graduate students. About half of participants were practitioners and professionals in engineering or design. A small percentage chose “Other”; based on comments written in the survey, these tended to be people with professional training or experience who were now working in a different field.

Figure 22. Survey participants’ jobs

Figure 23 shows a histogram of the approximate ages of the participants. With approximately half of the participants as practitioners and professionals and the other half as undergraduates and graduate students, a significant portion of the participants were in the 21–30 year range. The median age was in the 26–30 year range. If we assume a uniform distribution of ages within all of the bins, the average age was 29.4 years.
Figure 23. Survey participants' ages
RESULTS PART I

Design Tools and Design Attributes

From the design attribute data, we can observe if there are any associations between design attributes and other design attributes, as well as between design attributes and the design tool used. Table 2 shows Spearman correlations between design tools and design attributes, as well as the respective p-values. Note that correlations are in bold, p-values are in parentheses, and for p<0.05 the text is in red.

Table 2. Spearman correlations between attributes and design tools. Correlations are in Bold and p-values are in (). p<0.05 are shown in red.

Looking at Table 2, we can see that "standard remote and game controller" forms have a positive correlation with "buttons and touchscreen/touchpad" input and "joystick" input, as well as a negative correlation with "other/novelty" input. Intuitively, these correlations make sense; it would be expected that standard remote controls and game controllers would have these types of input, as well as not have some novelty type of input. As another example, "body and other" interactions had a positive correlation with "other" forms. This correlation also makes sense; you would expect interactions that involve some type of body movement, or perhaps brainwaves, would deviate from the standard remote, game controllers, smartphone, or mouse forms and integrate with some other type of form entirely. In addition "other" input showed a positive correlation with both "other" forms and "body and other" interaction. Taken together, these results — although not particularly insightful — make intuitive sense and help confirm the validity of the design attributes and correlations analysis.
Moving beyond the correlations we might expect to see between various types of forms, interactions, and input methods, we can look at how these design attributes correlate with the tool used to create the design. For designs that were sketched, these showed a negative correlation with "phone / tablet" forms. Designs that were made from foam prototypes showed a positive correlation with "mouse" forms. Designs that were created in CAD showed a positive correlation with having "buttons and touchscreen" input. One explanation for this is that CAD can be helpful at creating duplicated, fine detail such as an array of buttons.

CAD designs also showed a negative correlation with "other" forms. This result is interesting when considered in context with other research — referenced in the introduction of this paper — that found using CAD too early in the design process can prematurely limit a designer's concept exploration. As in the case of this author's findings, this other research suggested that use of CAD in the early stage of design is associated with less novel designs, possibly because of premature fixation.

**Design Tools and Top Designs**

Recall that the survey displayed pairs of re-sketch images and asked respondents to compare the two based on a set of qualities such as "looks more useful", "looks more original", etc. Because this data exists only as pair comparisons, we need to do some processing in order to know how each design scores, or ranks, with regard to the rest of the set. In order to do this, we applied Colley ranking and optimized ranking to this pair-comparison data.

Colley ranking is an algorithm developed for use in ranking teams in the U.S. College Football Bowl Championship Series. Sports analysts would want to know a team's performance relative to all the other teams, but of course the only data available would be single games: pair comparisons. In analyzing the data from our survey, each re-sketch image comparison is processed like the outcome of a football game and we apply the Colley ranking algorithm [10]. As more data is collected from image-pair comparisons — more games played — the accuracy of the rankings improves. For two images that are measurably equivalent, the ranking accuracy should be near to 50%. This ranking accuracy is calculated as follows:

\[
\text{ranking accuracy} = \frac{\text{count of higher-ranked concept wins}}{\text{total number of comparisons}}
\]

Alternatively, it is possible that it is the people, not the tool; designers who choose to use CAD in the early stage simply may be less creative. However, addressing this question is beyond the scope of this research.
For the optimized ranking, we applied a brute force heuristic optimization technique to the Colley ranking so that ranking accuracy is maximized. Table 3 shows the weighted accuracies of Colley and optimized rankings. This portion of the analysis, including the Colley and optimized rankings, was conducted by Dr. Tomonori Honda in collaboration with this research.

<table>
<thead>
<tr>
<th>Useful</th>
<th>Creative</th>
<th>Comfortable</th>
<th>Purchase</th>
<th>Aesthetic</th>
<th>Clarity</th>
<th>Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colley Rank Weighted Accuracy</td>
<td>0.779</td>
<td>0.716</td>
<td>0.760</td>
<td>0.756</td>
<td>0.731</td>
<td>0.764</td>
</tr>
<tr>
<td>Optimized Rank Weighted Accuracy</td>
<td>0.817</td>
<td>0.751</td>
<td>0.791</td>
<td>0.793</td>
<td>0.776</td>
<td>0.803</td>
</tr>
</tbody>
</table>

Table 3. Weighted rank accuracy of qualities

In a product design process, with a concept selection phase that follows an ideation phase, we are often less concerned with the average output ideation and more concerned with the top results. Using the optimized ranking results, we can look at the top designs and what tools created them. Figure 24, Figure 25, and Figure 26 show the top 20 ranked designs from the optimized rankings. During the design exercise phase of this experiment, there were a different number of participants for sketch, prototype, and CAD. In order to normalize for this, the number of designs in the top ranking was divided by the number of participants for that tool. This normalized result is presented as a percentage of the whole in the figures.

Creative
Top 20

Figure 24. Top 20 most creative designs, normalized by number of participants
Figure 24, Figure 25, and Figure 26 all show foam prototypes as the majority of top-ranked “creative”, “comfortable”, and “aesthetically pleasing” designs. This result was somewhat surprising as we had expected sketching to have produced a larger share of the top-ranked ideas. This expectation was built on the assumption that sketching would allow for designers to work quickly to generate many ideas — many ideas being associated with more creative ideas.

However, in examining the sketches produced, many of them resembled slower, more detailed “communication” sketches rather than faster...
“thinking” sketches. While sketching could have provided designers an ability to generate many ideas quickly, it tended not to be used this way. At the same time, many of the foam prototypes produced exhibited low detail — presumably because higher detail would have been difficult to create with the tools provided — and were faster to create. Figure 27 shows that the number of ideas per designer was highest for foam prototypes, then followed by sketching. Although designers were not specifically instructed to create as many ideas as possible, it is worthwhile to note that users of foam prototyping did appear to use the tools in this way while users of sketching did not.

Other research discussed previously [22] has demonstrated a strong correlation between the number of ideas generated and the participant’s creativity score, and this research supports this finding. Figure 28, on the horizontal axis shows the number of ideas generated by an individual participant; on the vertical is plotted the rank of the most creative idea from that participant.9

---

9 There were no apparent trends for any of the other qualities and number of ideas.
Design Tools, Design Attributes, and Qualities

Previously, we have explored the connections between design tools and design attributes and design tools and quality. In this section, we will examine the connections between all three: design tools, design attributes, and qualities. Table 4 shows the results of discrete choice modeling to calculate the variable importance of design tools and design attributes with respect to design quality.

![Figure 28. A participant's top-ranked creative idea plotted against the number of ideas generated](image)

Table 4. Variable importance of design tools and design attributes to design quality.

Darker backgrounds are shown for greater magnitude.
In Table 4, each cell of a given column can be compared. For each column, the background shading color is keyed like a heat map and best indicates order-of-magnitude changes. While it is useful to compare cells within the same column, different columns should not be compared to one another.

Although columns should not be compared to each other, it is worth noting that in every column except for one, designs with “novelty / other” forms were perceived very negatively. The only instance where “novelty / other” form was perceived positively was for the “creative” measure. That “novelty / other” forms were perceived so negatively across the table suggests that novel form alone is not enough to create user preference.

Looking at a more specific example of this, consider “novelty / other” forms being perceived negatively for clarity. This could be explained by “novelty / other” forms being more difficult to communicate their use. Figure 29 shows a design with “novelty / other” form, as well as high “creative / novel” quality, and low “clarity”. So while respondents rated this design as creative, it is not difficult to imagine how operating the remote might seem unclear.

Figure 29. Design with “novelty / other” form, high “creative / novel” quality and low “clarity”

Where “creative” was perceived negatively was for designs that used “buttons & touchpad / touchscreen” as the method of input. Figure 30 shows an example design that uses this type of input and was rated low for “creative / novel” quality.
Designs with "standard remote & game controller" forms were perceived highly for "useful" and "comfortable" (Figure 31). It is understandable that participants would be comfortable with these standard forms with which they are already familiar.

Designs with a "smartphone / tablet" form were perceived highly for "aesthetically pleasing". An example of this is Figure 32. In addition, "smartphone / tablet" form designs were also perceived highly for "creative", "more likely to buy", and "overall better idea". It is not obvious why designs of this form were perceived so positively for all of these qualities, but it perhaps could be due to participants being generally enthusiastic about smartphones and tablets, rating such-based designs positively as a result.
Figure 32. Design with “smartphone / tablet” form and high “aesthetically pleasing” quality
RESULTS PART II

Accuracy — Whole Population, Students, and Professionals

Given a 3-choice question, the average accuracy would be 33% if participants responded purely randomly. If we had observed this result, this would have perhaps indicated there were no discernible differences amongst the re-sketches, even though the designs had begun independently with either sketching, foam prototyping, or CAD. However, we would expect participants to do better than randomly guessing, particularly professionals with significant experience using these tools. In the results from this study, participants correctly identified the original tool used to create the design with 48% accuracy on average — likely indicating some artifact of the original design tool had been passed along during the re-sketching process.

Participants of this survey ranged from first-year engineering and design students to professionals with decades of experience in industry. How would different levels of experience affect their ability to correctly identify the original tool? Figure 33 shows the participants' averages broken out into groups for “undergraduate students”, “graduate students”, and “professionals”. Recall that “other” was used to describe people with professional training who were currently working in a different field. “Other”, as well, represented a small portion of the participants. Therefore, because of their similarities and relative sizes, “professionals” and “other” were combined into one group.
As Figure 33 shows, the differences in accuracy between the three groups is not statistically significant. This would seem to suggest that experience beyond a certain level does not improve an individual's ability to correctly identify the original tool. In future work, it could be informative to re-run the survey with participants who have no prior experience with these tools. Likely, these participants would perform worse than 48% accuracy, potentially equivalent to random 33% accuracy.

**Accuracy — Sketch, Foam, CAD**

We saw previously that undergraduate students, graduate students, and practitioners performed with highly similar levels of accuracy. What about the different tools: sketch, foam, and CAD? Is one tool easier, or more difficult, to identify accurately? Figure 34 shows accuracy as it is segmented by tool, rather than by experience. When viewed this way, there appears little difference between participants' ability to identify sketch, foam, or CAD.

In the survey, each time a participant was presented with an image and asked to identify if it was originally produced via sketch, foam, or CAD, there was a follow-up question about how this decision was informed. In the survey, this follow-up question read, “Was there something about the design that indicated to you a certain tool was used to create the original?”. Participants were given the option to respond “yes” or “no”. Further analysis of the accuracy to identify each tool split the responses for each tool based on whether or not a participant had indicated either “yes” or “no” that “something about the design [indicated] a certain tool was used”.

![Figure 34. Accuracy by tool](image-url)
Figure 35 shows this data with "yes/no" responses paired by tool, with the columns on the left of each pair, in more saturated color, representing "yes" and the columns on the right of each pair, in less saturated color, representing "no". Looking at the data this way, it is somewhat surprising to see that accuracies for sketch and CAD do not show a significant difference when split across "yes/no" responses, and that foam does appear to show a significant difference.

![Figure 35. Accuracy by tool with yes/no response](image)

Initially, it was expected that for all three tools accuracy would be higher for "yes" responses and lower for "no" responses. The breakdown for foam does exhibit this, however it is not clear why sketch and CAD appear to show no difference in accuracy between "yes" and "no" responses.

**Accuracy — Individual Concepts**

So far, this analysis has looked at the accuracy of participants with different levels of experience and the accuracy of identifying one tool over another. However, what about individual images? Do participants more reliably identify the original tool for some images rather than others? Figure 36 shows all the images presented during the survey, ranked in order from least-accurately identified to most-accurately identified.
The chart shows that the images do vary greatly in terms of what is accurately identified, spanning the entire range from 0–100%. Although the average is that participants correctly identified the original tool with 48% accuracy, clearly some images were easier to identify than others. What makes these images different?

As an example, image #74 was correctly identified as a sketch 13/13 times, see Figure 37. Upon inspection, the image does appear to exhibit certain attributes we might expect of a sketch: looseness of the forms, a roughly two-dimensional appearance, and certain organic and delicate forms (the wires) that might be difficult to produce in foam or CAD.

In another example, image #29 was correctly identified as foam only 2/108 times, see Figure 38 (note, this image was part of the adaptive conjoint training set, which is why it was reviewed by every participant of the survey). Upon inspection, this image appears atypical of what one might expect of a foam model. It exhibits a large amount of detail, complexity, and delicate features that might be difficult to produce in foam. Based on this information, the design looks more likely to have been produced as a sketch. In fact, 101/106 participants agreed — however incorrectly — that this image was a sketch.
Overall, what we’re seeing with these accuracies is the result of multiple, somewhat independent, factors. One, because participants correctly identified the original tool with 48% accuracy, some amount of detail from the original design and representation is being carried along into the re-sketched. This gives participants with some degree of familiarity with sketch, foam, and CAD a better-than-random likelihood to identify the original tool.

Two, some designs may be atypical of other designs produced with the same tool. Table 6 of [Attributes — Correlation with Surveyed Tool and Actual Tool] goes into detail about what attributes participants associate with a specific design tool, as well as what attributes are associated with the actual tool used in the experiment. However, it is clear that not every design exhibits all the attributes associated with the original tool, nor doesn’t exhibit all the attributes not associated with that tool.

Three, some re-sketched are misleading. In the translation process of taking an existing design — whether sketch, foam, or CAD — and creating a new sketch, some amount of interpretation was required on the part of the artist. Inspection of the set of re-sketched and comparing those to their original counterparts showed that three designs exhibited more significant interpretation than others. The best example of this kind of more-significant interpretation can be seen in Figure 39.
Figure 39 shows the original foam model for comparison with Figure 38. Focusing on the car, you can see the sketch has added significantly more detail to both the exterior and interior of the car. Based on the notes for the foam model, the sketch also added contextual environmental elements including the TV and a coin for scale. As mentioned previously, this image was correctly identified as foam only 2/108 times. Of the set of 83 images, three images, including this one, could be described as receiving more significant interpretation during the re-sketching process. For the two other images, both originally foam prototypes, one was more often mis-identified as a sketch, and the other was mis-identified equally as either a sketch or CAD model.

Compared to Figure 39 and the amount of interpretation, Figure 40 shows how this process is supposed to proceed with examples of sketch, foam, and CAD. In this process, annotations from sketches are carried over to the re-sketch, while notes about the foam model or CAD are translated to annotations in the re-sketch.

Quality — Correlation with Surveyed Tool
In Part 1 of this study, each image received a score for qualities such as “useful”, “creative”, “comfortable”, “intent to purchase”, “aesthetics”, “clarity of the idea”, and “overall better idea”. From this data, foam was associated with generating ideas with high creativity, comfort, and aesthetics. CAD was not directly related to low creativity, however it did produce the fewest number of ideas and more ideas was associated with higher creativity.

However, in this study, participants were asked what was the original tool used for creating the design, i.e. what was the surveyed tool. Table 5 shows correlations between the surveyed tool and the qualities data collected previously. There is only one correlation with a p-value less than 0.05, and it is noted by bold text: the negative correlation between CAD and...
creativity. Previously, this direct association between CAD and low creativity did not exist before in the original study. However, this new result seems to indicate, regardless of whether CAD actually produces less creative designs, that participants seem to think it does.

<table>
<thead>
<tr>
<th></th>
<th>Useful</th>
<th>Creative</th>
<th>Comfort</th>
<th>Purchase</th>
<th>Aesthetic</th>
<th>Clarity</th>
<th>Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sketch</td>
<td>-0.130</td>
<td>-0.037</td>
<td>-0.216</td>
<td>-0.159</td>
<td>-0.048</td>
<td>-0.130</td>
<td>-0.137</td>
</tr>
<tr>
<td>Foam</td>
<td>+0.061</td>
<td>-0.096</td>
<td>+0.096</td>
<td>+0.114</td>
<td>+0.144</td>
<td>-0.059</td>
<td>-0.076</td>
</tr>
<tr>
<td>CAD</td>
<td>+0.199</td>
<td>-0.260</td>
<td>+0.030</td>
<td>-0.184</td>
<td>-0.167</td>
<td>+0.164</td>
<td>+0.179</td>
</tr>
</tbody>
</table>

Table 5. Quality and tool correlations

This distinction between reality and perception of reality — what tool was actually used to create the original versus what tool participants chose as the original tool — is important beyond just seeing where there is agreement or disagreement. It can highlight what biases designers and engineers may have about tools and how designs are represented.

Attributes — Correlation with Surveyed Tool and Actual Tool

In this section, we continue with the idea introduced in the previous section: the distinction between reality and perception of reality. Table 6 shows correlations between attributes of the re-sketch and the surveyed tool, as well as attributes of the re-sketch and the actual tool used.

<table>
<thead>
<tr>
<th></th>
<th>SKETCH</th>
<th>FOAM</th>
<th>CAD</th>
<th>ACTUAL</th>
<th>SKETCH</th>
<th>FOAM</th>
<th>CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocky</td>
<td>0.254</td>
<td>0.247</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Rounded</td>
<td>0.292</td>
<td>0.247</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Planar faces</td>
<td>0.292</td>
<td>0.247</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Complementary</td>
<td>0.234</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Symmetry</td>
<td>0.234</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Large scale</td>
<td>0.234</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Delicate parts</td>
<td>-0.258</td>
<td>0.233</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Hinges</td>
<td>0.234</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Chamfer</td>
<td>0.234</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Detail out of plane</td>
<td>-0.222</td>
<td>0.232</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Buttons</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>People &amp; hands</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Multiple views</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Ground plane</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Sketch hairs</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>2D appearance</td>
<td>0.464</td>
<td>0.222</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Shading</td>
<td>-0.481</td>
<td>0.336</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Stylized text</td>
<td>0.328</td>
<td>0.413</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Wide tolerances</td>
<td>0.328</td>
<td>0.413</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
<tr>
<td>Isometric view</td>
<td>0.328</td>
<td>0.413</td>
<td>0.233</td>
<td>0.366</td>
<td>0.356</td>
<td>-0.328</td>
<td>0.313</td>
</tr>
</tbody>
</table>

Table 6. Attributes and tool correlations

In the vertical direction, Table 6 lists 21 of the image attributes related to how the idea was represented as a sketch. The first set of three columns contains Spearman correlations for sketch, foam, and CAD for the
surveyed tool (what tool surveyed participants chose as the original tool). The second set of three columns contains correlations for sketch, foam, and CAD for the actual tool used to create the designs.

Cells with bold numbers are correlations with a $p$-value less than 0.05. One way to read this table is that if participants of the survey were extremely accurate in their choices for what tool was used to create the original, then all of the correlations for surveyed tool would match the correlations for actual tool. Conversely, if survey participants were extremely inaccurate in their choices, then none of surveyed tool correlations would match the actual tool correlations.

Although we are typically interested in where there is a significant correlation, it can be equally interesting where there is no significant correlation. This leads us to a second way to read this table: pick a specific attribute and look at the related surveyed and actual correlations with sketch, foam, and CAD.

For example, pick the attribute “blocky”. “Blocky” showed a negative correlation with sketching, a positive correlation with foam, and no significant correlation with CAD amongst the tools participants picked in the survey. By comparison, “blocky” shows a very similar set of correlations with the actual tool used. Interpreting this, designers and engineers tend to associate foam with creating blocky designs, sketching with creating non-blocky designs, and no association of CAD and blocky designs. (This last result of no correlation between CAD and blockiness is somewhat surprising given our initial assumption that participants would associate CAD with blockiness.) These assumptions, on the part of the surveyed participants, about tools creating design attributes are supported by reality. Given the correlations between actual tool used and blocky design, “blocky” would be productive as an indicator.

As another example, pick the attribute “chamfer”. The surveyed tool showed a positive correlation with CAD and no correlation with either sketch or foam. However, this is in contrast to the actual tool which showed no correlation for any of the tools. This suggests that participants assumed a relationship, that chamfers were likely from CAD and not from other tools, when in actuality no such association exists. Or put differently, in the design process, presence or absence of chamfers did not correlate with sketch, foam, or CAD, yet surveyed participants assumed that chamfers were more often from CAD and not from other tools. One possible explanation for why participants made this assumption is the ease with which chamfers are created in CAD — often a matter of a click of a button.
Qualitative — Why a Specific Tool Was Chosen

In the previous section we looked at correlations between image attributes and the surveyed tool as well as the actual tool. This allowed us to see what attributes participants associate with what tools and what attributes are associated with the original tools used. It allowed us to intuit where participants’ assumptions align or misalign with reality. In this section, we look explicitly at these assumptions by examining the free, written responses from participants.

At the end of the survey, one of the questions participants were asked was “What about the designs in this survey seemed to indicate a certain tool was used?”. This free-response question was open-ended and optional. Out of the 108 participants who completed the entire survey, 46 provided helpful, descriptive responses of what specific details they associated with what tools. For example, one engineer wrote:

Careful attention to ergonomics suggested foam. Complex free-form designs seemed to suggest sketching, just because I figure using foam or CAD in those cases would have pushed the design to something more rigid.

In another example, a product designer wrote:

- Organic shapes -> Sketches, blue foam
- Simple, geometric or parametric-looking shapes -> CAD
- Things that fit in your hands -> blue foam

These were some of the better, more complete responses that dealt with all three tools, and there were many more like these. Overall, the responses explicitly described the assumptions and biases participants had for sketch, foam, and CAD.

In order to get a sense for trends in this free-response section, the text was analyzed for tools mentioned in conjunction with physical or representational attributes. For example, a professional fabricator wrote:

Some of the forms that were blocky seemed to be a function of foam or CAD which can be more difficult to get free form shapes.

This response was then manually coded as \texttt{foam:blocky} and \texttt{CAD:blocky}. In other cases, some responses were interpreted towards a phrasing more common with others. For example, one graduate engineering student wrote:

The more ‘out there’ an idea was, the more likely I was to think it was originally a sketch.
Since other responses mentioned “novelty” or “creativity”, in order to make this response conform to the majority — and to make it more concise — it was coded as sketch:novel.

All of these coded responses were collected, each attribute counted, and sorted by tool in order to see what attributes were discussed in reference to that tool. Figure 41 show the relative mentions of an attribute with respect to that tool.

Looking at the sketch area of Figure 41, “loose”, “complex”, “organic”, “concept”, and “novel” were among some of the most mentioned attributes. In responses from participants, “concept” meant that image tended to focus on the concept of an idea more so than the physical or tactile nature of it.

In the foam area of Figure 41, “organic”, “ergonomic”, “tactile”, “cut”, “interaction” were among some of the most mentioned attributes for foam. In responses, “cut” referred to the idea that some of the designs appeared more obviously to have been created using a cutting tool. “Interaction” referred to images that showed hands or some kind of human body interacting with the design. It is interesting to note that for foam, attributes mentioned include “organic” and “ergonomic” but also “cut”, “blocky”, and “geometric”. These two sets of attributes seem almost at odds with each other, and likely due to the different ways in which
foam can be worked: cut with a knife or hot wire versus shaped by sanding.

In the CAD area of Figure 41, CAD attributes mentioned included “geometric”, “precise”, “complex”, “extrude”, and “blocky”. Focusing on “geometric” as an attribute, it is interesting this appeared as the most commonly mentioned attribute associated with CAD. Certainly, geometric forms are possible, without much difficulty, to create with sketch or foam. From the written responses, participants mentioned how it was easy to create geometric or blocky forms in CAD:

...more solid/symmetric geometries seemed to be easier to create in CAD

Or, the converse was mentioned: that non-geometric forms, i.e. “organic”, are often difficult:

...forms that were blocky seemed to be a function of foam or CAD which can be more difficult to get free form shapes

planar/regular shapes were more likely foam or CAD, while more organic shapes are easier with a sketch

Overall, common occurrences in these written responses support some of the same associations as the correlations between surveyed tool and attribute. For example, the relatedness of “loose” and “wide tolerances” for sketch, “blocky” and “blockiness” for foam, “precise” and a negative correlation of “wide geometric tolerances” for CAD. That these two sets of data — quantitative and qualitative — are in agreement suggests that participants were self-aware of their choices for surveyed tool and had conceptualized their personal tool criteria specifically enough they could identify that criteria in the written free-response section.
SUMMARY PART I

Design Productivity
During the design activity, foam prototyping produced more ideas per designer, resulting in faster generation of ideas than sketching or CAD. Participants who made foam prototypes tended to create quick prototypes with relatively low levels of detail. In contrast, while sketching is generally considered a quick ideation generation tool, participants tended to create high-detail “communication” sketches rather than fast, low-detail “thinking” sketches. Because sketching can be used in these different fashions, it should be the level of fidelity of the design representation that is considered most relevant to the effects on the speed and quantity of the designs produced.

Creativity
Participants who generated more ideas tended to create designs that were more creative. This may not rule out the possibility that the designers who generate the most creative ideas are also the most prolific, but it does support the popular aphorism “the best way to have a good idea is to have a lot of ideas”.  

Design Tool and Quality
Of the top-20 ranked concepts for each quality measure, foam prototypes produced the majority of top-ranked creative, comfortable, and aesthetically pleasing designs. By comparison, CAD designs were negatively correlated with novel physical forms, possibly because of the constraining nature of using the tool during the early stage.

Design Attributes and Quality
Novel form alone is not sufficient for achieving positive user preference. Novel physical embodiments were perceived negatively across all quality measures except for the measurement of creative/novel. By comparison, standard remote and game controller forms were perceived as useful and comfortable, while smartphone and tablet forms were perceived as novel, likely to be purchased, and aesthetically pleasing. One of the fundamental tenets of an innovation-focused design process is that more creative ideas

10 Linus Pauling
will lead to an improved design outcome [20]. This result is not a contradiction, but provides some nuance: novelty may be a necessary component for design success but not to the exclusion of other qualities.
SUMMARY PART II

Accuracy
In the survey, participants demonstrated 48%, better-than-random accuracy at identifying the original design tool used. This suggests not only that the design tool imparts a design with certain attributes, but also that those attributes often are maintained through the re-sketching process. Segmenting the participants by experience (undergraduate student, graduate student, and professionals) revealed that beyond a minimal level, additional experience did not increase a participant’s accuracy.

Quality Correlation with Surveyed Tool
In the original study, the plurality of top-rated designs for creativity, comfort, and aesthetics came from foam prototypes. In comparison, low creativity correlation with CAD as the tool chosen during the survey.

Attribute Correlation with Surveyed Tool and Actual Tool
Correlations between attributes and the surveyed tool showed that participants perceived blockiness as being positively correlated with foam and negatively correlated with sketch. In comparison with the actual tool used to create the original designs, this same positive/negative correlation suggests that participants’ assumptions about blockiness and tools is accurate.

Participants’ responses showed wide geometric tolerances positively correlated with sketch and negatively correlated with CAD. In this case, participants’ assumptions are not supported by reality where there is no significant correlation between wide geometric tolerances and any tool.

Written Responses about Attributes and Tools
In the written responses from participants about what attributes indicated to them a certain tool was used to create the original design, common words or concepts demonstrated similar thinking and assumptions about the capabilities and artifacts or sketch, foam, and CAD. Trends in the written responses support some of the same associations as the correlations between attributes and surveyed tool.
Tool and Designer

For the designer using sketching, foam, and CAD to develop and represent ideas, it’s important to consider how the chosen design tool can affect the design outcome. Furthermore, the influence of the design tool on a design concept persists, even when the concept is re-represented as a new sketch. This is relevant for designers who may want to dissociate the idea from the tool that created it.

Designers are aware of and can articulate design artifacts from sketch, foam, and CAD. This awareness of design tool artifacts translates to bias of associating a design tool with a set of attributes. This was illustrated in the written responses, including those attributes that were frequently mentioned, such as “blocky”, “precise”, or “organic”. That these biases have reasonable alignment with actual performance characteristics of different design tools is demonstrated by participants’ better-than-random likelihood to identify the original design tool.
CONTRIBUTIONS

Research Methods
This dissertation explored the use of novel research methods in the design space. There were often challenges associated with these methods, involving the development of custom tools and evaluation workflows. However, now that these methods have been tested, they could be adopted by other researchers in the design space.

Through the use of re-sketching all the original designs, we made an attempt to normalize the representation of ideas that would have otherwise been presented differently. I propose that this approach would benefit other contexts where ideas are being evaluated, whether in a research or industry environment. For this research we had hired an industrial designer to re-sketch all of the designs, however Fišer et al. are developing software to automate re-representation of 3D models based on the style of a particular sketch [12].

Following the generation of multiple design options, it is necessary to perform some kind of concept evaluation. Commonly, this could include rating the ideas based on a number of criteria. For design research, these concepts might be presented in a serial fashion to an audience who then must rate each one in sequence. Due to the unpredictable effects one idea may have on the rating of another, as well as the effects of primacy and recency, this approach may have unreliable results. The research in this dissertation took a different approach by constructing the evaluations as pair-comparisons. Although, this necessitates different methods of analysis, the solution is no less tractable and the potential benefits are appealing.

One of the challenges of doing research involving human subjects is recruiting enough subjects to get meaningful data. Compounding this issue is often the need for participants with a specific level of experience. By borrowing a method from the market research domain, this dissertation used adaptive choice-based conjoint to measure designer’s implicit preferences for design-tool/design-attribute associations. Adaptive methods like this should be given greater consideration when they have the potential to gather high-quality data about a large set of values from a more moderate sample set.

Research Questions
In this dissertation, I studied the results of designers using different tools in early-stage design process, as well as how designs are perceived by
potential users and designers even when those designs are translated to another medium. Returning to the first research question:

How does the use of a specific design tool influence the design outcome?

In the early stage design process, the choice of design tool — sketch, foam prototype, or CAD — influences the number of designs produced during a working period, the form of those designs, and the likelihood of if those designs are creative, comfortable, or aesthetically pleasing. We might think that the creation of sketches, physical prototypes, or CAD models, would be a way to support mental models — a manifestation of an idea already formed in the mind. However, this research demonstrates it is not just a one-way street from mental model to manifestation; the interaction is two-way and the manifestation influences the mental model. When we choose to use one tool over another, we implicitly make a set of decisions, including: how many designs we will create, how creative will those designs be.

This is not to say that we should pick the one, best tool and use that exclusively, but rather the tools, as they are most often used now, lead to these differences. By understanding how different tools can influence the design outcome, we can make more informed choices about the tools we use and how we use them. While in this study sketching produced fewer designs per designer than foam prototypes, it would be interesting to study the effects of coaching designers to sketch more quickly and more prolifically.

As use of these different design tools has demonstrated an influence on design attributes, designers should be able to take advantage of the strengths and weaknesses of the tools. Further, designers should develop a greater fluency in the array of tools available to them so that lack of experience with a tool does not limit the design space.

From prior research, we know that how an idea is represented can influence user’s perception. In this dissertation, I also investigated the perception of designers:

What design attributes do designers associate with these tools?

Experience with design tools has informed designers what design attributes to expect to see from the use of a specific tool. This research created a preliminary ontology describing design attributes and how designs can be represented. With this knowledge, we understand that designers may be prejudiced to think CAD is associated with low creativity; that CAD is precise, but sketching is loose; that sketching isn’t blocky, but foam is. Actual results from these tools support only some of these biases, but these biases are nevertheless relevant to how designers perceive a design.
It may not be practical, or effective, to neutralize some of these design-tool/design-attribute biases, so designers should work with an awareness and respect for what a design representation may communicate beyond the idea itself. This research studied what design-tool artifacts remained after the design was translated to a sketch. With this as a guide, further research could take a more active approach and seek to modify a design representation so as to mimic design attributes that would be expected from a different tool.
FUTURE WORK

Original and Re-Sketch Comparisons
So far this dissertation has focused on reviewing responses to the re-sketches of the original designs. This step of the process was based on prior research that studied the influence of representation fidelity and representation mode on product perception. There is an opportunity to contribute further to this research by also studying the original design artifacts produced from the design prompt. Specifically, these designs, in their original forms, could be rated for creativity, comfortability, etc. or directly compared against their re-sketch counterparts. The implications of this work could help improve our understanding of how the representation’s fidelity and mode influence user perception.

Design Context
The study in Part I of this dissertation examined the use of sketching, prototyping with blue foam, and CAD with SolidWorks in the development of designs for remote controls. Based on the designs produced using these tools, we developed some conclusions about how these tools can influence the design outcome. However, we know that these are not the only tools available or in use in a design process. There are different tools for sketching; physical prototyping can be accomplished via any variety of materials not just blue foam; SolidWorks is but one CAD package amongst many others.

The original design prompt was to design remote controls for a family entertainment system. From this prompt, we made observations of differences in the design outcome based on the tool used, yet it is not known to what degree the prompt has on the significance of the tool used. Perhaps some design prompts may provide a greater advantage or disadvantage to certain design tools. For example, the original prompt focused mainly on user-facing features and the form of the design; future work could expand to include fundamental functional engineering aspects [17] of the design from the engineer’s perspective. Continuing this research by expanding our study to include greater variation of tools and design prompts will build on our knowledge for the interrelationships of design tools and design context and their influence on design outcome.
REFERENCES


preferences in early stage design. *Journal of Mechanical Design*, 137(7), 071408.


