Prototyping for User Feedback: How Visual Detail Can Influence Feature Preference

by

James Loren Christian

Bachelor of Arts, Bachelor of Science in Mechanical Engineering
University of Michigan, 2012

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

Massachusetts Institute of Technology

June 2014

© 2014 Massachusetts Institute of Technology. All rights reserved.

Signature redacted

Author

Department of Mechanical Engineering
May 9, 2014

Signature redacted

Certified by

Associate Professor of Mechanical Engineering & Engineering Systems
Thesis Supervisor

Signature redacted

Accepted by

Professor of Mechanical Engineering
Chairman, Department Committee on Graduate Students
Prototyping for User Feedback: How Visual Detail Can Influence Feature Preference

by

James Loren Christian

Submitted to the Department of Mechanical Engineering on May 9, 2014, in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

Abstract

User-centered design focuses on insights drawn from users to drive product development. Provisional design prototypes are often used as a tool for collecting feedback from users that can then be incorporated into future design iterations. However, experience from design practice suggests that user feedback should not always be taken verbatim, and that the particular context in which a prototype is presented can greatly influence the user’s perception of the concept. The theory of preference construction helps design researchers understand how users’ preferences are dependent on the particular context of a design representation. Thus, it is important for designers to have insight into the potential effects of contextual factors. This thesis presents the results of a survey of ~200 users for their responses to four prototype designs for a handheld GPS unit with exactly the same screen size but differences in two key visual details: a bezel surrounding the screen and an overmolded edge around the unit. Results showed that screen size was a key consideration in the respondents’ perception of the product, but that the addition of a bezel seemed to reduce the likelihood that they would state a preference for a larger screen. These findings can be incorporated into guidelines to aid design practitioners in maximizing the value of the feedback they receive from users on provisional prototypes and sketches.

Thesis Supervisor: Maria C. Yang
Title: Associate Professor of Mechanical Engineering & Engineering Systems
Acknowledgements

I would like to thank my thesis advisor, Maria Yang, for her constant support and positive reinforcement over the past two years. She has been the most patient and selfless mentor. This exceptional research opportunity has been an exploration in a direction that truly fascinates me, and I feel lucky to have had the opportunity to take it on in such a rich environment.

Thank you to my friends and colleagues of the Ideation Lab. As a team, your presence is inspiring. Thank you to Jesse, Geoff, Bo, Jasmine, and Mike. A special thank you to Maxime, for your loyal support and enthusiasm for challenging my ideas, to Qifang for your willingness to share knowledge and help me think through my thoughts, and to Anders for your support and collaboration on multiple projects.

Thank you to my co-founders and incredible friends at AvaTech. Brint, your unstoppable energy and positive attitude are a source of daily inspiration. Working with you has redefined my definition of interdisciplinary design, and strengthened my confidence in the ability for design to make a difference. Your leadership has literally led me to places I would never have gone. And Sam, your constant forward progress and never-give-up attitude help me to keep going. As partners in design I think we bring out the best in each other, and we have the potential to make a real difference.

Thank you to Mom, for always reminding me what is truly important in life. Thank you to Dad, for helping me to reset and level myself when the going gets tough.

And finally, thank you to Amanda. Your loyalty, sincerity, and patience are incredible. You always know the right thing to say, and your positivity is the spark that keeps me going through the stressful times.
# Table of Contents

Abstract 2

Acknowledgements 3

Introduction 5

Literature Review 8

- Prototyping and the form-giving process 8
- The Designer-Prototype-User Relationship 9
- Preference Construction and Inconsistency in Design 10
- Research Gap 11

Methods 13

- Product Selection 13
  - Handheld GPS Receivers 13
  - What about Smartphones? 14
- Product Design and Visual Details 15
  - Bezel 15
  - Overmold 17
- Materials: Prototype Concept Variants 18
- Survey Procedure 21
- Mechanical Turk Limitations 22
  - Worker Requirements 23
  - Response Quality Checks 23
- Data Analysis Methods 24

Results 25

- Participants 25
- Implicit Initial Responses 25
- Explicit Feature Preferences 26

Conclusions 28

- Limitations 29

Future Work 30

References 31
Introduction

Product form, or industrial design, is highly important for the desirability of a product, and ultimately its success in the marketplace (Creusen and Schoormans, 2005; Bloch, 1995). It has been suggested that many undesirable products are the result of poor integration between product form and function (Wallace and Jakiela, 1993). As such, the field of user-centered design helps product designers consider how a product takes form at early stages in the design process (Abras, et al., 2004).

The early, front-end stages of the design process are marked by ambiguity as the design team formulates a design direction. It is particularly difficult at this stage to know if a direction is appropriate, as the problem and solution continue to evolve while the designers gather knowledge and familiarity with the space (Dorst and Cross, 2001). A traditional strategy is to show provisional sketches and/or prototypes to target users for feedback. More specifically, designers often use ranking and/or scoring methods to elicit quantitative preferences from users (Thompson, 2003). The quantifiable results (often coded from qualitative, or open-ended, user feedback) inform designers and allow for more objective decision-making.

It is well known among designers that user feedback should rarely be taken verbatim, because the context in which prototypes are presented can influence a user’s opinion (Ulrich and Eppinger, 1995). However, “context” is a broad term. The list of contextual parameters that may influence feedback is nearly endless. One could consider anything from the environment and social setting, to the product attributes and form, or even to the chosen prototype representation style and craftsmanship. In considering the designer-prototype-user relationship, some theorists even stress the importance of considering the user’s metaphysical awareness of themselves being observed by designers (Crilly, 1995).

One particular design experience stood out as inspiration for this study. The author co-founded a company called AvaTech, with the mission of helping backcountry recreationalists avoid dangerous avalanches. During the development of their first product - a portable device used to measure snow structure properties and other critical factors relevant to avalanche risk assessment - the designers created a set of ultra-durable and bulky prototypes for an engineering testing trip in South America. In an attempt to collect data on target user interest, these prototypes were painted and detailed to communicate as a consumer product. As a result of time and monetary limitations, the designers planned to collect feedback while explaining to users that the final product would be much smaller and more compact. However, the feedback received from target users was unexpected. Multiple users communicated a
strong desire for an even larger device with a bigger screen for showing measurement data. This came as a surprise because the designers knew that backcountry recreationalists were extremely volume- and weight-conscious. So, the designers chose to ignore the feedback – rather than making the device and display larger, they chose to hold the display constant and make the device significantly smaller. After another round of prototypes with the smaller form-factor, the designers stopped receiving requests for a bigger screen. While it seems that the correct path was chosen, the designers lacked confidence in how to interpret and act upon the user feedback. It became clear that the appearance of the device played a role in target user perception and preference elicitation. But what exactly was the cause? Moreover, how could the designers anticipate this and curate more reliable user feedback scenarios?

Preference elicitation toward design alternatives has been widely studied by the design research community. The concept of preference construction – that a person’s preference toward something does not exist prior to it being asked (Slovic, 1995) – is commonly used to understand variations in preference elicitation tasks. It is widely accepted that user preferences are dependent on how questions are asked and how design concepts are presented (Crilly, et al., 2004). Furthermore, research has been done to show that the construction of preferences can result in inconsistencies in user preference data (Reid, et al., 2013). This can complicate and often mislead the design process.

Obtaining consistency in preferences is a crucial step in product design. Without consistency, decision-makers find it difficult to support or feel confident in their choices (Kulok and Lewis, 2005). Thus, it is important for designers to collect reliable and consistent preference data from target users. Design researchers have studied this topic from multiple angles, with the intent of understanding how contextual factors can impact user preferences and preference consistency. In one particularly relevant study, researchers showed how user preferences toward product attributes are highly dependent on the level and clarity of information given (MacDonald, et al., 2009). The authors suggest that users may, correctly or incorrectly, assume a relationship between two separate product attributes that may influence their preferences. In other studies, it has been shown that factors such as representation style (Macomber and Yang, 2011), product aesthetic details (Bao, et al., 2014), and concept framing and background (Chen, et al., 2013) play a significant role.

In this thesis we explore a different relationship – the connection between visual details of product prototypes and user preferences toward core product attributes. We will use the phrase “visual details” to describe the sort of surface-level form
language of products such as corner and edge treatments, coloring, materials, texture, etc. As user-centered design has shown us the importance of form development in early stages of the design process, we seek to understand the subtleties in form development that could actually influence user feedback and alter the direction of the design process to follow. In the AvaTech device example, the designers asked themselves: how did the overly bulky and robust form language of the original prototypes influence user preference elicitation toward the core product attributes, such as screen size?

The following research questions guided this study:

- Do the visual details on a prototype impact user preference toward core product attributes?
- If so, how do the visual details on a prototype influence user feedback toward core product attributes?

In this study we focus on two commonly used visual details in product design – overmolds and bezels – and apply them to generate multiple prototype variants of a consumer product – a handheld GPS receiver. We hypothesize that the application of these details will have a notable impact on user preference and feedback toward the size of the product’s screen.
Literature Review

The use of prototyping as a means to elicit user feedback and understand preferences has been widely studied by the design community. This review looks at the current literature on prototyping methods, user interaction and perception of design prototypes, and broader studies on preference elicitation in the context of design decision-making. The intersection between these bodies of work is discussed, and a gap for new exploration is identified.

Prototyping and the form-giving process

Product form is a crucial element in the success of marketable consumer products (Creusen and Schoormans, 2005; Bloch, 1995). It has been shown that form and design can be used to differentiate products in competitive markets (Berkowitz, 1987) and help to communicate information about product attributes (Nussbaum, 1993). While historically considered a task for industrial designers, the development of product form has been incorporated into the early stages of product design (Abras, 2004).

“Virtually all innovative designs emerge in the interplay between two dueling representations: the wish list of specifications that describe and define the new ideas, and the prototypes that attempt to embody them.” (Schrage, 1993)

Prototypes are the medium through which designers communicate product form and function (Schrage, 1993). They are used to answer questions, both internally among designers and externally to elicit feedback from target users. As tools and methods for prototyping get more complex, there has been a movement to focus prototype design to answer critical questions (Houde and Hill, 1997). The concept of “looks-like” and “works-like” prototypes is commonly used to help distinguish the question a prototype seeks to answer.

As we attempt to understand how prototype form can influence user preferences, it is helpful to first understand how prototype forms originate. In other words, how does a designer give form to begin with? Wallace proposed a model for form giving (Wallace and Jakiela, 1993). The model suggests a four-step process to separate the work into a series of hierarchical problems. The discrete steps of this model include (1) product organization, (2) surface design, (3) surface detailing, and (4) graphics. The value of this model is that it allows us to visualize how product feature choices impact the form-giving process. For example, the selection of a particular component for a handheld electronic product will most heavily influence the first step – product
organization, whereas the selection of a certain typeface for product decals will most heavily influence the final step—graphics.

**The Designer-Prototype-User Relationship**

There is an interesting body of research focusing on the relationship between designers, prototypes, and users. The concept of the “design stance” was originally coined by Dennett as a mindset that users adopt when trying to learn a system where they focus on what the system was designed to do (Dennett, 1989). Crilly expanded on this concept and applied it directly to the communication between designers and target users in prototype testing and user feedback scenarios. The diagrams he creates serve as excellent summaries, and suggest a deep level of insight into the factors that influence the outcomes of these scenarios. In the particular diagram shown in Figure 1 below, we point out a few key factors in this complex relationship:

- The designer has an image of how the product will turn out, which is influenced by personal beliefs, motivations, capabilities etc.
- The designer has a plan for the prototype, which may be different than the realization of the prototype.
- The user has an interpretation of the prototype as it is realized, which is influenced by personal experiences, expectations, motivations, etc.

![Figure 1: Diagram from The Design Stance in User-System Interaction (image source: Crilly, et al., 2008)](image)

Many other studies have been conducted to understand how design representation impacts user judgment of (and feedback toward) prototypes. For a good summary of the relevant literature, refer to Reid, et al., 2013. The summary proposes a framework for three categories of judgments: Opinions, Objective Evaluations, and Inferences.
The authors explain how these three categories can be expressed through information completeness and the existence of a "correct" answer. When a judgment task has a correct answer, providing incomplete information will result in an inference while providing complete information will result in an objective evaluation. When a judgment task has no correct answer, the result is an opinion. The author specifies that opinions are most often a complex combination of inferences and objective evaluations. An even more complex question is "which design do you prefer?" This is commonly asked in design research experiments, where researchers use pairwise comparisons of prototypes to find a relationship between contextual factors and user preference.

Preference Construction and Inconsistency in Design

The concept of preference construction was developed through observations of common themes in behavioral decision research (Slovic, 1995). In his paper, Slovic suggested that preferences don't exist prior to being asked, but rather people construct preferences during preference elicitation tasks. This concept was widely accepted as it helped to explain why contextual factors can result in preference inconsistencies and reversals previously found in behavioral psychology, economics, and marketing.

Research in the field of behavioral psychology has shown that the framing of decision problems can affect choices. Tversky and Kahneman provide a thorough review of psychological principles that govern the perception of decision problems (Tversky and Kahneman, 1985). They use examples of choices involving monetary outcomes and potential loss of human life to illustrate the phenomenon of preference reversal.

Furthermore, the field of behavioral economics has used context effects to challenge the assumption that humans are rational decision makers. In his book Predictably Irrational, Ariely uses a series of small experiments to show how the framing of decisions, pricing, and product alternatives can be used to controllably lead buyers to irrational and/or inconsistent decisions (Ariely, 2008). In one particular example using newspaper subscriptions, he shows how the presence of an inferior and undesirable subscription option can effectively shift customer preferences between the more desirable alternatives. This example is particularly intriguing as an analogy to product design and target user feedback scenarios. It begs the question "how does the presence of bad design concepts influence user preference between the better concepts?"
In recent years, the concept of preference construction has been applied to engineering and design decision-making methods. For a thorough review of preference elicitation research, see MacDonald, et al., 2009. The authors propose a framework for understanding the different types of preference inconsistencies that can occur as a result of constructed preferences. The framework suggests three components – comparative, internal, and external inconsistencies – and the study provides a unique example of how each component can impact the design process. The authors develop the concept of crux and sentinel attributes, suggesting that users may, correctly or incorrectly, assume a relationship between the two. In the most relevant example, a comparative inconsistency is illustrated with a study of user preference toward paper towel products. They show that a paper towel's quilting (the sentinel attribute) is only seen as important to users if the users are not given information about its absorbency (the crux attribute). In other words, users infer a relationship between two product attributes and elicit preference based on that inference. This can falsely bolster an attribute when that attribute is assumed to impact another. The take-away is that care must be taken if users think that relationships exist between product attributes. The three examples from this study support the notion that users have limited abilities to assess their own needs and preferences, and that it is critical for designers to understand those limitations when collecting this information.

Further design research has shown inconsistencies in other product preference elicitation tasks. For example, Bao uses the case of residential solar panels to show how preferences toward the likeability of these products are dependent on the context in which they are shown (Bao, et al., 2014). In their experiment, they found that preferred colors of solar panels was dependent on the color of the roof of the house in which it was shown. This is a strong example of how the visual details of the context can actually influence the preferences toward the visual details of a product.

**Research Gap**

Product designers and entrepreneurs have widely adopted the notion of user-centered design, and with it the iterative, co-evolutionary process of problem/solution development. This means that new products take form as the understanding of design problems takes place. As such, many opportunities for user feedback come mid-process and quite often with imprecise prototypes and partial design representations.

Behavioral psychology, economics, and product design decision-making research have all shown us that preferences toward products and concepts are constructed
based on the context of the decision-making scenario. Thus, we should be aware of how differences in designer and user knowledge and experiences can impact perceptions of prototypes and hence the outcomes of user feedback scenarios.

Much of the work in this area has been focused on design representation styles or on variations in product visual details to understand what users “like” more. Digging a little deeper and beyond the surface-level likeability of a product, we find there are many unanswered questions about how visual details in prototypes can impact user preferences toward core product attributes. We feel these questions are particularly relevant when we consider how product designers are trained to iterate on concepts based on user feedback.
Methods

To test our research questions we conducted an online survey to collect user feedback on a product prototype. By randomly showing each participant a visual detail variant of the prototype, we can begin to understand how this can influence both implicit and explicit preference elicitation toward product features. This section explains how the product example and visual details were chosen, and how the survey was conducted.

Product Selection

The first step was to select a type of product with which we could effectively test our hypothesis. While experiences with the AvaTech device were strong inspiration for the research questions and hypothesis, there were logistical challenges that made it not feasible as an example product. Instead, we focused on a handheld GPS receiver designed for outdoor activities. We chose this product for the following reasons:

- **A handheld GPS receiver is recognizable.** Most people are familiar with the GPS and can understand what it is and what it does with little explanation.
- **The ratio between product size and visual details is close to 1:1.** We can capture the full product and the user interface details (buttons, screen, graphics) of a handheld GPS receiver in one image.
- **The market is large.** There are a lot of people who use handheld GPS receivers because they are convenient for a variety of outdoors activities. This makes it feasible to collect large samples of data.

Handheld GPS Receivers

The satellite-based Global Positioning System (GPS) allows a GPS receiver to determine its position using four or more satellites. They are often packaged into small, portable electronics devices, and are convenient for outdoor activities that involve navigation. We researched the current handheld GPS receivers on the market to understand the typical product architecture, features, and details. Two brands – Garmin and Magellan – dominate the market. Each company sells multiple models of these products, including both touch-enabled and button-enabled models. We chose to focus on button-enabled models, to avoid too strong a similarity with popular smartphone design. Figure 2 below shows a few of the most popular button-enabled models.
Garmin eTrex 10

Dimensions
4.0" x 2.1" x 1.3"

Screen Size
2.2" diagonal

Screen % of Face
24.3%

Garmin GPSMAP 62

Dimensions
6.3" x 2.4" x 1.4"

Screen Size
2.6" diagonal

Screen % of Face
28.0%

Magellan Triton 2000

Dimensions
5.8" x 2.5" x 1.1"

Screen Size
2.5" diagonal

Screen % of Face
27.6%

Figure 2: Handheld GPS receivers in the current market span a wide range of sizes. The viewable screen area is a small fraction of the total device face. (image sources: www.garmin.com, www.magellangps.com)

What about Smartphones?
The current market for handheld GPS receivers experiences significant competition from smartphone devices. For many users, a GPS-enabled smartphone is sufficient and there is no need for a separate, dedicated handheld GPS receiver. However, these dedicated units are often enabled with specialty features for advanced users that make them superior to their smartphone competitors.
Product Design and Visual Details

As you can see in the product examples shown above, button-enable GPS receivers have a relatively busy aesthetic. They use solid colors and bold detailing to communicate durability and utility. Notice that the ratio of screen size to device size for all of these examples is quite small. In other words, there is a lot of bulk around a small viewing area. According to an infographic titled "How much of your phone is actually screen" (author unknown, image source: http://c2.bgr.com/2014/02/slim-bezel-smartphone-infographic-full.jpg) the screen ratio for smartphones ranges from 50-76%. In a world of smartphones, touch-enabled products, and infographics like this one, how do these GPS units retain their desirability? A product designer would remind us that the use case is different. These GPS units are designed to withstand harsh weather and abuse during outdoor activities, while a smartphone is not. This means that these units need physical buttons that operate when wet or with gloved hands, and a strong case that won't break easily when dropped. Also the screen on a handheld GPS is designed for navigation while users are enjoying the outdoors, while the screen on a smartphone is designed as a canvas for a massive range of applications and functionality. When compared to smartphones, handheld GPS receivers favor a smaller screen and a more rugged form-factor.

So, if we accept that the screen ratio on a handheld GPS receiver is going to be relatively small, what details can product designers use to make the device look desirable? We chose to focus on two details that are commonly used and appear to have a very strong visual effect: the bezel and the overmold.

Bezel

A bezel is the margin between the screen area and the outer contour of a product. This is a commonly used detail in consumer electronic products, as it can achieve multiple design objectives. The following is a non-exhaustive list of four different applications of a bezel with examples.

1. **For functionality and manufacturability.** First and foremost, a bezel is often a useful feature for the product functionality and manufacturability. A bezel can be useful for protecting the screen and providing a way to secure it in place.

2. **To frame the screen with a color detail.** The bezel can also serve to create a visual separation between the image on the screen and the rest of the environment.
Figure 3: The Pebble Smartwatch boasts a large bezel around the screen area. It comes in a variety of colors. (image source: www.getpebble.com)

3. **To change the visual form language of the screen itself.** Screens are inherently rectangular in their form. To avoid wasting any screen pixels, a bezel will be rectangular on the inner contour. However, the outer contour can have any form. The Garmin eTrex handheld GPS is a good example, where the bezel around the (relatively small) screen has an outer contour curved to match the outer contour of the product as a whole.

Figure 4: The bezel on the Garmin eTrex 10 gives the display area a form language common with the contour of the device.
4. **To give the illusion that the screen is bigger.** In certain conditions, a color-matched bezel can make it visually challenging to understand where the screen ends and the plastic begins. Take the Macbook Pro Retina, for example. It features a 0.5” black bezel on all sides of the screen. When the screen is turned off, or when the background image is dark, it becomes next to impossible to see the boundary.

Figure 5: The black bezel on the Macbook Pro Retina blends directly into the screen when powered off or when there is a dark background image. (image source: www.apple.com)

**Overmold**
An overmold is a secondary material molded over the surfaces of a product. This process is used to incorporate multiple materials into a single part. Often, overmolds are made of a softer material and can be used to improve the durability, grip comfort, and visual appearance of a product. Examples of these effects are shown below.

1. **For product durability.** An overmold made of rubber can improve shock absorption for high impact forces (such as dropping).

Figure 6: The Dell Latitude ST Tablet has a TPU overmold for durability. (image source: www.dell.com)
2. **For holding comfort and grip.** Overmolds can be used to introduce a new material with better grip properties.

![Image of Oral B Advantage toothbrush](image source: www.oralb.com)

**Figure 7:** The Oral B Advantage toothbrush uses overmolds to provide good grip while wet.

3. **For visual appearance.** The exact design of overmolds takes into consideration mechanical requirements, aesthetics, and user interaction. A good overmold can contribute to the visual interest of the product and also serve as clear indication how to hold or interact with the product.

![Image of DeWalt Bare Tool DC385B](image source: www.dewalt.com)

**Figure 8:** The DeWalt Bare Tool DC385B has overmolds that clearly indicate where the users' hands are supposed to go – one on the grip in the back (right side) and one underneath the barrel in the front.

**Materials: Prototype Concept Variants**

The prototype concept for this study was designed to be easily recognizable as a handheld GPS receiver. As shown above, the bezel and the overmold are both commonly seen in products on the market today. By alternating the presence of the bezel and the overmold, we created four visual detail variants of the same handheld...
GPS concept. We hypothesize that these detail variants will have an impact on the implicit and explicit preference elicitation toward product features such as screen size. The variants are shown in Figure 9 below.

![Figure 9: Alternating the presence of the bezel and the overmold creates four prototype visual detail variants.](image)

Prototype renderings were created using a combination of Dassault Systems SolidWorks 2013-2014 Educational Version, Adobe Photoshop CS6, and Adobe Illustrator CS6. A 3-dimensional CAD model of the prototype was created in
Solidworks, using the feature configuration tools to toggle on/off the bezel and overmold features. Simple lighting effects and surface colors were applied to create renderings in Solidworks. The display image – a black & white topographical map – was created using Illustrator. The prototype rendering and display image were combined in Photoshop. The drawing of the hand and the image of a ruler were both added to help give participants a sense of scale. It was shown to participants in grayscale to help focus on form and design rather than color preferences. An example of a final prototype variant shown to participants is shown in Figure 10.

Figure 10: Prototype D – Overmold + Bezel. Each prototype visual detail variant was presented to users held in a hand and with a 6-inch ruler to help give a sense of scale.
At the end of the introduction we hypothesized generally that the application of visual details will have a notable impact on user preference and feedback toward the size of the product's screen. After exploring the overmold and bezel features in detail, we specify exactly how we predict these features to influence preference elicitation:

- Participant elicitation of preference toward screen size will be influenced by the visual detail variant shown.
- The presence of a visual detail (either bezel or overmold) will decrease the chances that a participant will suggest an increase in screen size. We predict that the bezel will have a stronger effect, thus the order from most- to least-likely to recommend an increase in screen size is expected to be: A, C, B, D.
- The likelihood to suggest a screen size change will increase when prompted specifically about the attribute, (i.e. participants are more likely to agree to a specific change than they are to suggest the change without prompting)

### Survey Procedure

The survey was targeted at gathering feedback from potential users about the overall design and product features of the handheld GPS receiver. Each participant was randomly shown one of the four prototype visual detail variants, and asked to provide both open-ended and structured feedback. The open-ended feedback provided the opportunity to collect data about implicit user preferences, while the structured feedback with directed questions allowed for the collection of explicit preference data. This simple design allowed us to search for correlations between prototype variant shown and feedback provided. (i.e. the independent variable was prototype variant and the dependent variable was the participant response) A full document of the survey can be found in the appendix.

We created the survey using Qualtrics, and collected responses using Amazon's Mechanical Turk, a web-based crowdsourcing marketplace where requesters post and workers complete Human Intelligence Tasks (HITs). To verify completion, we directed participants to copy a 7-digit code generated in Qualtrics at the end of the survey back to Mechanical Turk. Participants were paid $0.75 for the HIT. According to timing methods provided by Qualtrics, completion time for the survey ranged greatly, from as little as 80 seconds to as many as 66 minutes. The median completion time (chosen to ignore the effect of extreme outliers) was approximately 5 minutes. This comes out to a median wage of $9/hour.

We conducted multiple pilot surveys to verify that the survey was clear and understandable by a range of respondents. Feedback from these pilots helped us to
design the product visual representations and create the language to encourage valuable participant responses.

The survey began with demographic info (age, gender) and GPS use information. To boost participant interest in the example product used for this study, we included a short blurb about the product features. In addition to boosting interest, it also helped to combat the problem with participants fixating on elements irrelevant to the task (e.g. focusing on how they would prefer to use a smartphone, rather than giving feedback to the design itself).

Our company is developing a new handheld GPS receiver for outdoor enthusiasts. While some smartphones can be used for navigation, our customers have told us that they want a product designed specifically to meet their advanced needs. Here are some key advantages/features that our product will have:

- New technology for better location accuracy and tracking capabilities
- High-resolution sunlight-readable display
- Simple, easy to use button interface
- Ultra-low power design for superior battery life
- Durable, waterproof, and shockproof for harsh environments

In the core of the survey, participants were randomly shown one of the prototype variants and asked to provide their open-ended feedback to the design. The approach of starting with open-ended questions was intended to gather initial reactions to the design before prompting users to think specifically about the chosen product features. After this response was submitted, they were asked to provide more focused feedback on both the overall size of the device and the size of the display. We phrased the prompts from the perspective of “how would you improve the design”, and provided a 3-point rating scale for the width and height dimensions of these two attributes. We followed up by asking participants to explain the ratings provided.

**Mechanical Turk Limitations**

While Mechanical Turk is a powerful tool for collecting data from large populations quickly and with relative ease, it does have its flaws. Gauging participant honesty and effort are the biggest challenges. In the interest of earning money with minimal effort and/or optimizing their wages, participants may rush through surveys without putting thought into them. Responses from overly rushed participants could invalidate the experimental data. Thus, we designed the survey and set worker requirements to encourage quality responses. We also imposed a few quality checks to flag dishonest and/or low-effort responses.
Worker Requirements
Mechanical Turk allows requesters to set eligibility requirements for workers who might be interested in completing a HIT. Research in best practices has shown that setting a high requirement on the HIT approval rating for workers can improve quality of data received (Mason and Suri, 2012). While this cannot guarantee that only quality data will be submitted, it can help to improve the odds significantly. In order to support the best quality survey responses, we set the following Mechanical Turk worker requirements:

- HIT Approval Rate for all Requesters’ HITs greater than or equal to 97%
- Number of HITs Approved greater than or equal to 100
- Location is United States

Response Quality Checks
We took two precautions in assessing the quality of participant responses to the survey. Firstly, we used Qualtrics’ built-in timing question to observe the amount of time a participant spent on each page. The purpose of this question was to identify any participants who spent exceptionally short amounts of time on any given page. The purpose of this question was to identify any participants who spent exceptionally short amounts of time on any given page.

Secondly, we added a test question at the end of the survey asking the user to recall one of the five new product features listed at the beginning of the survey. This was formatted as a multiple-choice question with only one correct answer. The other two possibilities were written to be theoretically feasible product features but clear indicators that the participant did not remember the product description. To persuade participants to not refresh the survey and hunt for a correct answer, we made it clear that their answer to the question would not impact their ability to complete the survey or receive compensation.

At the beginning of the survey we listed five product features. Which one of the following was in the original list? (your response here alone won’t impact your ability to complete the survey or receive compensation)

- Preloaded 100K topo mapping
- Ultra-low power design for superior battery life
- 5 megapixel auto-focus digital camera

These two approaches provide insight to the participant’s survey procedure, but they are not direct measurements of response effort or honesty. While relatively short page times suggest a rushed participant, we cannot conclude that long page times indicate a thoughtful or honest participant. While an incorrect response indicates that the participant either forgot or didn’t read the product features, a correct response could be the result of a lucky guess or a participant who read the features carefully but didn’t put effort into their response. It is important to remember that
workers operate with different speed and performance characteristics that are out of our control. In this paper we argue that responses from participants who completed the survey in less than two minutes and/or answered the test question incorrectly should be omitted.

**Data Analysis Methods**

From this survey we will focus primarily on the outcomes of two pieces of data. First, we look at the initial response feedback from participants. Second, we look at the 3-point rating scale data focused on feature preference toward overall device size and screen size. We compare both sets of the data against the visual detail variant of the prototype shown to the participants.

Initial open-ended responses from participants were coded according to whether or not the participant suggested a concern with the screen size on the device. The responses were coded in a binary fashion – as either showing evidence of this concern or not. Fisher’s Exact Test was used to conduct a pairwise comparison between visual detail variants, to determine if the number of participants expressing screen size concern toward one variant was significantly different than another variant.

We asked users to rate all product attributes on a 3-point scale with the options: (1) Make it larger, (2) Don’t change it – I like it as is, (3) Make it smaller. This is an ordinal scale because there is a clear ordering of the choices, but there is no measurable distance between each rank. In other words, a participant rating the screen height as “Make it smaller” provides no distinction of how much smaller. As such, we used the Kruskal-Wallis Test, designed specifically to analyze the significance of ordinal data sets. This test allows us to determine if the ratings provided for each visual detail variant were significantly different.
Results

Participants

In total we received 212 responses; 116 male, 96 female, over age ranges from <20 to 61+. We found that 5 users completed the test too quickly (fewer than 2 minutes), 28 users answered the test question incorrectly, and 4 users failed both quality checks. After elimination of the responses that did not pass these checks, we had a total of 175 participant responses.

Overall, our survey population showed a high engagement in outdoor activities. Over 90% of participants claimed to take part in at least one outdoor activity that could include the use of a handheld GPS. We found it interesting that 33% of participants claim they do not currently use a GPS unit of any kind (on their smartphone, or otherwise). However, when told about the product features, fewer than 10% of participants claimed they would not be interested in the proposed GPS based on the initial product blurb. After observing a random product detail variant, fewer than 10% of participants claimed to have a neutral or unfavorable preference toward the proposed GPS. From this data we see no significant difference in the overall level of interest of the product throughout the survey.

Implicit Initial Responses

In total, 33 initial responses were characterized as an elicitation of preference toward a larger screen. Table 1 below shows this data broken down according to the visual detail variant of the device shown to the participant.

<table>
<thead>
<tr>
<th>Response</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggest screen size increase</td>
<td>13</td>
<td>4</td>
<td>9</td>
<td>7</td>
<td>33</td>
</tr>
<tr>
<td>No mention of screen size</td>
<td>33</td>
<td>41</td>
<td>33</td>
<td>35</td>
<td>142</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>45</td>
<td>42</td>
<td>42</td>
<td>175</td>
</tr>
</tbody>
</table>

Table 1: Responses with evidence of preference elicitation toward a larger screen, broken down by visual detail variant shown.

We conducted a pairwise comparison using a one-tailed Chi-Square Test to identify statistical significance between test conditions. In summary, we found that there was a significant difference (p=0.0327) between visual detail variant A and B. This can be interpreted to say that, under the given test conditions, participants were
significantly more likely to raise concern about the screen size when shown variant A than they are when shown variant B. Other conditions were not significantly different, though a trend was seen in the data. The breakdown of participant responses is shown in Figure 11 below.

![Initial Response: Raised Concerns About Screen Size](image)

Figure 11: Participants were more likely to express screen size concerns when shown visual detail variant A than when shown B. Variants C and D fall somewhere in between, without statistical significance.

**Explicit Feature Preferences**

Explicit preference data toward screen size and overall device size was analyzed for trends. Ranked preferences toward overall device size suggest that participants were generally happy with the design. In total, 109 participants (62%) suggested no changes in either the width or the height of the device. Using the Kruskal-Wallis Test, no significance was found between the ratings for each visual detail variant of the product. This data suggests that the visual detail variations used in this study have minimal impact on preferences toward overall device size.

However, the explicit preference data toward screen size was more interesting. As compared to overall device size, screen size preferences tended toward much more
dissatisfaction. In total, 88 participants (50%) suggested changes to the width of the screen and 100 participants (57%) suggested changes to the height of the screen. Combined together, 111 participants (63%) suggested at least one screen dimension should change.

A pairwise comparison with the Kruskal-Wallis Test showed that preferences toward the screen height of variants A and B were significantly different (p=0.0207). Further analysis by comparing one concept against the rest showed that variants A and B lie on opposite ends of the spectrum while C and D fall somewhere in between. This data suggests that participants who view the prototype without any visual details (variant A) are significantly more likely to suggest an increase in screen height than those who view the prototype with a dark bezel (variant B).

Preferences toward screen width follow a similar trend, however with less significance. It is possible that the only reason for similarity is that participants generally gave the same rating for screen height and screen width. For example, over 90% of participants who rated the screen height as “make it larger” also rated the screen width in the same way. A graphical breakdown of preferences toward all four product attributes according to visual detail variant can be seen in Figure 12. Statistical significance data (p-values from Kruskal-Wallis Test) are listed in Table 2.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Screen Height</th>
<th>Screen Width</th>
<th>Device Height</th>
<th>Device Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>A vs. B vs. C vs. D</td>
<td>0.09</td>
<td>0.54</td>
<td>0.60</td>
<td>0.87</td>
</tr>
<tr>
<td>A vs. B</td>
<td><strong>0.02</strong></td>
<td>0.15</td>
<td>0.24</td>
<td>0.90</td>
</tr>
<tr>
<td>A vs. C</td>
<td>0.69</td>
<td>0.36</td>
<td>0.31</td>
<td>0.45</td>
</tr>
<tr>
<td>A vs. D</td>
<td>0.13</td>
<td>0.36</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>B vs. C</td>
<td>0.07</td>
<td>0.63</td>
<td>0.88</td>
<td>0.46</td>
</tr>
<tr>
<td>B vs. D</td>
<td>0.43</td>
<td>0.63</td>
<td>0.37</td>
<td>0.96</td>
</tr>
<tr>
<td>C vs. D</td>
<td>0.29</td>
<td>1.00</td>
<td>0.45</td>
<td>0.55</td>
</tr>
<tr>
<td>A vs. (BCD)</td>
<td>0.08</td>
<td>0.17</td>
<td>0.38</td>
<td>0.83</td>
</tr>
<tr>
<td>B vs. (ACD)</td>
<td><strong>0.04</strong></td>
<td>0.32</td>
<td>0.36</td>
<td>0.71</td>
</tr>
<tr>
<td>C vs. (ABD)</td>
<td>0.31</td>
<td>0.85</td>
<td>0.52</td>
<td>0.40</td>
</tr>
<tr>
<td>D vs. (ABC)</td>
<td>0.48</td>
<td>0.85</td>
<td>0.49</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 2: Kruskal-Wallis Test p-values for different visual detail variant comparisons.
Figure 12: Preference rating toward device and screen size attributes according to visual detail variant. Preference categories with less than 10% are shown but not labeled.
Conclusions

In this study we sought out to understand how the visual details of prototypes can influence target user preferences toward core product attributes. We used the example of a consumer product—a prototype handheld GPS receiver—with overmold and bezel variants to observe implicit and explicit preference elicitations toward overall device size and screen size. We hypothesized that both visual details would impact user feedback. Our hypothesis was only correct for certain comparisons. A summary of the results is presented and future work is discussed.

In summary, it was found that certain visual detail variations do have a significant impact on preference elicitation toward certain core product attributes. Specifically, it was found that the presence of a bezel can strongly influence user preference elicitation toward the size of the product screen. On the other hand, no significant correlations were found between the presence of an overmold and preference elicitations toward overall device size or screen size.

One conclusion from this study is that, under these conditions, the bezel has a stronger effect on screen size preference than the overmold. This is most-likely because the bezel interacts directly with the screen, and the color similarity can reduce the contrast, and therefore take the users' focus away from the screen edges. As expressed in the methods section, this effect is not uncommon in current consumer electronics products today.

Limitations

One particular limitation of this study was the choice of a 3-point rating scale for explicit product attribute preference elicitation. While this scale really boils it down to the key takeaways that designers may be seeking, the limited options could inhibit user expression of their preferences. For example, how does a user rate the screen width when they have a strong preference toward making it slightly larger? Does that participant rate their preference as “make it larger” or “I like it as is”? Or, does the coarseness of the rating scale unintentionally suggest that small changes are not even a possibility?

Another potential limitation is that, in the way our study was designed, there is no common reference point for participants to compare their preferences against. This makes it difficult to characterize individual participant tendencies, and therefore impossible to decouple the effects of pre-existing versus constructed preferences.
Finally, this thesis only studies one example product, two visual detail types, and four visual detail variants. This small experiment grants us a lot of insight into the potential effects of visual detail variants, but it is hard to make generalizable claims from such a study. Additionally, the sample sizes for this study were relatively small. It is important to note that we are inherently investigating small changes in preference. A change from 5% to 10% of participants eliciting their implicit preferences toward a core product attribute may be an interesting finding. However, when we expect such small differences in populations, larger sample sizes are needed to establish significant findings.

**Future Work**

Future work building on this preliminary experiment can both improve the strength of the findings as well as branch out to test other relevant contextual factors that may be at play. Increasing the sample size, adjusting the rating scale to a finer resolution, and/or asking more clarifying questions to survey participants could benefit this research. More drastic changes, such as bringing the prototype variants off the screen and into the physical world with off-line and in-person survey participation, could improve data quality and strengthen the claims.

To find more generalizable answers to these research questions, future work would need to look at a broader range of products and visual detail effects. Some possible products include: computers, smartphones, or vehicles. Other visual details include: corner and edge treatments, material, color, texture, or even prototype craftsmanship.

Finally, the findings and conclusions from this study are not intended to be kept at the academic level. Work should be done to explore how we can help product designers take this into consideration in their design process. A set of useful guidelines and themes could be identified and connected to other bodies of work such as Gestalt theory and/or Shape Grammars. The formalization of visual detail influences on target user feedback is an exciting prospect – one that we believe deserves attention in the product design research community.
References


Conferences and Computers and Information in Engineering Conference (pp. 291-300). American Society of Mechanical Engineers.


Appendix

The survey created in Qualtrics is shown in the images below. Each participant was randomly shown one of the four possible visual detail variants, however we only show one variant as an example here.

Product Concept Evaluation

Our research group is conducting a survey to understand how users provide design feedback. Your thoughts and opinions will be very valuable in helping us understand the design process, and we appreciate your time and effort. Your participation is entirely voluntary. However, you will only receive compensation for completing the survey. In any publications that may result from this research, any identifying information will remain confidential, and the data will only be presented in the aggregate.

You will receive a confirmation code for Mechanical Turk at the end of the survey.

Age Group

Gender

Do you participate in any of the following outdoor activities? Check all that apply.

- [ ] Hiking
- [ ] Geocaching
- [ ] Mountaineering
- [ ] Backpacking
- [ ] Backcountry skiing/snowboarding/snowshoeing
- [ ] Riding sports machines (ATV, Dirt Bike, Snowmobile, etc.)
- [ ] Other (please specify)

For any of the activities above, do you use a GPS? Please check all that apply.

- [ ] I use a handheld GPS unit
- [ ] I use the GPS on my smartphone
- [ ] I use a GPS unit other than a smartphone or handheld
- [ ] I don't use a GPS
Our company is developing a new handheld GPS receiver for outdoor enthusiasts. While some smartphones can be used for navigation, our customers have told us that they want a product designed specifically to meet their advanced needs. Here are some key advantages/features that our product will have:

- New technology for better location accuracy and tracking capabilities
- High-resolution sunlight-readable display
- Simple, easy to use button interface
- Ultra-low power design for superior battery life
- Durable, waterproof, and shockproof for harsh environments

Would you be interested in owning a handheld GPS unit with these features? (Please respond honestly. It will not affect your ability to complete the survey.)

Yes  Maybe  No

Your feedback is still valuable to us! For the remainder of the survey please do your best to answer the questions as if you were interested in owning this product.

Form/Design Feedback (part 1)

We are looking for your feedback on the form/design of the product. Below is a sketch of the concept, set in grayscale so you can focus on product structure and size.

Product Features:
- New technology for better location accuracy and tracking capabilities
- High-resolution sunlight-readable display
- Simple, easy to use button interface
- Ultra-low power design for superior battery life
- Durable, waterproof, and shockproof for harsh environments

Overall Product Dimensions
4.8" x 2.4" x 0.6" (12.2cm x 6.1cm x 1.5cm)
What are your thoughts about the form/design of this product concept? Please write as much as you need.
Form/Design Feedback (part 2)

Now, we'd like to get your feedback on some specific product features. Here is the same sketch from the last page.

**Overall Product Dimensions**
4.8” x 2.4” x 0.6” (12.2cm x 6.1cm x 1.5cm)

---

**Overall device size.**

To improve the design, how would you modify the overall device size?

<table>
<thead>
<tr>
<th></th>
<th>Make it smaller</th>
<th>Don't change it - I like it as is</th>
<th>Make it larger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device Height: 4.8” (12.2cm)</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Device Width: 2.4” (6.1cm)</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
Device display/screen.

To improve the design, how would you modify the size of the display/screen?

<table>
<thead>
<tr>
<th></th>
<th>Make it smaller</th>
<th>Don't change it - I like it as is</th>
<th>Make it larger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screen Height: 2.4&quot; (6.1cm)</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Screen Width: 1.6&quot; (4.1cm)</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

In your own words, please explain your selections above. If you would modify anything about the device display/screen, what would you do differently?

[Blank space for text]

Do you have any other design recommendations for this product?

[Blank space for text]

If your design recommendations were taken into account, how much would you like this product?

<table>
<thead>
<tr>
<th></th>
<th>Dislike very much</th>
<th>Dislike</th>
<th>Neither like nor dislike</th>
<th>Like</th>
<th>Like very much</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
Reviewer Feedback

At the beginning of the survey we listed five product features. Which one of the following was in the original list? (Your response here alone won’t impact your ability to complete the survey or receive compensation)

- Preloaded 100K topo mapping
- 5 megapixel auto-focus digital camera
- Ultra-low power design for superior battery life

Reviewer Feedback

Your results are important to us. Please take one minute to give us some feedback on this survey experience. Thanks!

How clear was your task?

- Very Clear
- Somewhat Clear
- Neutral
- Somewhat Unclear
- Very Unclear

Do you have any comments or feedback about the survey?