CrowdCrit: Crowdsourcing and Aggregating Visual Design Critique

Kurt Luther  
Carnegie Mellon University  
kluther@cs.cmu.edu

Maneesh Agrawala  
University of California, Berkeley  
maneesh@eecs.berkeley.edu

Amy Pavel  
University of California, Berkeley  
amypavel@gmail.com

Björn Hartmann  
University of California, Berkeley  
bjoern@eecs.berkeley.edu

Wei Wu  
University of California, Berkeley  
wuhuwei@gmail.com

Steven Dow  
Carnegie Mellon University  
spdow@cs.cmu.edu

Jari-lee Tolentino  
University of California, Irvine  
jltolent@gmail.com

Abstract

People who create visual designs often struggle to find high-quality critique outside a firm or classroom, and current online feedback solutions are limited. We created a system called CrowdCrit which leverages paid crowdsourcing to generate and visualize high-quality visual design critique. Our work extends prior crowd feedback research by focusing on scaffolding the process and language of studio critique for crowds.

Author Keywords
Design; critique; crowdsourcing; scaffolding; feedback

ACM Classification Keywords
H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organizational Interfaces

Introduction

For centuries, the "studio critique" model has provided a foundational exercise for art and design, and more recently, for project-based education in computing and engineering. Critiques can help novices to understand key principles in a domain, to articulate the goals and assumptions behind their work, and to recognize how others perceive their work [3]. Effective design critiques typically involve face-to-face interactions between people in a small group setting.
Unfortunately, high quality critique can be difficult to obtain outside of a design firm or classroom, especially for novice designers, who may need help with everyday tasks like posters, flyers, or slide decks. Designers may seek critiques in online communities, but even those who spend the time building a reputation are often unsatisfied with the low quality and quantity of critiques they receive [8]. Novice designers may experience evaluation apprehension and avoid sharing their works-in-progress alongside experts [4]. Paid crowds are attractive because of their speed, low cost, and scalability, but most crowd workers provide poor quality feedback because they lack design knowledge. When crowd workers have been integrated into a design process, their role has usually been to provide high-level impressions or contribute outside perspectives [2]. One recent study [9] showed that structured interfaces can help crowds provide more targeted feedback, but the study did not emphasize design principles or the process and language of critique.

The CrowdCrit System

We created the CrowdCrit system (Figure 1) to facilitate high-quality crowd critique using a learning theory called scaffolding [7]. CrowdCrit scaffolds the critique process and language for crowd workers by allowing them to select from a series of 70 pre-authored critique statements, based on widely followed visual design principles, and visually annotate relevant areas of the design. This system helps even novice crowd workers produce detailed, actionable feedback. Designers receive the critiques within hours and explore them using a novel aggregation interface. CrowdCrit is implemented as a web-based tool with Python, JavaScript, and the Mechanical Turk API.

<table>
<thead>
<tr>
<th>Layout</th>
<th>Good alignment</th>
<th>Poor alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readability</td>
<td>No spelling errors</td>
<td>Poor kerning</td>
</tr>
<tr>
<td>Simplicity</td>
<td>Simple and clean</td>
<td>Overuse of images</td>
</tr>
<tr>
<td>Emphasis</td>
<td>Strong focal point</td>
<td>Lacks hierarchy</td>
</tr>
<tr>
<td>Balance</td>
<td>Good use of symmetry</td>
<td>Lacks movement</td>
</tr>
<tr>
<td>Consistency</td>
<td>Good repetition</td>
<td>Lacks unity</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>Reaches intended audience</td>
<td>Mixed messages</td>
</tr>
</tbody>
</table>

Table 1. Design principles and sample critique statements (positive and negative).

![Figure 1. The CrowdCrit system allows designers to submit preliminary designs to be critiqued by online crowds and clients. The system then aggregates and visualizes the feedback for designers.

Critique Statements

We compiled a holistic set of 70 critique statements based on principles of effective visual design compiled from design textbooks (e.g. [1]). After several design iterations, this resulted in a group of seven high-level principles (see Table 1). Each statement has a short title, a more detailed description of the issue, and a possible, generic solution. This format is meant to embody Sadler’s [5] criteria for good feedback (specific, conceptual, and actionable) and provide a basic level of utility. Because traditional critiques involve both positive and negative statements, we included both strengths and weaknesses.

Eliciting Critique

The CrowdCrit critique interface is shown in Figure 2. Critique statements are organized into eight tabs: one for each of the seven design principles, plus an “Other” tab for critiques that our list may overlook. Within each tab, a definition of the design principle appears at the
The right side of the Overview tab displays the total number of critiques for the design (75 in Figure 3, left), along with a stacked bar chart showing one bar per principle. Bars are ordered by decreasing critique count so that the principle with the most critiques appears first. Hovering over any bar segment displays a particular critique statement identified by workers. Color encodes critique valence; positive critiques are green and negative critiques are purple. We use the labels “strengths” and “suggestions” to emphasize improvement through iteration, rather than admonishment for mistakes. Crowd expertise is encoded in lightness values; darker shades indicate critiques from workers with higher expertise.

Below the bar chart, a “Top Feedback” section presents the most frequently used critique statements, along with their parent principle, valence, and the percentage of total critiques that they represent. This section is intended to help guide designers towards the most serious issues with their design.

The Principle tabs embody a “zoom and filter” approach [6], filtering critiques to only those associated with the selected principle (e.g., Layout in Figure 3, right). Critique statements for a given principle ordered by how many worker critiques each contains. To the right of each statement name is a corresponding histogram bar, reflecting the proportion, positive/negative valence, and average expertise of critiques for that statement.

For each critique statement, the critiquer’s comment is shown along with an expertise label. We label workers in the bottom quartile of design expertise as novices, workers in the middle quartiles as competent, and
workers in the top quartile as experts. Hovering over any critique causes all corresponding annotations to appear on the design itself. "Details on demand" [6] are provided through options like collapsible critique lists, and hoverable definitions and annotations.

Evaluation and Future Work
We evaluated CrowdCrit by organizing a poster design contest in which 14 participants with a range of design experience received CrowdCrit critiques midway through their design processes. Our interviews and quantitative results showed that designers generally found the crowd feedback helpful, and designs with more negative critiques were more likely to improve. In future work, we are exploring whether focusing crowds on one design principle at a time leads to higher quality critiques.

Acknowledgments
This material is based upon work supported by the National Science Foundation under Grant No. IIS 1210836 and an Okawa Research Grant. We thank Pengcheng Tang for development assistance and Brian Bailey for helpful conversations.

References