VizCrit: Exploring Strategies for Displaying Computational Feedback in a Visual Design Tool

Mingyi Li
Northeastern University
Boston, Massachusetts, USA
li.mingyi2@northeastern.edu

Mengyi Chen
University of Pennsylvania
Philadelphia, Pennsylvania, USA
mengyic@seas.upenn.edu

Sarah Luo
Purdue University
West Lafayette, Indiana, USA
luo459@purdue.edu

Yining Cao
University of California, San Diego
La Jolla, California, USA
rimacyn@ucsd.edu

Haijun Xia
University of California, San Diego
La Jolla, California, USA
haijunxia@ucsd.edu

Maitraye Das
Northeastern University
Boston, Massachusetts, USA
ma.das@northeastern.edu

Steven P. Dow
University of California, San Diego
La Jolla, California, USA
spdow@ucsd.edu

Jane L. E
National University of Singapore
Singapore, Singapore
ejane@cs.stanford.edu

Figure 1: We introduce *VizCrit*, a design feedback tool that offers feedback with three levels of actionability: *textbook-based* feedback as static text, and *awareness-centered* and *solution-centered* feedback as adaptive visual annotations. To design the interactive annotations, we observed how expert designers and instructors provide situated feedback (left), then collaboratively co-designed a set of visual annotations (right) for four core design principles (Alignment is shown here). For each annotation, we designed algorithms for heuristically computing the annotations to be displayed as overlays on the visual design (including issue detection for solution-centered feedback). The user evaluation with novices explores how different *actionability* in feedback influences novices' process-related behaviors, learning of principles, perceptions of creativity, and overall outcomes.

Abstract

Visual design instructors often provide multi-modal feedback, mixing annotations with text. Prior theory emphasizes the importance of actionable feedback, where "actionability" lies on a spectrum—from surfacing relevant design concepts to suggesting concrete fixes. How might creativity tools implement annotations that support such feedback, and how does the actionability of feedback impact novices' process-related behaviors, perceptions of creativity, learning of design principles, and overall outcomes? We introduce *VizCrit*, a system for providing computational feedback that supports the actionability spectrum, realized through algorithmic issue detection and visual annotation generation. In a between-subjects study (N=36), novices revised a design under one of three conditions: textbook-based, awareness-centered, or solution-centered feedback. We found that solution-centered feedback led to fewer design issues and higher self-perceived creativity compared with textbook-based

feedback, although expert ratings on creativity showed no significant differences. We discuss the implications for AI in Creativity Support Tools, including the potential of calibrating feedback actionability to help novices balance productivity with learning, growth, and developing design awareness.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**.

Keywords

visual design; feedback; annotations; creativity support tools; human-AI collaboration

1 Introduction

In design education, effective feedback is central to scaffolding creative growth [6, 24, 72]. Instructor- and peer-led studio critiques (or "crits") can help students reflect, iterate, and develop disciplinary judgment, both during the act of designing and after drafts are complete [6, 24, 72]. Instructors often provide multi-modal feedback directly on students' work, externalized as a mix of visual annotations [80] and spoken or written commentary to highlight principles in context. These practices do more than convey corrections: they shape how novices see their work, explore alternatives, and build evaluative judgment [23, 40]. To investigate the potential for computational scaffolding, we explore how multi-modal feedback, including visual annotations, can be integrated into a design tool and how this impacts design learners.

Existing research characterizes high-quality feedback as *specific*, *justified*, and *actionable* to effectively guide learning [41, 66, 71, 76]. Given recent advances in AI, feedback of all styles is becoming readily available. We therefore explore *actionability*—how directly feedback prescribes changes—because it can shape both the creative outcomes and the potential for learning. On one hand, instructors typically avoid imposing specific solutions on students; they ask students to notice issues and reason about how design choices relate to underlying concepts [23, 24]. On the other hand, AI tools could be designed to provide a range of feedback, with a norm leaning towards prescriptive feedback with concrete fixes that can accelerate practice [73, 74], but might reduce opportunities for reflection [45, 66, 69, 89].

While prior work in design education has offered hints on potential tradeoffs between different levels of feedback actionability [23, 63, 71, 76, 91], we still lack solid empirical data on how different choices affect novices' learning and creativity. We explore key questions within the context of a visual design tool:

- **[RQ1]** How might we design visual annotations to support different levels of actionability? and,

- **[RQ2]** How does the actionability of computational feedback affect design process-related behaviors, perceptions of creativity, learning of principles, and overall outcomes?

As tools like generative AI can now provide immediate, always-available, and the full spectrum of design feedback [25, 35], these concerns about different feedback styles have important implications for how to position AI within design workflows and human-AI collaboration [11, 38, 92]. Understanding the impacts of different feedback actionability can help Creativity Support Tools (CSTs) designers tailor feedback in ways that support both productivity and creative growth [10, 55].

To answer **[RQ1]**, we observed how visual design experts use visual annotations as part of their critique process [23, 24]. We conducted co-design interviews with design experts (N=9) who provided critiques using different feedback actionability on a set of designs. Our analysis of these critiques informed the designs of our visual annotations for hierarchy, alignment, whitespace, and unity design principles. We embed these into *VizCrit*, a prototype for a visual design tool that delivers on-demand and principled feedback. *VizCrit* provides feedback that spans the actionability spectrum: static, *textbook-based* feedback via text only, and adaptive visual annotations that are either *awareness-centered*, focusing on drawing awareness to design principles, or *solution-centered*, both calling out issues within a design and suggesting a concrete change. To implement these annotations, we carefully designed algorithms to computationally generate the expert-informed annotations, and for the solution-centered annotations, to detect issues within a visual design, and to propose reasonable fixes.

For **[RQ2]**, drawing on prior work in education and design, we hypothesize a tradeoff: solution-centered feedback would improve immediate performance, confidence, and polish [32, 63, 75, 91], but perhaps at the cost of learning, creativity, and ownership [45, 69]. In contrast, awareness-centered feedback would be harder to act on in the moment but could scaffold better reflection, transfer, and the development of evaluative judgment [71, 76]. We also hypothesize that generic textbook-based feedback would leave too big of a gap for novices to close, or to successfully transfer what they learn to address design principle violations [12, 41]. Using *VizCrit* allows us to move beyond theoretical accounts to collect empirical data on how feedback actionability impacts novices' design behaviors. To investigate these hypotheses, we conducted a between-subjects study with novice designers (N=36) using *VizCrit* to revise a flawed visual design. Participants were randomly assigned to one of three conditions: textbook-based, awareness-centered, or solution-centered feedback, and were asked to revise a design using the feedback provided. To understand the effects of feedback actionability, we measured participants' process-related behaviors (e.g., interaction timing and frequency of the feedback requests), self- and expert-rated creativity, design principle learning from knowledge tests, and design outcomes (e.g., number of issues and degree of change).

We found that solution-centered participants gave significantly higher ratings on self-assessed creativity than textbook-based participants, although there were no statistically significant differences in how independent experts assessed creativity across conditions. Our behavioral data on how participants responded to feedback suggests that solution-centered participants may have developed a

false sense of creativity, stemming from their trust in the system's suggestions and feelings of progress. Inline with our hypotheses, solution-centered participants showed better design outcomes as measured by a statistically significant decrease in design issues compared to textbook-based participants. Contrary to our hypotheses that solution-centered feedback would hinder learning, there were no significant differences in the learning gains across conditions. Our interview data suggests that solution-centered participants still reflected on how they applied principles to their own designs.

We discuss implications for leveraging AI within CSTs and how to strike a balance between providing immediate assistance and helping novices in the long term to develop a critical eye and deepen their creative expertise.

This paper makes the following contributions:

- **Co-design study of feedback annotations with experts:** We uncover insights from visual design experts and contribute a set of visual annotations across textbook-based, awareness- and solution-centered feedback actionability.
- **System design and implementation:** We introduce *VizCrit*, a prototype that delivers feedback as visual annotations with linked text, supported by algorithms for detecting issues and generating suggestions.
- **Study on feedback actionability with design novices:** In a between-subjects study (N=36), we found preliminary evidence, despite our hypotheses, that solution-centered feedback led to fewer design issues and increased self-perceived creativity compared with textbook-based feedback. However, expert ratings on creativity showed no significant differences, suggesting that more actionability may inflate novices' perceived creativity through feedback-driven progress.

2 Related Work

We review related literature on feedback in education as well as design, highlighting the dimensions of "actionability" in each domain. We then review work exploring the potential of computational guidance for creativity, generative AI for design feedback, and the potential influence of such tools on creativity and learning.

2.1 The "Actionability" Factor in Design Feedback

Educational research emphasizes the centrality of feedback for learning, yet its effects vary with content, timing, and learner expertise [12, 22, 41, 71]. Foundational work emphasizes that feedback only improves learning if learners can close the gap between current and desired performance [12, 41, 71]. Educational researcher Shute [76] defines different feedback styles, from directive (tells learners what to fix) to facilitative (prompts learners to engage in their own correction). In our work, we expand on this to define a spectrum of actionability in feedback from the most generic, unactionable (i.e. like a textbook) to awareness- and solution-centered feedback and beyond. Studies show that while facilitative feedback such as principle-based comments support reflection, students revise more effectively when feedback includes concrete suggestions [63, 91]; in STEAM classrooms, directive feedback also produced higher creative gains than opinion-only feedback [75].

However, directive feedback reduced students' autonomy and self-reflection on the learning process, as the main focus became fixing errors [87]. These frameworks show that while directive input can accelerate revision, overly prescriptive feedback risks limiting learners' engagement in reflection and transfer.

Critique structures help learners reflect and iterate [6, 56, 61, 72]. In line with education, work in design highlights a continuum from comments that raise awareness of principles to those that propose specific changes. Researchers found that design studio critique often emphasizes principle-based comments that build disciplinary understanding without prescribing exact moves [23, 24]. Nguyen et al. [65] similarly showed that open-ended or affective framing of critique can foster receptivity and preserve a sense of ownership in student work. At the same time, studies of fixation warn that prescriptive examples can constrain creativity: designers often remain anchored to given examples, highlighting how prescriptive examples can anchor design trajectories in unproductive ways [45, 69]. Motivated by prior work, we compare the tradeoffs in learning, creativity, and performance in the visual design domain, particularly when facilitated by computational support.

2.2 Computational Guidance in Creativity Support Tools

By embedding certain assumptions and abstractions, creative support tools hold power over how users create [55, 88]. Therefore, they influence not only outcomes but also creative ownership and agency. HCI has long argued that tools should empower users, not merely optimize productivity [20, 21, 37]. The growth in AI-driven creativity support intensifies these concerns: when systems readily provide ready-made solutions or feedback, they can accelerate work but risk shifting authorship away from the creator [10, 11, 38].

Considering the lens of power in CSTs [55] provides insights on feedback actionability and how it imposes different ways of thinking. Existing tools cover a spectrum from *surfacing awareness* (which gives users power over their next move) to guidance that *proposes potential answers* (where the feedback imposes power over the user's next move). Awareness-centered tools aim to widen perception and scaffold reflection: e.g., adaptive conceptual prompts in *Shöwn* [64], principle-oriented critics and analytic overlays [36, 53], and composition guidance for photography [31]. Other systems offer more specific suggestions to accelerate revision, such as interactive layout suggestions in *DesignScape* [67] or guidance towards achieving specific portrait lighting styles [30].

Recently, a growing body of work investigates generative AI for feedback in various contexts, such as UI and web design [28, 29, 42, 90], 3D design [18], presentation slides [95], and art education [96]. Most related to our work, *UICrit* [28] generates design critiques for UI screenshots by collecting a dataset of critiques and using few-shot and visual prompting techniques with LLMs. Also related, *DesignChecker* [42] provides suggestions on web design issues by applying design heuristics-based algorithms. Most of these focus on a few specific feedback designs (often solution-centered) and tend to be primarily text-based. Our research builds on this prior work by studying the difference in impact across the spectrum of actionability and the potential of visual annotations for feedback.

Since novices are still developing domain expertise and evaluative judgment [34, 39], they are especially susceptible to tool-related influence. Recent empirical work on feedback timing shows that when and how guidance appears (in-action versus on-action) risks novices' overreliance on the tool for creative decisions [32]. These perspectives suggest that the *form* of guidance matters for both learning and creativity. Building on this literature, we use the awareness-versus-solution distinction as an analytical lens to examine *how* guidance impacts novices' learning, creativity, and performance in visual design. Inspired by work exploring how AI explanations can help calibrate trust and reduce overreliance [85], we deliver guidance as *visual annotations* that act as explanations, aiming to make system reasoning legible (mitigating overreliance) while preserving room for user judgment and exploration.

2.3 The Role of Annotations in Design Communication

Annotations are a key medium for design communication: by externalizing what matters in a layout, they help people see relations, surface issues, and coordinate critique. Work on sketching and external representations shows that marks "on the page" are not merely records but cognitive scaffolds that prompt noticing and reasoning about alternatives [80, 84]. Annotations also carry a social–interpretive function: through practices like marking and coding, experts cultivate *professional vision*, training novices to see work through disciplinary criteria and to connect abstract principles to concrete choices [39]. In addition to this conceptual role, prior systems have used interactive annotations to make principles visible in context [7, 31, 53, 57, 67]. We build on this prior work by operationalizing annotations as a form of computational feedback that seeks to make design principles explicit while varying their actionability and studying how they impact novices' design processes and understanding of design principles.

3 Co-Design Study on Feedback Annotations with Instructors

How might we design visual annotations to support different levels of actionability? To explore this question, we co-designed with experts a set of annotations for design principles. For each principle, our goal was to create a parallel pair of annotation designs: one awareness-centered and one solution-centered. To ground these designs, we relied on experts in visual design for two stages: (1) collecting existing annotation examples from educational materials to generate preliminary design probes and (2) conducting a co-design study with experts where they annotated novice designs with freeform, awareness-centered, and solution-centered feedback styles. All study materials can be found in Supplemental Materials.

3.1 Design Probes

We first collected annotations that instructors and experts in design had already created around prior materials and then performed an informal content analysis. We searched for descriptions of design principles with examples using a range of keywords around design, principles, descriptions, visualizations, examples, etc. Across 43 articles, 10 books, and 7 videos, we extracted hundreds of examples and grouped them by principles (balance, alignment, whitespace, hierarchy, unity, readability). From these, we iteratively brainstormed a preliminary set of annotations. Balance was removed since many annotations from our content analysis mixed it with whitespace. For each principle, we produced two awareness-centered and two solution-centered probes to test in parallel, to reduce bias [27]. We used these artifacts as exploratory design probes to elicit experts' annotation practices and their insights into the actionability of feedback. The results were later synthesized into final design annotations for the *VizCrit* system.

3.2 Participants

To understand professional feedback practices and evaluate our annotation designs, we recruited 9 expert visual designers (6F, 3M; age: 21–50, $\mu = 28.0$) from online design communities or our personal networks. Participants had an average of 6.4 years of experience with visual design and feedback-giving (range 1–20+), and included design instructors as well as professional product/UX/graphic designers. We administered a pre-study screening survey to collect portfolios or websites, and then selected participants based on their demonstrated expertise.

3.3 Procedure

Participants were first asked to review several novice designs and then provide feedback: first freeform, then awareness-centered, and finally solution-centered. This produced parallel examples across the spectrum and insights into design experts' feedback communication strategies. We administered the co-design study on Figma, since it is widely used in the design community and supports both quick, low-fidelity annotations (e.g., freehand drawing) and high-fidelity operations (e.g., applying blur effects). When a change was difficult to execute quickly, the interviewer assisted by implementing it in Figma based on the participants' verbal descriptions and then confirmed its accuracy.

Each study was conducted over Zoom and lasted for approximately 1-1.5 hours. Each participant was compensated with a $60 Amazon gift card. All sessions were screen-recorded to capture participants' actions and think-aloud protocols, which enabled us to observe participants' strategies in real time and collect rich qualitative data for analysis.

3.3.1 Interview. After completing a demographics survey, participants were asked questions about their current process of providing feedback on visual designs. This included the integration of feedback into their visual design process or teaching, challenges they see novices encounter while applying or addressing feedback, and whether they used visual annotations or followed design principles to communicate feedback.

3.3.2 Freeform Design Feedback. Participants were asked to critique three novice designs in Figma, each with two principle violations (hierarchy, alignment, whitespace, unity, readability). Participants were asked to annotate or write their feedback for the novice designs using freeform feedback practices. Researchers then explained the spectrum of awareness-centered to solution-centered feedback and asked the participants to annotate those same designs using the awareness- and solution-centered feedback styles.

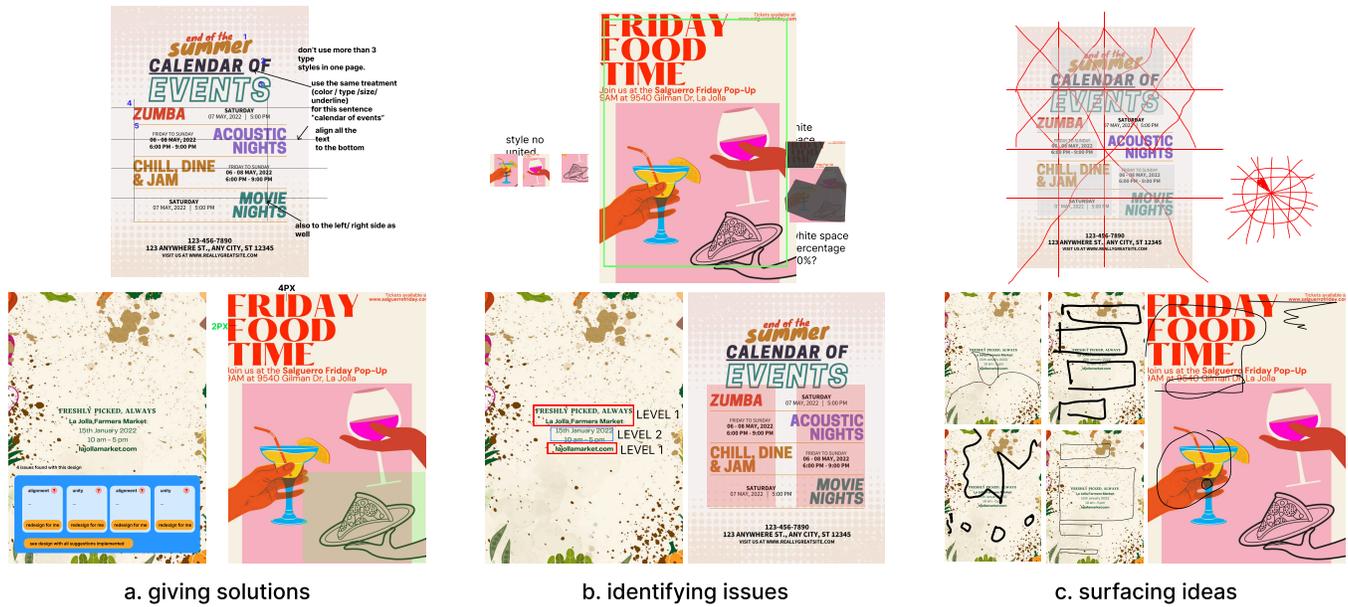a. giving solutions   b. identifying issues   c. surfacing ideas

Figure 2: Expert instructors' freeform feedback shown here across the actionability spectrum: (a) solution-centered, (b) calling awareness to issues, but not providing solutions , and (c) awareness-centered.

3.3.3 Feedback on Design Probes. Researchers showed and explained our preliminary annotation designs (two awareness-centered and two solution-centered per principle). Each annotation idea was shown on a visual design that violates exactly that same principle. Participants then reviewed the designs and proposed their preferred annotations on a new canvas: either from scratch, or borrowing ideas from our design probes.

3.4 Findings

We summarize key observations, participant practices, and perspectives on awareness-centered versus solution-centered feedback from the co-design study.

3.4.1 Experts Fluidly Move Along the Actionability Spectrum. Experts emphasized that feedback rarely fell into purely awareness-centered or solution-centered categories. Instead, they shifted styles depending on context, using neutral framing to highlight design principles and encourage reflection, and direct suggestions when issues were concrete or time-sensitive. In their freeform feedback, experts also fell along the spectrum (see Figure 2), with 3 giving solutions (2.a), 3 identifying issues (2.b), and 2 providing a mix of identifying issues or surfacing ideas to consider (2.c).

One participant reflected on the tension between efficiency and agency, noting both that "*sometimes you do want direct feedback like when you're on a crunch and need a fix right away*" (E1), and that "*not giving this right or wrong answer is important, otherwise automated feedback could dictate creativity*" (E1). Many described combining the two, beginning with neutral prompts and escalating to directed recommendations when problems persisted: "*first give neutral feedback, they might be able to identify the issue, and if it's still a problem, give directed feedback*" (E7).

3.4.2 Annotations as a Core Communication Medium. Visual annotations were a central way for experts to communicate critiques, further validating our visual annotation approach. While some solely used comments, most participants used arrows, circles, or highlights to draw attention to specific areas, often embedding explanations directly on the canvas. These marks were paired with structured mechanisms, such as sketching guide lines (e.g., to show a margin), displaying font sizes, or outlining alignment groups, to help novices better visualize the underlying design parameters and concepts. Annotations were often complemented with brief textual clarifications that allowed participants to connect what they saw from annotations with the reasoning behind them. As one participant explained, combining annotations with text made it easier to "*judge if I want that to be the case or if it's completely off base*" (E2).

3.4.3 Expert Perspectives on Awareness-centered Versus Solution-centered Feedback. Across principles, experts viewed awareness-centered annotations as lighter in tone and color, mainly serving to raise awareness or prompt reflection. By contrast, solution-centered annotations were associated with more explicit cues, often using strong colors like red and green to indicate problems and solutions. Several emphasized that awareness-centered designs provided space for creative interpretation, whereas solution-centered ones were appreciated for efficiency in straightforward cases.

During the initial interview (prior to researchers introducing the concepts), E2 mentioned the difference between what he called "*critique and direction,*" which corresponded to our awareness- and solution-centered concepts. He clarified, awareness-centered critiques should highlight what is happening, while solution-centered goes further to "*tell you how to fix it*" (E2). He noted that in practice, "*some (learners) need more critique than others. Some need a lot of direction because they just get stuck... some students don't*

Figure 3: Our design probes with a set of four annotations for each design principle (hierarchy, whitespace, alignment, unity). Participants were asked to iterate on these probes based on their own feedback-giving style. The examples shown here are design annotations resulting from their iterations: experts used Figma's drawing and editing tools to annotate, adjust layouts, and propose alternative arrangements while explaining their rationale.

need that" (E2), showing how the balance between awareness- and solution-centered feedback shifts with learner experience. While providing critiques, he provided various ideas for the same concept (see bottom-left of Figure 2.c), illustrating that there can be multiple appropriate annotations for the same principles.

3.5 Insights on Annotations and Actionability

Experts confirmed that annotation styles and mechanisms needed to be individually tailored to each design principle. Here we describe some of the lower-level annotation design ideas from the experts, and introduce our final annotation designs by principles. Figure 3 shows a set of our design probes and expert proposals, and Figure 4 shows our final set of annotation designs. Since our goal was to embed annotations computationally into a visual design tool, we applied a consistent visual language across different principles.

After rounds of iteration, we removed readability as a principle, as visualization ideas drastically differed from other principles and there were various existing accessibility checkers for issues like color contrast [86]. We also focused primarily on authoring the text elements as they were most directly related to the principles and most emphasized by experts' feedback designs. However, we envision this as an extensible platform where instructors can add custom annotations for additional principles and visual elements.

Our annotation designs (see Figure 4) use consistent language in how elements are annotated. Other than necessary exceptions,

text elements are highlighted with filled rectangles and dotted lines indicate guide lines in the canvas (such as for alignment or margins). For awareness-centered annotations, text elements are highlighted to visualize concepts relating to each design principle. Red and green are intentionally avoided to minimize potential misinterpretation of something as "correct" or "incorrect" due to their preexisting associations. For solution-centered annotations, red and green denote problematic regions and possible solutions: red boxes indicate an issue and green arrows suggest a possible action to fix it. When there is no issue, green is used to highlight elements related to the principle. All colors come from the Tableau color palette [60], chosen based on feedback from co-design experts, who emphasized the importance of neutral, non-ambiguous colors.

3.5.1 Hierarchy. Experts visualized text parameters to show relative importance, using font-size listings and overlays that previewed how larger titles would shift emphasis.

> *Awareness-centered.* Annotations use color (shades of gray to blue) to visualize levels of visual emphasis (high, medium, low). This shows how visual attention is distributed across the design.
> *Solution-centered.* Red highlights flag elements with misaligned emphasis, and green arrows suggest size adjustments to better match the level of importance (e.g., enlarging a title or reducing competing text).
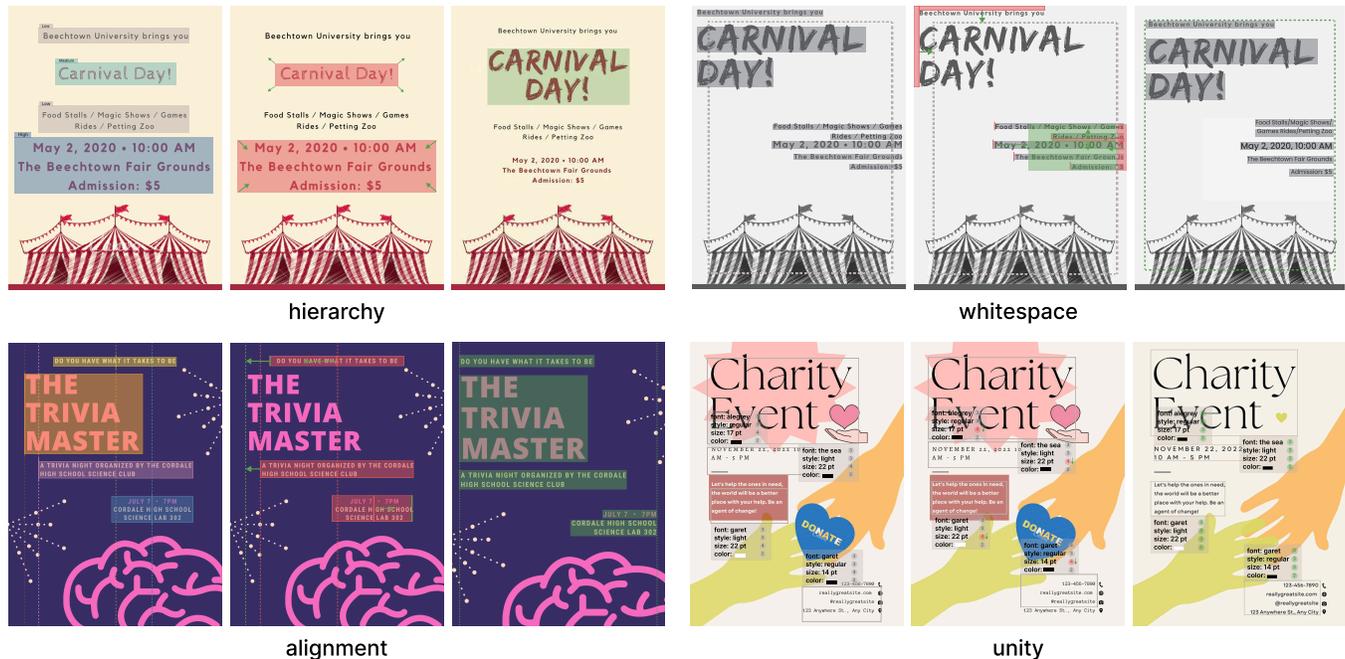
Figure 4: Final set of annotation designs across hierarchy, whitespace, alignment, and unity. For each principle, annotations are shown left to right: awareness-centered, solution-centered with issue(s), solution-centered with no issues.

No issues. When emphasis aligns with content importance, no issues are flagged. Correctly emphasized elements, such as the title, are confirmed in green.

3.5.2 Alignment.
Experts preferred structural overlays such as grid lines, margin guides, or colored grouping highlights to expose inconsistent axes, which they considered precise and easy for novices to interpret.

Awareness-centered. Annotations show external alignment across multiple elements (grouping by shared edges) and internal alignment within single text element (left, right, center, justified).
Solution-centered. Red highlights mark alignment inconsistencies, while green arrows suggest consolidating groups or matching internal alignment with external positioning. For example, the annotations could recommend left-aligning multiple text elements to reduce the number of alignment groups or making right-positioned text be right-aligned internally.
No issues. When alignment is consistent, no elements are flagged. Correct groupings and alignments are shown in green.

3.5.3 Whitespace.
Experts recommended margin indicators and highlighting the space between elements, as shown in many of their iterations on our designs (Figure 3). Some suggested simplifying annotations by removing pixel values, while many proposed using grayscale transformation to contrast positive and negative spaces.

Awareness-centered. Text is abstracted into gray boxes against white background, highlighting the empty space between elements and margins. Dotted lines indicate suggested boundaries for spacing.
Solution-centered. Red highlights mark insufficient or uneven spacing, while green arrows or boxes suggest increasing margins or evening out line breaks. For example, the annotation could recommend shifting elements inward to create more consistent spacing.
No issues. When spacing is sufficient and consistent, no issues are flagged. Margins and gaps are confirmed in green.

3.5.4 Unity.
Experts emphasized embedding parameters into the design, for example, generating color palettes, showing percentage breakdowns of area covered by different hues, and overlaying variations in fonts and styles to make inconsistencies more visible.

Awareness-centered. Each text element is annotated with its font family, style, size, and color, along with counts of distinct values. This shows how many variations exist in the design.
Solution-centered. Counts are highlighted in red to flag excessive variation, while green arrows suggest reducing counts (e.g., limiting the number of total font sizes) to strengthen consistency.
No issues. When the number of variations is within acceptable limits, no issues are flagged. Current counts are shown in green to confirm visual cohesion.
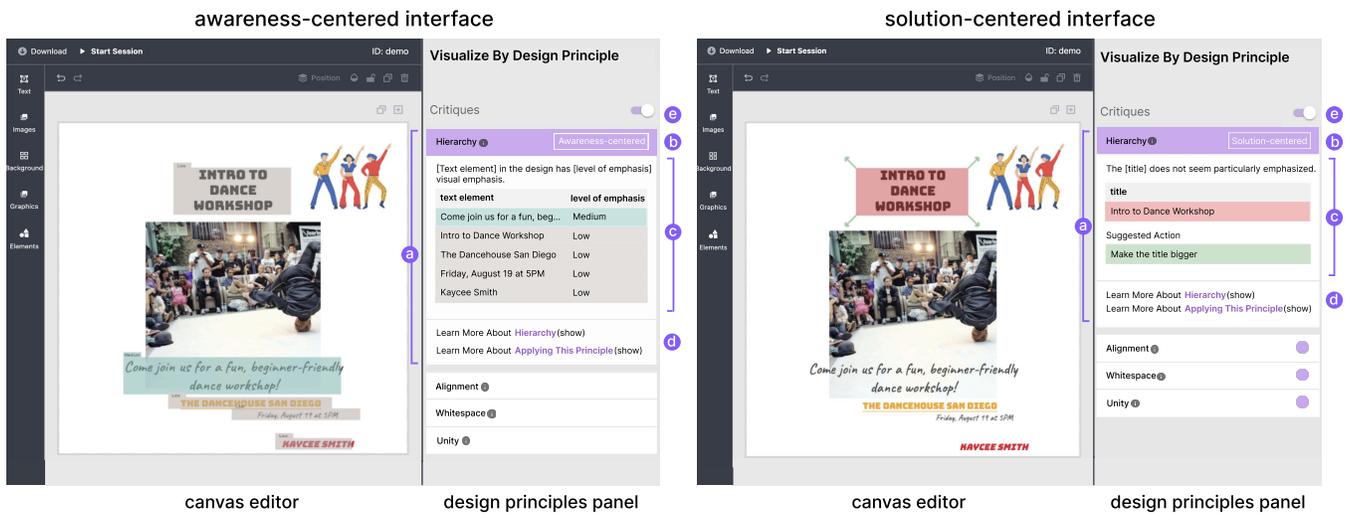
awareness-centered interface   solution-centered interface

Figure 5: *VizCrit* **interface: pictured here with hierarchy annotations, awareness-centered on the left and solution-centered on the right. The interface is primarily split into two components: the** *Canvas Editor* **where users design and see their situated visual annotations, and the** *Design Principles Panel* **with further text information about the principle and annotation: (a) provides further information about the principle annotation and is expanded by selecting (b), (c) provides a text explanation of the annotation, and (d) provides additional textbook-based information about the principle. For solution-centered annotations, users can use (e) to toggle on and off the purple dot notifications that indicate when there is an issue for a given principle.**

4 VizCrit System

We introduce the design and implementation of *VizCrit*[1], a visual design tool that automatically generates and displays feedback as annotations (see Figure 5). We describe how we translate instructor critiques into computational heuristics that can generate design annotations, detect issues, and suggest solutions in real time. Our current implementation relies on heuristics-based algorithms rather than generative models, because we aim to provide granular and interpretable annotations grounded in principles and annotation practices from design instructors and professionals.

The system is composed of two main components: a *server* for performing algorithmic computations for awareness- and solution-centered annotations, and a *user interface* that provides the design canvas and surfaces annotations to the user (see Figure 6). The user interface is built on top of Polotno[2], an open-sourced web-based visual design tool and the server is a Python backend with a WebSocket that facilitates communication between them. A detailed description of how the backend server communicates with the user interface is included in Appendix A.4.

4.1 User Interface: Surfacing Annotations in a Visual Design Tool

The user interface has two main sections: *Canvas Editor* and the *Design Principles Panel*. The editor allows users to create visual designs using elements like text, images, and graphics. On the right side of the canvas, the panel offers control for real-time and on-demand design annotations.

4.1.1 Design Principles Panel. The main portion of the panel comprises expandable areas per principle: Hierarchy, Alignment, Whitespace, and Unity. The user can click on a principle name (5.b) to expand that section. Upon click, the action is also reflected in the design editor through showing the corresponding annotation as an overlay to the canvas. Within the principle, there are two components: *Annotation Explanation* (5.c) and *Principle Details* (5.d).

4.1.2 Annotation Explanation. To help users interpret design annotations, annotation explanations (5.c) are provided on the right panel as a text description of the visual annotation in a table format (see Figure 7). It has a consistent color-coding with the design annotation and replicates all visual information into a text representation. These are designed to be as close as possible to a direct mapping so they communicate equivalent information. For the solution-centered version of annotation, there is an extra "Suggested Actions" field that translates the green arrows into their implied actions for fixing design issues. When the annotation is occluded by the design, the user can also refer to the annotation explanation for more details. Appendix 2 shows a comprehensive list of annotation explanations.

4.1.3 Principle Details. Static textbook-based references for each design principle (5.d) are provided. To avoid information overload, all additional details are available only after the users click the "show" button. The first collapsed information helps in understanding the design principle, including its definition and a common issue. The second collapsed information helps in applying the principle by listing a relevant example and actions that could be taken, such as increasing the prominence of title text to follow Hierarchy principle. This is the design principle-related information available in

---

[1]VizCrit Code and other study materials can be found on our project page: https://ejane.me/vizcrit.html
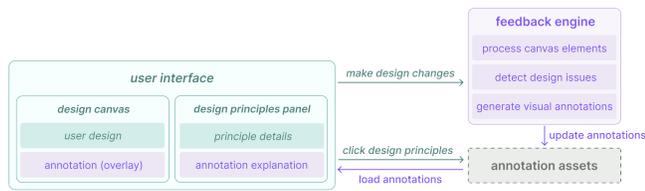[2]https://polotno.com/

Figure 6: A system diagram from the user's perspective. The purple color indicates interactions with feedback and design annotations, and the cyan color with italic text style indicates user interactions. Whenever the user makes any design changes on the canvas, the frontend (interface) will send a request along with the current design state to the server (feedback engine). The feedback engine then analyzes the design and makes updates to the annotation files in annotation assets. Annotations will be loaded to the interface whenever the user clicks design principles on the principles panel to request feedback. This ensures the efficiency of our real-time annotation rendering.

VizCrit across all feedback conditions. See Figure 1 and Appendix 3 for more principle details.

4.1.4 Critiques Toggle. In the solution-centered condition, the purple dot next to a principle (5.b) indicates potential design violations of that principle along with feedback. If there are no design issues, the dot will have a lighter color and a check sign. Users can use the critiques toggle (5.e) to turn on/off the notifications. Since issues are only surfaced in solution-centered annotations, this is hidden in awareness-centered and textbook-based conditions.

4.2 Server: Generating Annotations Computationally

We introduce heuristic algorithms for generating annotations for each design principle. For awareness-centered annotations, this mainly involves analyzing the design elements to generate visual annotations. For solution-centered annotations, this additionally involves detecting issues and determining reasonable solutions. We describe initial steps for processing text that are applied across multiple principles and then describe the algorithms per principle.

4.2.1 Processing Text Elements. OCR analysis is performed on text elements to measure their spacing, extracting text structure with OpenCV contour detection[3] to obtain their precise pixel-level bounding boxes. This is particularly needed for Hierarchy and Whitespace principles as they require inspection for the space physically occupied by the text. Then, for each text element, text line heights are determined using HarfBuzz text shaping library[4].

4.2.2 Hierarchy: Examine Visual Prominence. To analyze the level of emphasis of text elements, K-means clustering algorithm is performed on the elements' line heights to form two or three different groups. Then, the cluster centers are used to compare with font size thresholds for determining the low, medium, high priority groups. These are directly visualized for awareness-centered annotations.

---

[3]https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html
[4]https://harfbuzz.github.io/

For solution-centered annotations, the system additionally detects issues through identifying a weak (or no) point of entry. If there are no elements in the high priority group, the system will report a "No elements seem to be visually emphasized" issue. If multiple elements exist in the high priority group, the system will report non-title texts as competing elements. In both cases, the system will suggest that the user make the title larger. In the second, it will additionally propose reducing the size of the competing element.

4.2.3 Alignment: Obtain Horizontal Groups. From the canvas JSON object, text elements are filtered, and their x-coordinates and alignment attributes are used to calculate how elements are horizontally positioned. Next, they are grouped based on their horizontal alignment, and the system will indicate "too many alignment groups" issue if several text elements are aligned to different axes. Additionally, the group's width is computed to determine left/right boundaries and to check if their external positions on the canvas match their internal alignment. Issues will be reported if elements are misaligned (e.g., text on the left side of the canvas is center-aligned). In the first case, the system will suggest aligning elements to fewer axes. In the second, it will propose matching the external and internal alignment.

4.2.4 Whitespace: Perform Spacing Analysis. The awareness-centered annotation directly surfaces the text processing described earlier to showcase the boundaries around text elements and highlight the whitespace between them. For solution-centered issue detection, three distinct whitespace analyses are performed: canvas margin, inter-element spacing, and ragged text edges. To detect whether elements are positioned too close to the canvas edges, we set a minimum threshold based on the canvas size. If below the threshold, the system will direct the users to move elements farther from the edge. The detection of insufficient spacing between elements involves a pairwise analysis. First, filtering logic is conducted to exclude elements that are far apart. Next, elements are compared if they are horizontally and vertically adjacent to one another. The system will suggest moving two text elements farther apart from each other. Ragged edges occur when a text block has notably different lengths in each line, creating visual disruption for the reading flow. After obtaining the widths from the OCR analysis, the shorter lines are compared with the longer lines to determine whether they are significantly shorter. The system's suggestion will be to even out the vertical lengths.

4.2.5 Unity: Count Text Variances. The JSON objects of text elements are used directly to count the variations in font family, style, color, and size. These counts are surfaced per text element in the awareness-centered annotation. In the solution-centered annotation, if any text elements have more than three variances for any text properties, the system will show the count in red and indicate a "too many variances in text" issue. In the visual annotation, a down arrow is used as a signal to reduce the number of text properties to three or fewer.
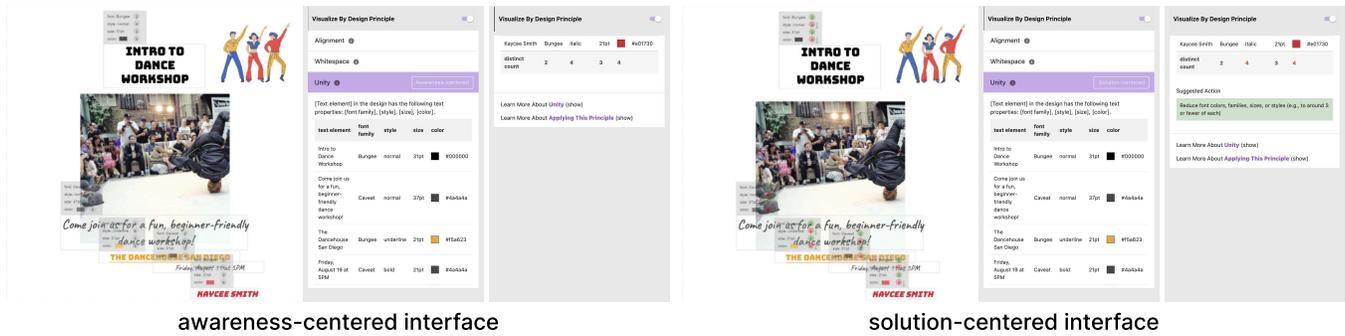
Figure 7: Each of our visual annotation had a corresponding text explanation on the right panel as tables. Pictured here are the unity annotations with their explanations to help interpret visuals: awareness-centered (left) and solution-centered (right).

5 Technical Evaluation

Annotation generation, issue detection, and solution suggestion algorithms are evaluated systematically. Among them, issue detection algorithms (for hierarchy, alignment, whitespace, and unity principles) are the primary focus.

The evaluation dataset came from a public dataset [32], containing 20 Polotno Studio participant designs. Authors generated 26 additional test cases during the implementation process to capture every issue instance within each principle and particularly focus on the edge cases. Algorithms were compared against design expert ratings from the original dataset as ground truth, and the authors reviewed all test cases for accuracy. Ambiguous cases (6 hierarchy, 3 alignment, 4 whitespace, 5 unity), where designs were labeled as issues in the dataset but not by our focused heuristics, or vice versa, were marked N/A and produced a modified ground truth set.

Our issue detection algorithm demonstrates effectiveness in identifying design issues, achieving > 0.85 (out of 1) accuracy across principles using modified ground truth. Here is a detailed breakdown for each principle (expert/modified): Hierarchy: 0.95/1; Alignment: 0.79/0.88; Whitespace: 0.84/0.93; Unity: 0.7/0.93. For the annotation generation and solution suggestion, three authors visually inspected system-generated annotations overlaid on top of original designs in Miro, producing 152 verified cases. We iterated and refined the heuristics until annotations appeared appropriate across all test designs.

6 Study on Feedback Actionability with Design Novices

Since *VizCrit* enables different levels of feedback actionability, it provides an opportunity to study how these differences impact novices' creative processes, particularly the tradeoffs between immediate performance and creativity. We conducted a study to explore RQ2, which we expand here into sub-questions:

- **[RQ2a]** How do novices engage with *VizCrit*'s feedback and annotations throughout their design processes?
- **[RQ2b]** How do different levels of feedback actionability influence perceptions of creativity, learning of underlying design principles, and design outcomes?

6.1 Participants

We recruited 36 participants (28F, 7M, 1 prefer not to say; age: 18–35, $\mu = 21.4$) through UC San Diego's SONA system[5], a participant recruitment and study management tool. All but one self-identified as having little to no experience in visual design (one identified as intermediate).

6.2 Study Procedure

We adopted a between-subjects design where participants had the same access to the feedback panel within *VizCrit*, which offered feedback with varying actionability: awareness-centered, solution-centered, and textbook-based. Participants were provided with an initial seed design to revise that consisted of one image, one graphic element, and five text elements and contained one design issue per principle (i.e., four total violations). We counterbalanced the three feedback types and two seed designs—each participant worked with one type of feedback actionability and one seed design. The study sessions were conducted over Zoom and we obtained verbal consent from participants to capture a screen recording. The sessions lasted 60-90 minutes and participants received credits through SONA.

6.2.1 Pre-Test. After completing the demographics survey, participants were given a knowledge test to examine their existing knowledge of design principles. We designed the test following Bloom's Taxonomy [13]: asking participants to remember, understand, and apply visual design principles. The test comprised questions with two levels of difficulty: easy (Vocabulary Test) and hard (Principles Application Test). Participants were first given a design principles reference sheet that contained each design principle's name, definition, common issue, and example. For the vocabulary test, participants were asked to name the principle that a poster pair addressed, and for the application test, they were asked to identify different design issues at specific locations within a design.

6.2.2 Tool Tutorial & Practice Task. Participants were guided to perform a series of editor operations (adding text, changing backgrounds, etc.) to get familiar with the tool's editing functionality. Next, they were provided with an intentionally simplistic starter design with clear violations across four principles as a practice task to address issues *without* feedback.

6.2.3 Annotation Tutorial & Design Task. Participants were provided with a document that explained every design annotation that they might encounter in the later design task. They could also explore the practice task with design annotations to further familiarize themselves with *VizCrit*. Participants were then introduced to the main design task: revise another (more visually complex) seed design to make it visually appealing and fit their creative preferences. We intentionally provided a shared flawed seed design to control for differences in participants' initial design skills—this helped ensure that all participants encountered at least one issue per principle, making evaluations on participants' design outcomes and creativity more comparable across conditions. Participants were asked to think aloud and use the task as a learning opportunity to develop their design skills.

6.2.4 Post-Test. Participants repeated the knowledge test as a post-task evaluation. The difference between the pre- and post-test captures the degree of learning as a result of the activity.

6.2.5 Survey & Interview. Participants completed a survey consisting of the Creativity Support Index [20] with 5-point Likert-scale ratings that assessed their perceptions of their designs, design knowledge, experience with *VizCrit* and annotations, and some AI-usability metrics [68]. The study concluded with an interview focusing on the participants' overall design and learning experience.

6.3 Data Collection and Analysis

We describe the data collected during the studies across knowledge tests, design tasks, surveys, and interviews and how we analyze the data.

6.3.1 Qualitative Analysis. To understand participants' perceptions of *VizCrit*'s feedback, we conducted qualitative analysis on the interview data. Three authors open-coded all the interviews on Condens[6], following a reflexive thematic analysis method [14] and taking a combination of inductive and deductive coding approaches. Codes captured participants' reactions and perceptions of annotations/principle panel and their desired improvements for the tool. Authors met regularly to discuss data and codes, which were then aggregated on a whiteboard to generate final themes.

6.3.2 Behavioral Data. We analyzed participants' interaction with *VizCrit*, including their responses to annotations and design principle clicks, along with the evaluation on *VizCrit*'s feedback quality.

Response to Annotations. To further capture how participants responded to design annotations, whether they immediately accepted the system's suggested solutions (in solution-centered condition) or made relevant changes to address design principles (in awareness-centered condition), and which design iterations were directly influenced by *VizCrit*'s feedback, we coded participants' subsequent design changes after they viewed annotations.

System Usage Logs and Analysis. We logged participants' interaction data using Firebase Firestore Database[7]. The logs captured 1) principle clicks (timestamp and principle name), 2) every design annotation viewed (either when a principle panel was opened or

when annotations were refreshed within an open panel), and 3) participants' on-canvas design states. Both design annotations and participants' designs were stored as JSON. The principle clicks data was cleaned by removing any clicks outside of the design task and visualized for analyzing the total number, frequency, and timing of feedback requests.

Feedback Quality. Our technical evaluation focused on accuracy for a test set. To further evaluate the feedback quality actually seen by participants, we randomly selected three design annotations that each participant viewed among their early, middle, and late feedback requests. Two authors with design expertise rated on the accuracy of annotations (0: inaccurate, 0.5: partial, 1: accurate).

6.3.3 Quantitative Analysis. Since our study involved three different types of feedback, we performed statistical analysis using Kruskal–Wallis H test [50]. If the Kruskal-Wallis test reported a significant effect, we then conducted pairwise post-hoc comparisons using Mann–Whitney U test [59] with Bonferroni correction. For comparisons only involving two independent groups (e.g., awareness-centered versus solution-centered annotations), we also used a Mann–Whitney U test [59]. Given the sample size, some comparisons were not statistically significant (see Table 1). We mark these all as 'non-significant' in our reporting.

Self Assessments. We administered the Creativity Support Index [20] but marked the 'Collaboration' dimension as 'Not Applicable' (i.e., 0 in the 0-10 scale) for all participants for consistency, since participants completed the design task individually.

Expert Assessments. To evaluate the participants' design outcomes, we also wanted to bring in a more objective perspective from experienced designers. Since several of the co-authors have significant expertise in design, we performed the following author-based expert evaluation (see Figure 9 for examples). All evaluations were individually graded, reconciled, and then averaged.

- **Degree of Change.** We measured the degree of change between participants' final designs and initial seed designs to understand the extent that they iterated on their design. Two authors developed a design iteration metric based on *each* canvas element: images (content, layout), graphic elements (content, layout), text elements (style, color, size, font, layout), and background color. For the images and graphic elements, authors coded participants' design iterations in a randomized order and reviewed the results together to resolve any inconsistencies. The text elements and background color iteration were programmatically compared through design JSON files.
- **Creativity.** We used a subjective score for the creativity of final designs (10-point, 1: bad, 10: good).
- **Design Quality.** We measured design quality by counting the number of design principle violations. Two authors with design expertise rated the number of issues in participants' final designs. We developed 4-point scoring rubrics covering four design principles.

---

[6]https://condens.io/
[7]https://firebase.google.com/

Table 1: An overview of quantitative results from participants' self-assessed Creativity Support Index (CSI), expert evaluations, feedback interaction analysis, and the learning test. The asterisks denote the level of statistical significance: * indicates $p < .05$, and ** indicates $p < .01$. Conditions denote the pairwise comparisons that were statistically significant, and for the comparisons that were not shown in table, they were all insignificant.

Category	Evaluation	Significance	Comparison
Design Quality	Number of design issues	$p < .05$ (*)	Solution-centered > Textbook-based
Creativity	Score from Creativity Support Index (CSI)	$p < .01$ (**)	Solution-centered > Textbook-based
	Enjoyment (CSI)	$p < .01$ (**)	Solution-centered > Textbook-based
	Results Worth Effort (CSI)	$p < .01$ (**)	Solution-centered > Textbook-based
	Creativity score from Expert Ratings	non-significant	non-significant
Degree of Change	Number of design iterations	non-significant	non-significant
Interaction with Feedback	Number of design principle clicks	$p < .01$ (**)	Awareness-centered & Solution-centered > Textbook-based
Design Principle Application	Improvement in the learning test	non-significant	non-significant

6.3.4 Knowledge Test (Learning). Two authors with design expertise developed 16-point scoring rubrics and graded pre- and post-knowledge tests. After grading individually, we reviewed any outstanding scores that had differences greater than one. We revisited these results and reconciled the differences. All scores were then averaged. Then, we measured learning by computing the differences between the post-test and pre-test scores. We analyzed the results separately for Vocabulary (easy) and Principles Application (hard) sections to capture more nuanced learning outcomes.

6.4 Results

In this section, we describe how participants engage with feedback, their perceptions of visual annotations, and the impact of different levels of feedback actionability on process-related behaviors, learning of design principles, perceptions of creativity, and overall design outcomes. Results are organized by sub-questions of **[RQ2]**.

6.4.1 RQ2a: How do novices engage with VizCrit's feedback and annotations throughout their design processes?

Motivations and Timing for Requesting Feedback. Participants across different feedback actionability conditions demonstrated distinct motivations and behavioral patterns for accessing *VizCrit*'s feedback. To analyze at what times and design phases participants sought feedback, we divided them into "in-action" participants who periodically requested feedback and "on-action" participants who mainly engaged with feedback at a later stage after completing a substantial amount of work [32]. For example, P30 (solution-centered) who was "in-action" reflected that she checked annotations throughout the design process: "*anytime after I adjusted something. I just wanted to make sure everything was lined up.*" For P20 (solution-centered), in contrast, exemplified "on-action" participant who only used annotations to "*fine tune*" his design.

Most textbook-based participants were on-action who checked feedback only at the end (on-action versus in-action in textbook-based = 8:3, N/A: 1; awareness-centered = 4:8; solution-centered = 6:6). Indeed, both awareness- and solution-centered participants requested significantly more feedback than textbook-based participants (awareness-centered: $M = 12.83, SD = 7.65$; solution-centered: $M = 9, SD = 4.59$; textbook-based: $M = 3.91, SD = 1.76$; $p < .01$). Since textbook-based participants only received generic and static educational content and they had already gained some design principle knowledge from the pre-test reading, they found less value in referencing principles repeatedly. As P12 stated, "*I don't think I was using it so much throughout the process... at the end I looked through it to review that I had done everything. I hadn't forgot anything.*" Therefore, for these on-action participants, design annotations or principles became like a final checklist to help identify remaining issues and validate participants' design choices.

Participants treated *VizCrit*'s feedback not only as a checker, but also as a visual inspection tool that offered insights into unnoticed details in awareness- and solution-centered conditions. Design annotations augmented participants' visual ability to locate issues that would otherwise be impossible to find. As P7 (awareness-centered) commented, "*I was able to see exactly what parts were not looking the best and where I should put them... the whitespace one was the most convenient because I could actually see the borders and the margins and stuff cause my eyeballing is very bad*" Many participants also used design annotations to find inspiration or to get unstuck. For example, P20 (solution-centered) mentioned that he previously wanted to center-align all text elements, but then *VizCrit* offered a left alignment suggestion, and he decided to adopt a new layout that aligned elements only on the left and right sides. He said, "*it (VizCrit) helped me explore different things that I wouldn't have thought of previously. And I really appreciate that.*" Relatedly, P15 (awareness-centered) reflected using *VizCrit* when he was hesitant

about where to place new elements: "*if I'm struggling with seeing where to add or where to place a text, I would use a design tool as (it) can help me with more direction.*"

Some Participants Desired More Actionability. Several participants expressed a desire to move even further along the actionability spectrum and shift away from reflective thinking toward productivity-focused moves. They wanted an automated system that could help them resolve issues more easily and efficiently. For example, P1 (awareness-centered) wished to have a click-and-apply feature that reduced the effort of self-identifying and addressing issues: "*you could just click on each principle and it was just automatically applied to it. And then you can see which one looks better or something. And then you could just like keep it.*" Similarly, P24 (awareness-centered) requested performing a simple action on the right panel, not the canvas, to make texts more consistent. She suggested, "*there's a different type of font and everything, click and then make all this (different) font to one same font from here (the right panel).*" Rather than engaging with their own designs, participants were ready to hand some creative agency to the system to achieve better and quicker design results. In addition to automatically fixing issues, some awareness-centered participants also requested more direct hints and guidance (similar to our solution-centered feedback) to reduce their effort in self-evaluation. As P22 proposed, "*a check to see if everything looks good... So hierarchy is done and hierarchy is checked or something like that to see oh that's good.*" This aligned with the predictions from our co-design experts who thought solution-centered annotations could accelerate the design process, but there should be a balance between efficiency and creative ownership. Participants' requests reflect an over-simplification of the iterative design process and entirely turn it into a series of automated check-click-apply steps that hinder creative exploration, agency, and ownership.

Participants further requested mechanisms to measure creativity or visually appealing aspects of a design. They requested "*a visual appeal score*" (P1) and "*some annotations on the creativity part or on the trends part*" (P2). This may indicate that novices could have more trust in the computational tool to the point that it overrides their own intuition, judgment, or even creative thinking abilities. Sliding too far towards full actionability and entirely trusting the tool may lead novices to over-reliance and limit learning opportunities, as they experiment less and reflect less critically on their own choices.

Awareness- and Solution-centered Annotations Were Effective and Comprehensible. In both awareness- and solution-centered conditions, participants found that visual annotations were easy to use and understand, simple, straightforward, responsive, and effectively communicated key principle information. Particularly, all awareness- and solution-centered participants (N=24) agreed on the comprehensibility of *VizCrit* annotations (see Figure 8), appreciating that the tool was "*super cool*" (P3), "*very easy and just made sense*" (P20), and "*very accurate*" (P24). Our expert ratings on the feedback quality also reflected high accuracy for both awareness- and solution-centered annotations (solution-centered: 89%, awareness-centered: 92%, among selected annotations).

One primary reason that participants found annotations comprehensible was that *VizCrit* applied effective color-coding to not only text elements, but also their corresponding explanation text on the

right panel. This way, solution-centered participants could easily locate issues and solutions by examining red and green highlights and awareness-centered participants could easily differentiate elements' level of emphasis or alignment groups by checking the group color. As P17 noted, "*I like the color coding again and the dotted lines. I think they're helpful visually because I'm a visual learner.*"

Furthermore, participants enjoyed design annotations that were situated and specific to their own designs. Since the annotation also quickly refreshed based on their design changes, participants found such frequent updates effective. P35 (awareness-centered) mentioned that "*it's not just telling me what it (design principle) is, it's actually giving me the info relative to my design currently.*" P16 (solution-centered) stated that "*when I changed something, it also changed really fast, which was good to keep up*" and P9 (solution-centered) further added "*getting live feedback on what your actions were doing ... I can like test out my ideas and see how the computer would react to it, which was really nice.*" This indicated high interactivity and usability of *VizCrit*, since participants not only passively received critiques, but also proactively engaged with annotations to explore different creative possibilities. For textbook-based participants who did not receive visual annotations, they requested more visual aids, examples, or on-canvas highlights that could help them better understand or apply design principles (P13, P23, P25, P29, P31). For example, P23 proposed having a more situated visual indicator that helps users better align their canvas elements: "*show some lines to see what's being aligned or out of place... editor could show different shapes of sections of the page to (help users) see if everything is aligned well.*" This resonated with *VizCrit*'s visual annotations that offered in-context visual signals for evaluating and improving the design.

However, participants still identified areas for improvement. They found some design annotations hard to interpret, especially for the solution-centered internal alignment annotation (see Figure 4), where *VizCrit* used arrows to suggest changing the internal text alignment. Participants misunderstood it as horizontally moving the element. To address this, *VizCrit* should provide a clearer annotation explanation to clarify the concept of internal/external alignment. Also, a different type of arrow could be used to represent different meanings (e.g., concrete arrows for moving elements, dashed arrows for changing alignment). In our expert evaluation on feedback quality, although there are no inaccurate annotations, some received partial scores (0.5/1) due to the pre-defined thresholds in our heuristics-based algorithms (e.g., size or distance thresholds for differentiating priority groups in hierarchy or alignment groups).

Moreover, while mostly efficient, updates to design annotations could experience inconsistent delays due to the cloud rendering service that converted canvas JSON to images. As a workaround, some participants chose to wait or work on other parts of their designs during the refresh, but for participants who wanted immediate feedback after an iteration, this caused confusion and frustration.

Annotations Enhanced Confidence. Many awareness- and solution centered participants commented that they gained confidence after receiving feedback annotations. Participants considered *VizCrit* as "*an objective standard*" (P3, solution-centered) that could reflect how audiences perceive their design. Thus, when *VizCrit* reported no issues or aligned with participants' self-evaluation, participants
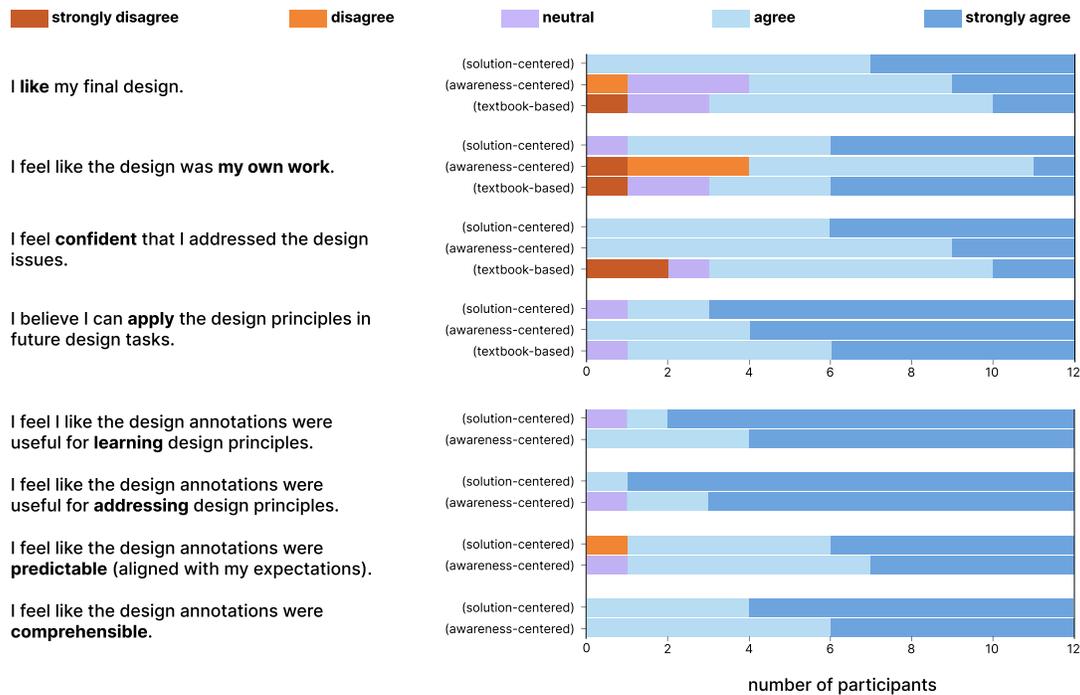
Figure 8: Distribution of participants' ($n = 36$, 12 per feedback condition) self-ratings on their final design, creative ownership, confidence, design knowledge, and overall experience with design annotations (awareness-centered or solution-centered). There were no statistically significant differences in these ratings. However, the ratings on whether the design felt like "my own work" seem to contradict our hypotheses about creative ownership. While we expected solution-centered participants to have the least ownership, they felt it equally as much if not more compared with other feedback conditions.

gained confidence as their designs were objectively validated. P33 (awareness-centered) explained, "*with eyeballing you can kind of make mistakes... but with this tool, it helps to make sure confidently everything's aligned, everything is visually going to be appealing... it just helps boost confidence (so) you don't have to double check.*" The computational inspection and external validation from *VizCrit* helped participants build confidence in their design outcomes.

Furthermore, having an objective inspection could foster collaboration and bridge peer communication. P3 recalled an example of her roommate, who thought designs from collaborators were "*just not right, design-wise, conceptually,*" but she did not know how to communicate her frustration and concerns. Therefore, P3 felt "*something like this is really helpful and diffusing that tension and making design principles more accessible and known.*" In this case, *VizCrit* could be a mediator of design conversations, which quantified and visualized design issues using established standards. It provides references for peer discussion and smooths the tension of providing and receiving design critiques.

6.4.2 *RQ2b: How do different levels of feedback actionability influence perceptions of creativity, learning of underlying design principles, and design outcomes?*

Solution-centered Feedback Inflated Self-perceived Creativity. Participants with solution-centered feedback felt significantly more creative than textbook-based participants (solution-centered: $M =$

83, $SD = 12.06$; awareness-centered: $M = 71.61, SD = 18.78$; textbook based: $M = 65.08, SD = 12.86$; $p < .01$). They enjoyed using *VizCrit* more ($p < .01$) and felt that their designs were more worth the effort than textbook-based participants ($p < .01$). However, despite participants' stronger sense of creativity, the expert ratings showed no significant differences on creativity (solution-centered: $M = 5.04, SD = 2.47$; awareness-centered: $M = 4.96, SD = 2.02$; textbook-based: $M = 7.04, SD = 2.12$, non-significant) or the degree of change across conditions (solution-centered: $M = 3.17, SD = 1.39$; awareness-centered: $M = 3.44, SD = 1.35$; textbook-based: $M = 4.12, SD = 1.18$, non-significant).

To contextualize this misalignment, we conducted an exploratory analysis of participants' behavioral responses to annotations. Of all design annotations viewed, the immediate feedback adoption rate [8] was 82% for solution-centered participants and 42% for awareness-centered participants. We also observed several concrete instances suggesting that solution-centered participants felt an urge to implement *VizCrit*'s feedback even when they felt confused. Taking the 'weak point of entry' issue as an example, after viewing annotations, P6 (solution-centered) increased the title size but the alert persisted. Although surprised, she continued following the feedback instead of questioning *VizCrit* or considering alternative

---

[8]We measured the immediate feedback adoption rate as the proportion of viewed feedback instances that were followed by a design change that acted on that feedback (see Section 6.3.2, Response to Annotations).

		design		design quality	creativity metrics		
		seed	final	issues	parameters changed	degree of change	creativity score
textbook-based	P14			1. weak point of entry 2. insufficient/ragged spacing	1. image content 5. text style 2. graphic content 6. text color 3. graphic layout 7. text size 4. background 8. text font color 9. text layout	9.25	8.5
textbook-based	P29			1. weak point of entry 2. arbitrary alignment of elements 3. insufficient/ragged spacing 4. too many variations in text	1. image content 6. text color 2. image layout 7. text size 3. graphic content 8. text font 4. graphic layout 9. text layout 5. background color	11.75	9.5
awareness-centered	P15			1. weak point of entry 2. insufficient/ragged spacing	1. image content 6. text color 2. image layout 7. text size 3. graphic content 8. text font 4. graphic layout 9. text layout 5. text style	13.25	8
awareness-centered	P24			1. weak point of entry 2. too many variations in text	1. image content 6. text size 2. graphic content 7. text font 3. graphic layout 8. text layout 4. background color 5. text color	12.5	5
solution-centered	P36			• No issues	1. image content 2. text color 3. text size 4. text font 5. text layout	5.5	4
solution-centered	P28			• No issues	1. image layout 6. text size 2. graphic content 7. text font 3. graphic layout 8. text layout 4. background color 5. text color	9.75	4.5

Figure 9: Example participant designs and expert ratings across conditions. Experts annotated design issues and parameter changes, and provided ratings of creativity (1: low, 10: high) and degree of change relative to the seed design.

solutions: "*bigger? OK, maybe I'll make this (size) 55.*" She felt compelled to act and even prepared to make title larger again while the new annotations were loading. After seeing the updated annotation indicated no issue, she immediately moved her cursor away from the title. For awareness-centered participants, however, they better retained their creative agency and were more likely to stop iterating even if the annotations did not completely align with their expectations. For instance, P33 began by making the title bigger and then felt unsure about the feedback: "*I'm making it (title) bigger. It's still saying medium level of emphasis, but I'm making it bigger.*" Then she decided, "*I'll be satisfied with that,*" and stopped iterating despite the annotation still marked the title as medium level. The more actionable and solution-centered annotations seem to enable feedback-driven design iterations, which could bring feelings of progress and creativity, but potentially limiting agency and opportunities for self-experimentation on different design options.

Solution-centered Feedback Reduced Design Issues and Did Not Hinder Learning. Solution-centered participants improved their design quality by reducing the number of issues in their final designs (see Table 1). Our expert ratings revealed that solution-centered designs had significantly fewer design principle violations than textbook-based designs (solution-centered: $M = 0.75, SD = 0.72$; awareness-centered: $M = 1.54, SD = 0.78$; textbook-based: $M = 1.92, SD = 1.04; U = 118, p < .05$). This aligns with our hypothesis that the most actionable feedback can boost performance.

Contrary to our hypothesis that emphasizing performance might harm learning, we found no significant differences in the principle application learning test across conditions (solution-centered: $M = 1.04, SD = 2.16$; awareness-centered: $M = 0.73, SD = 1.64$; textbook-based: $M = 0.33, SD = 1.01$, non-significant). Our interview data suggest that solution-centered feedback did not hinder learning and participants still reflected on how they applied design

principles to their designs. P9 articulated how annotations solidified his understanding of principles by providing specific visual examples: "*I really think it was (helpful for learning) because reading the design principles, I had an idea for it, and then the visuals, it gave me another example, and the pictures really (show) what it should look like, which added onto the readings which set it in place, OK this is what I should do.*" Other solution-centered participants echoed this point. P3 initially found it difficult to address design principles, but *VizCrit* guided her to apply them. She explained, "*I needed it to point out where the issues were ... even with the alignment, it looked good to me. But (VizCrit) pointed out that it wasn't conceptually correct and then I had to change*" This indicates that solution-centered annotations did not hinder learning by virtue of showing example solutions and prompting participants to apply them within their own design task.

Awareness-centered Feedback Encouraged Self-Reflection. Participants with awareness-centered feedback regularly reflected on their designs based on annotations and made iterations. Of all feedback viewed, awareness-centered participants actively reflected on 34% of 149 feedback and solution-centered participants reflected on 15% of 116 feedback during the think-aloud design task. This aligned with what instructors from the co-design study anticipated: even without exact instruction, students could still understand awareness-centered annotations since they provided enough information to help novices reflect and make design decisions.

Indeed, as awareness-centered participants frequently reflected on the rationale behind their design choices, it led to concrete actions for iterations or reinforced their current decisions. For example, after checking the unity feedback, P22 noticed that she used two different font colors. She then justified that this aligned with her intention since the title should be emphasized with a different color. In another case, when she noticed that two texts had different priority levels from the hierarchy annotation, she updated the font sizes immediately to give title more visual prominence. P22 reflected, "*It seems like this one is high, this one's medium, but this one should be in the high hierarchy because it is the title of the whole design. So that's why in order to make that the title, I'm going to make the font bigger, maybe 60, super big.*"

Participants' reflection also extended beyond the highlighted text elements and led to creative exploration for other design elements, such as images and graphics. From the whitespace annotation, many participants noticed the large empty space on the canvas, which inspired them to add extra elements for a more appealing and balanced design. As P15 mentioned, "*there's still a lot of white space on the bottom, so maybe I can potentially add an icon somewhere that I can find in the graphics.*" However, none of the solution-centered participants realized such empty space since they were focused on the identified issues and overlooked other visual cues and information communicated by the annotation.

7 Discussion

We co-designed with experts a set of visual annotations across the feedback actionability spectrum and embedded them into *VizCrit*, an interactive visual design tool supported by algorithmic issue detection and annotation generation. With *VizCrit*, we studied the impact of feedback actionability with design novices and found that solution-centered feedback improved design quality but may have inflated novices' self-perceived creativity. Some novices requested even more actionability than solution-centered feedback, suggesting a preference for performance over learning. We review the results and contextualize the limitations and potential future work to extend the exploration around feedback actionability. We then discuss how to calibrate feedback actionability to balance productivity and creative growth, roles of AI in Human-AI co-creativity, and end with broader ethical considerations of using AI in creativity support tools.

7.1 Limitations and Future Work

We contextualize our study on feedback actionability for design novices by discussing its limitations and future work.

7.1.1 Impacts of Design Expertise on Desire for Actionability. Our study was conducted with novice designers who are university undergraduate students with little to no prior design experience. Therefore, our findings on how different levels of feedback actionability affect participants' learning, creativity, and design outcomes are situated within this specific novice population. We believe that it is important to conduct further studies with other populations (e.g., novices with an interest in design, intermediate designers with some art training, novices with less familiarity with AI tools) to discover additional insights. For instance, novice designers' self-interest in learning visual design could create a difference in their perceptions of feedback. Future studies could target those who are more motivated to develop design skills in settings like design workshops. Additionally, current college students extensively use AI tools to "get answers" in various learning contexts [2, 4, 5, 94], which might be the reason for prioritizing productivity over reflective thinking. Future work could further investigate populations who use AI tools less frequently or have less efficiency-oriented motivations for using AI (e.g., getting explanations on learning materials). Other than design novices, more experienced designers who have more design knowledge could perceive and engage differently with awareness- and solution-centered feedback. Future studies should explore how the level of expertise shapes the interaction strategies and interpretations of critiques [93] and whether awareness-centered design annotations better align with professional workflows with more creative opportunities and freedom.

7.1.2 Task Influence on User Motivation. During the study, we provided participants with an initial seed design with design issues. Although the current setup exposed them to a full set of design annotations and set a starting stage for novices, it could limit their sense of creative ownership. The seed design might also bias outcomes as solution-centered feedback could appear more effective when issues were obvious and inevitable. First, future studies could vary the design task by asking participants to start from scratch (ideally on a design task that they are self-motivated to complete) and investigate how novices interact with awareness-centered feedback and *emerging* solution-centered feedback that identifies issues arising from their own design choices. Additionally, researchers should explore an alternative study setup that includes a comprehensive *VizCrit* version where both awareness- and solution-centered annotations are available. This could allow researchers to observe

whether design phases, individual preferences, or design habits affect their choices regarding feedback actionability.

7.1.3 Longer Term Studies of Learning. We did not find statistically significant differences in the learning tests across conditions. Our qualitative results suggest that solution-centered feedback did not hinder learning by the virtue of offering novices concrete experiences of solving design issues [26, 47, 72] and worked examples [19] of designs with no issues. However, we recognize that learning was hard to measure within a short timeframe, since newly acquired knowledge might not be immediately captured by the evaluation [13] and we anticipate the benefits of awareness-centered annotations to occur over a longer period of time. This is because participants' regular self-reflection could support deeper understanding of design principles and awareness-centered design annotations left more space for learning by iterating and exploring different design possibilities [56, 61]. Longitudinal studies could therefore better capture the learning trajectories and long-term influence of feedback actionability.

7.1.4 Moving Beyond Design Principles. VizCrit was primarily designed to visualize explicit knowledge of design principles. However, these principles could be further extended to other design concepts (e.g., color contrast) and even tacit design knowledge [78]. For example, future systems could incorporate design experts' gaze patterns as an overlay or animation. For awareness-centered annotations, experts' attention patterns or gaze on focal design elements could be highlighted on the canvas, and for solution-centered annotations, the experts' potential next actions could be further visualized.

7.2 Calibrating Feedback Actionability to Balance Productivity and Creative Growth

Our results showcased evidence of a range of the influence scaffolding can have on novices' design processes—awareness-centered feedback encouraged participants to actively reflect on their design choices, while solution-centered feedback led to better design quality. On one hand, the more actionable scaffolding supported novices by making them feel empowered to consider principles and to explore [52, 62]. On the other hand, we saw how it structured and limited the breadth of their exploration [55, 70]. While more actionability may support performance and productivity, it has the greatest risk of overreliance, removing the need for deeper self-reflection and learning [15, 26, 46]. This can be especially harmful if novices prioritize a productivity-focused mindset valuing efficiency over reflective thinking. Feedback with less actionability, on the other hand, requires designers to engage in self-reflection [52, 62] and experimentation with their own design language rather than being steered prematurely toward normative solutions [55, 70].

Therefore, we suggest design opportunities for more dynamic computational design tools by calibrating feedback actionability along several dimensions, including interaction timing, user goals, and experience levels. In terms of timing, designers might consider creating tools to progressively increase actionability throughout a design process. Early on, novices can use less actionable feedback to support ongoing reflection and exploration of the design space, and later, use more actionable feedback to resolve any remaining issues and make refinements. This potentially requires future AI

tools to actively identify design phases. To support diverse user goals (e.g., doing a task for the creative experience or just for a quick outcome), users should be able to tailor actionability [17, 32, 70]. For example, when they are under time constraints [16, 81] and desire efficiency, most actionable feedback can support them in boosting productivity (potentially accompanied by multiple outcomes in parallel for designers to choose from [27]). Additionally, users with more experience or who want to prioritize their learning or agency may prefer the flexibility of less actionable feedback, while novices may need more actionability when decision-making still feels intimidating.

7.3 Understanding the Roles of AI in Human-AI Co-Creativity

By computationally inspecting the design, *VizCrit* not only acts as a collaborator to deliver constructive feedback, but also as an audience to simulate how others perceive the design [9, 43]. Both awareness- and solution-centered participants described designing "in others' eyes" and become more aware of how a professional, artistic reviewer may view and evaluate their draft [33, 39], also aligning with prior insights on a tendency to conform too closely to the audience perspectives [43]. Regardless, this can be seen as a form of external validation to users, especially for novices or inexperienced creators, who typically lack confidence in their creative outcomes [82] or lack access to timely feedback [24]. This may help encourage novices to adventure a new domain as social apprehension is reduced [48, 49] and experiment with more creative possibilities and seek new feedback after each attempt [44].

Beyond individual creative experience, AI tools can also be a social mediator that facilitates collaboration while having a cordial atmosphere [31, 79]. As creativity is highly subjective and personal [77], with AI's objective analysis, collaborators can have a shared reference for discussion and psychologically safer spaces for negotiation and creative alignment. AI tools can further enable collaborators to build and customize creative and learning experiences [54]. For example, researchers can investigate how design instructors and professionals adopt or customize *VizCrit*'s feedback in teaching contexts and what additional scaffolds may be needed to support their uses.

7.4 Navigating Ethical Considerations of AI in Creativity Support Tools

Our results highlighted ethical considerations of AI in CSTs regarding its power and the normativeness of scaffolding based on the well-known guidelines, especially given novices' tendency to feel the compliance pressure from rule-based feedback, and their tendency to overly rely and trust the tool.

7.4.1 Compliance Pressure from Rules-based Feedback. Novices often perceive design principles as strict rules to follow and memorize instead of flexible guidelines to adapt or break [1]. As novices tended to lack confidence in breaking preset rules and inappropriately relied on "AI" solutions [58], they could easily interpret solution-centered feedback as perfect answers that delineated the correct and problematic attempts. As a result, novices might be compelled to make changes to satisfy the system, giving it more

power over their creative decision-making process [55]. In contrast, awareness-centered feedback allowed more exploration within the constraints of design principles. Without knowing a system-side definite answer, novices had spaces to reflect, justify, and reinforce their design choices, even though they might be "wrong" based on principles. Therefore, to better maintain creative agency, CSTs with AI could aim to empower novices to decide when and how to follow feedback and provide breathing room to question computational analysis [27, 83].

7.4.2 Overreliance on AI and False Sense of Creativity. The contrasting result from solution-centered participants' self-perceived and expert-rated creativity uncovered their false sense of creativity and unconscious overreliance on the system. Novices can easily overly trust AI [8, 15] to evaluate and revise their designs, which not only risks agency, but leads to even worse outcomes if the system provides feedback that follows principles (algorithmically correct), but is not applicable in a personal creative context. Prior work has shown that explanations can help reduce overreliance and facilitate creative outcomes [3, 51, 85]. However, in our case we thought the annotations would serve as easier-to-understand *explanations* (than traditional text feedback), but it instead might have further created a false sense of trust. This implies that AI should encourage more critical self-reflection to avoid the possibility of being a substitution of creative decision-making. Computational tools could not only help novices with efficiency gains, but also skill building and creative growth for better retaining user agency and avoiding overreliance.

8 Conclusion

As AI systems make instant, on-demand design feedback widely available, it is essential to understand how feedback actionability shapes creative work. Informed by co-design interviews with nine experts, we introduce *VizCrit*, a prototype visual design tool that automatically detects design issues and generates multi-modal feedback, including visual annotations, across the actionability spectrum: textbook-based, awareness-centered, and solution-centered. In a between-subjects study with 36 novices, we discover that solution-centered feedback supported design quality, did not hinder learning, and increased self-perceived creativity. However, this self-perception of creativity may be inflated by virtue of feedback-driven iterations and sense of progress. These results argue that less actionable, awareness-centered feedback encourages self-reflection, while more actionable, solution-centered feedback supports performance and productivity but may risk overreliance. For designers of AI-powered creativity support tools, we recommend calibrating feedback actionability for design novices considering the interaction timing and users' goals. In doing so, systems can pair immediate assistance with long-term growth in judgment, reflection, and creative exploration.

Acknowledgments

References

[1] Andrew Abbott. 2014. *The system of professions: An essay on the division of expert labor.* University of Chicago press.

[2] Rudaiba Adnin, Atharva Pandkar, Bingsheng Yao, Dakuo Wang, and Maitraye Das. 2025. Examining Student and Teacher Perspectives on Undisclosed Use of Generative AI in Academic Work. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25).* Association for Computing Machinery, New York, NY, USA, Article 1071, 17 pages. doi:10.1145/3706598.3713393

[3] Faez Ahmed, Nischal Reddy Chandra, Mark Fuge, and Steven Dow. 2019. Structuring Online Dyads: Explanations Improve Creativity, Chats Lead to Convergence. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19).* Association for Computing Machinery, New York, NY, USA, 306–318. doi:10.1145/3325480.3325486

[4] Matin Amoozadeh, David Daniels, Daye Nam, Aayush Kumar, Stella Chen, Michael Hilton, Sruti Srinivasa Ragavan, and Mohammad Amin Alipour. 2024. Trust in Generative AI among Students: An exploratory study. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1* (Portland, OR, USA) *(SIGCSE 2024).* Association for Computing Machinery, New York, NY, USA, 67–73. doi:10.1145/3626252.3630842

[5] Alex Barrett and Austin Pack. 2023. Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 59.

[6] Terry Barrett. 1988. A comparison of the goals of studio professors conducting critiques and art education goals for teaching criticism. *Studies in Art Education* 30, 1 (1988), 22–27. doi:10.2307/1320648

[7] Matthew T Beaudouin-Lafon, Jane L E, and Haijun Xia. 2023. Color Field: Developing Professional Vision by Visualizing the Effects of Color Filters. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23).* Association for Computing Machinery, New York, NY, USA, Article 101, 16 pages. doi:10.1145/3586183.3606828

[8] Natalie C Benda, Laurie L Novak, Carrie Reale, and Jessica S Ancker. 2021. Trust in AI: why we should be designing for APPROPRIATE reliance. *Journal of the American Medical Informatics Association* 29, 1 (2021), 207–212. doi:10.1093/jamia/ocab238

[9] Karim Benharrak, Tim Zindulka, Florian Lehmann, Hendrik Heuer, and Daniel Buschek. 2024. Writer-Defined AI Personas for On-Demand Feedback Generation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24).* Association for Computing Machinery, New York, NY, USA, Article 1049, 18 pages. doi:10.1145/3613904.3642406

[10] Cynthia L Bennett, Renee Shelby, Negar Rostamzadeh, and Shaun K Kane. 2024. Painting with Cameras and Drawing with Text: AI Use in Accessible Creativity. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24).* Association for Computing Machinery, New York, NY, USA, Article 5, 19 pages. doi:10.1145/3663548.3675644

[11] Oloff C. Biermann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Proceedings of the 2022 Designing Interactive Systems Conference (DIS '22).* Association for Computing Machinery, New York, NY, USA, 1209–1227. doi:10.1145/3532106.3533506

[12] Paul Black and Dylan Wiliam. 1998. Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice* 5, 1 (1998), 7–74. doi:10.1080/0969595980050102

[13] Benjamin S. Bloom. 1956. *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain.* Longmans, Green and Company.

[14] Virginia Braun and Victoria Clarke. 2021. *Thematic Analysis: A Practical Guide.* Sage Publications.

[15] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 188 (April 2021), 21 pages. doi:10.1145/3449287

[16] Shiye Cao, Catalina Gomez, and Chien-Ming Huang. 2023. How Time Pressure in Different Phases of Decision-Making Influences Human-AI Collaboration. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 277 (Oct. 2023), 26 pages. doi:10.1145/3610068

[17] Liwei Chan, Yi-Chi Liao, George B Mo, John J Dudley, Chun-Lien Cheng, Per Ola Kristensson, and Antti Oulasvirta. 2022. Investigating Positive and Negative Qualities of Human-in-the-Loop Optimization for Designing Interaction Techniques. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22).* Association for Computing Machinery, New York, NY, USA, Article 112, 14 pages. doi:10.1145/3491102.3501850

[18] Chen Chen, Cuong Nguyen, Thibault Groueix, Vladimir G. Kim, and Nadir Weibel. 2024. MemoVis: A GenAI-Powered Tool for Creating Companion Reference Images for 3D Design Feedback. *ACM Trans. Comput.-Hum. Interact.* 31, 5, Article 67 (Nov. 2024), 41 pages. doi:10.1145/3694681

[19] Ouhao Chen, Endah Retnowati, BoBo Kai Yin Chan, and Slava Kalyuga. 2023. The effect of worked examples on learning solution steps and knowledge transfer. *Educational Psychology* 43, 8 (2023), 914–928. doi:10.1080/01443410.2023.2273762

[20] Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Trans. Comput.-Hum. Interact.* 21, 4, Article 21 (June 2014), 25 pages. doi:10.1145/2617588

[21] Sasha Costanza-Chock. 2020. *Design Justice: Community-Led Practices to Build the Worlds We Need.* MIT Press. doi:10.7551/mitpress/12255.001.0001

[22] Maria Cutumisu and Daniel L. Schwartz. 2018. The impact of critical feedback choice on students' revision, performance, learning, and memory. *Computers in Human Behavior* 78 (2018), 351–367. doi:10.1016/j.chb.2017.06.029

[23] Deanna Dannels, Amy Housley Gaffney, and Kelly Norris Martin. 2008. Beyond content, deeper than delivery: What critique feedback reveals about communication expectations in design education. *International Journal for the Scholarship of teaching and Learning* 2, 2 (2008), 12. https://digitalcommons.georgiasouthern.edu/ij-sotl/vol2/iss2/12/

[24] Deanna P Dannels and Kelly Norris Martin. 2008. Critiquing critiques: A genre analysis of feedback across novice to expert design studios. *Journal of Business and Technical Communication* 22, 2 (2008), 135–159. https://doi.org/10.1177/1050651907311923

[25] Nicholas Davis, Chih-PIn Hsiao, Kunwar Yashraj Singh, Lisa Li, and Brian Magerko. 2016. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) *(IUI '16)*. Association for Computing Machinery, New York, NY, USA, 196–207. doi:10.1145/2856767.2856795

[26] John Dewey. 1938. *Experience and Education.* Macmillan Company, New York, NY, USA.

[27] Steven P. Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L. Schwartz, and Scott R. Klemmer. 2011. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Trans. Comput.-Hum. Interact.* 17, 4, Article 18 (Dec. 2011), 24 pages. doi:10.1145/1879831.1879836

[28] Peitong Duan, Chin-Yi Cheng, Gang Li, Bjoern Hartmann, and Yang Li. 2024. UICrit: Enhancing Automated Design Evaluation with a UI Critique Dataset. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 46, 17 pages. doi:10.1145/3654777.3676381

[29] Peitong Duan, Jeremy Warner, Yang Li, and Bjoern Hartmann. 2024. Generating Automatic Feedback on UI Mockups with Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 6, 20 pages. doi:10.1145/3613904.3642782

[30] Jane L. E, Ohad Fried, and Maneesh Agrawala. 2019. Optimizing Portrait Lighting at Capture-Time Using a 360 Camera as a Light Probe. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 221–232. doi:10.1145/3332165.3347893

[31] Jane L. E, Ohad Fried, Jingwan Lu, Jianming Zhang, Radomír Mech, Jose Echevarria, Pat Hanrahan, and James A. Landay. 2020. Adaptive Photographic Composition Guidance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376635

[32] Jane L. E, Yu-Chun Grace Yen, Isabelle Yan Pan, Grace Lin, Mingyi Li, Hyoungwook Jin, Mengyi Chen, Haijun Xia, and Steven P. Dow. 2024. When to Give Feedback: Exploring Tradeoffs in the Timing of Design Feedback. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 292–310. doi:10.1145/3635636.3656183

[33] David B Elliott and Amanda Skaff. 1993. Vision of the famous: the artist's eye. *Ophthalmic and Physiological Optics* 13, 1 (1993), 82–90.

[34] K Anders Ericsson, Ralf T Krampe, and Clemens Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological review* 100, 3 (1993), 363. doi:10.1037/0033-295X.100.3.363

[35] Gerhard Fischer. 1994. Turning breakdowns into opportunities for creativity. *Knowledge-Based Systems* 7, 4 (1994), 221–232.

[36] Gerhard Fischer, Andreas C Lemke, Thomas Mastaglio, and Anders I Morch. 1991. Critics: An emerging approach to knowledge-based human-computer interaction. *International Journal of Man-Machine Studies* 35, 5 (1991), 695–721.

[37] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 337, 18 pages. doi:10.1145/3290605.3300619

[38] Katy Ilonka Gero, Vivian Liu, and Lydia B. Chilton. 2022. Sparks: Inspiration for Science Writing Using Language Models. In *Proceedings of the 2022 Designing Interactive Systems Conference (DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1002–1019. doi:10.1145/3532106.3533533

[39] Charles Goodwin. 1994. Professional vision. *American Anthropologist* 96, 3 (1994), 606–633. doi:10.1525/aa.1994.96.3.02a00100

[40] Colin M. Gray. 2013. Informal peer critique and the negotiation of habitus in a design studio. *Art, Design & Communication in Higher Education* 12, 2 (2013), 195–209. doi:10.1386/adch.12.2.195_1

[41] John Hattie and Helen Timperley. 2007. The power of feedback. *Review of educational research* 77, 1 (2007), 81–112.

[42] Mina Huh and Amy Pavel. 2024. DesignChecker: Visual Design Support for Blind and Low Vision Web Developers. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 142, 19 pages. doi:10.1145/3654777.3676369

[43] Yu-Kai Hung, Yun-Chien Huang, Ting-Yu Su, Yen-Ting Lin, Lung-Pan Cheng, Bryan Wang, and Shao-Hua Sun. 2025. SimTube: Simulating Audience Feedback on Videos using Generative AI and User Personas. In *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*. Association for Computing Machinery, New York, NY, USA, 1256–1271. doi:10.1145/3708359.3712146

[44] Nujaree Intasao and Ning Hao. 2018. Beliefs about creativity influence creative performance: The mediation effects of flexibility and positive affect. *Frontiers in psychology* 9 (2018), 411756.

[45] David G Jansson and Steven M Smith. 1991. Design fixation. *Design studies* 12, 1 (1991), 3–11.

[46] Charlotte Kobiella, Yarhy Said Flores López, Franz Waltenberger, Fiona Draxler, and Albrecht Schmidt. 2024. "If the Machine Is As Good As Me, Then What Use Am I?" – How the Use of ChatGPT Changes Young Professionals' Perception of Productivity and Accomplishment. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1018, 16 pages. doi:10.1145/3613904.3641964

[47] David Kolb. 1984. *Experiential Learning: Experience As The Source Of Learning And Development.* Prentice-Hall.

[48] Yasmine Kotturi and McKayla Kingston. 2019. Why do Designers in the "Wild" Wait to Seek Feedback until Later in their Design Process?. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 541–546. doi:10.1145/3325480.3326580

[49] Sneha R. Krishna Kumaran, Wenxuan Wendy Shi, and Brian P Bailey. 2021. Am I Ready to Get Feedback? A Taxonomy of Factors Creators Consider Before Seeking Feedback on In-Progress Creative Work. In *Proceedings of the 13th Conference on Creativity and Cognition* (Virtual Event, Italy) *(C&C '21)*. Association for Computing Machinery, New York, NY, USA, Article 31, 10 pages. doi:10.1145/3450741.3465255

[50] William H. Kruskal and W. Allen Wallis. 1952. Use of Ranks in One-Criterion Variance Analysis. *J. Amer. Statist. Assoc.* 47, 260 (1952), 583–621. doi:10.1080/01621459.1952.10483441

[51] Todd Kulesza, Simone Stumpf, Margaret Burnett, Sherry Yang, Irwin Kwan, and Weng-Keen Wong. 2013. Too much, too little, or just right? Ways explanations impact end users' mental models. In *2013 IEEE Symposium on Visual Languages and Human Centric Computing.* 3–10. doi:10.1109/VLHCC.2013.6645235

[52] Chinmay Kulkarni, Steven P. Dow, and Scott R. Klemmer. 2014. *Early and Repeated Exposure to Examples Improves Creative Work.* Springer International Publishing, Cham, 49–62. doi:10.1007/978-3-319-01303-9_4

[53] Chunggi Lee, Sanghoon Kim, Dongyun Han, Hongjun Yang, Young-Woo Park, Bum Chul Kwon, and Sungahn Ko. 2020. GUIComp: A GUI Design Assistant with Real-Time, Multi-Faceted Feedback. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376327

[54] Jingyi Li, Sonia Hashim, and Jennifer Jacobs. 2021. What We Can Learn From Visual Artists About Software Development. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 314, 14 pages. doi:10.1145/3411764.3445682

[55] Jingyi Li, Eric Rawn, Jacob Ritchie, Jasper Tran O'Leary, and Sean Follmer. 2023. Beyond the Artifact: Power as a Lens for Creativity Support Tools. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 47, 15 pages. doi:10.1145/3586183.3606831

[56] Louis Lousberg, Remon Rooij, Sylvia Jansen, Elise van Dooren, John Heintz, and Engbert van der Zaag. 2020. Reflection in design education. *International Journal of Technology and Design Education* 30, 5 (2020), 885–897.

[57] Jiaju Ma, Chau Vu, Asya Lyubavina, Catherine Liu, and Jingyi Li. 2025. Computational Scaffolding of Composition, Value, and Color for Disciplined Drawing.

In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. Association for Computing Machinery, New York, NY, USA, Article 161, 15 pages. doi:10.1145/3746059.3747605

[58] Shuai Ma, Xinru Wang, Ying Lei, Chuhan Shi, Ming Yin, and Xiaojuan Ma. 2024. "Are You Really Sure?" Understanding the Effects of Human Self-Confidence Calibration in AI-Assisted Decision Making. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 840, 20 pages. doi:10.1145/3613904.3642671

[59] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.

[60] Maureen Stone. 2016. How we designed the new color palettes in Tableau 10. https://www.tableau.com/blog/colors-upgrade-tableau-10-56782

[61] Jason McDonald and Esther Michela. 2022. "This Uncertain Space of Teaching:" How Design Studio Instructors Talk About Design Critiques Along with Themselves when Giving Critiques. *Journal of the Scholarship of Teaching and Learning* 22, 1 (Mar. 2022), 48–66. doi:10.14434/josotl.v22i1.30888

[62] Michael Mose Biskjaer and Kim Halskov. 2014. Decisive constraints as a creative resource in interaction design. *Digital Creativity* 25, 1 (2014), 27–61.

[63] Melissa M Nelson and Christian D Schunn. 2009. The nature of feedback: How different types of peer feedback affect writing performance. *Instructional science* 37, 4 (2009), 375–401.

[64] Tricia J. Ngoon, Joy O Kim, and Scott Klemmer. 2021. Shöwn: Adaptive Conceptual Guidance Aids Example Use in Creative Tasks. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1834–1845. doi:10.1145/3461778.3462072

[65] Thi Thao Duyen T. Nguyen, Thomas Garncarz, Felicia Ng, Laura A. Dabbish, and Steven P. Dow. 2017. Fruitful Feedback: Positive Affective Language and Source Anonymity Improve Critique Reception and Work Outcomes. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1024–1034. doi:10.1145/2998181.2998319

[66] David J Nicol and Debra Macfarlane-Dick. 2006. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education* 31, 2 (2006), 199–218. doi:10.1080/03075070600572090

[67] Peter O'Donovan, Aseem Agarwala, and Aaron Hertzmann. 2015. DesignScape: Design with Interactive Layout Suggestions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1221–1224. doi:10.1145/2702123.2702149

[68] Changhoon Oh, Jungwoo Song, Jinhan Choi, Seonghyeon Kim, Sungwoo Lee, and Bongwon Suh. 2018. I Lead, You Help but Only with Enough Details: Understanding User Experience of Co-Creation with Artificial Intelligence. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3174223

[69] A Terry Purcell and John S Gero. 1996. Design and other types of fixation. *Design studies* 17, 4 (1996), 363–383.

[70] Samuel Rhys Cox, Helena Bøjer Djernæs, and Niels van Berkel. 2025. Beyond Productivity: Rethinking the Impact of Creativity Support Tools. In *Proceedings of the 2025 Conference on Creativity and Cognition (C&C '25)*. Association for Computing Machinery, New York, NY, USA, 735–749. doi:10.1145/3698061.3726924

[71] D Royce Sadler. 1989. Formative assessment and the design of instructional systems. *Instructional science* 18, 2 (1989), 119–144.

[72] Donald A. Schön. 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

[73] Daniel L Schwartz, Catherine C Chase, Marily A Oppezzo, and Doris B Chin. 2011. Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of educational psychology* 103, 4 (2011), 759. doi:10.1037/a0025140

[74] Daniel L. Schwartz and Taylor Martin. 2004. Inventing to Prepare for Future Learning: The Hidden Efficiency of Encouraging Original Student Production in Statistics Instruction. *Cognition and Instruction* 22, 2 (2004), 129–184. doi:10.1207/s1532690xci2202_1

[75] Sha Shen, Saidi Wang, Yun Qi, Yanli Wang, and Xiangdong Yan. 2021. Teacher suggestion feedback facilitates creativity of students in STEAM education. *Frontiers in Psychology* 12 (2021), 723171. doi:10.3389/fpsyg.2021.723171

[76] Valerie J Shute. 2008. Focus on formative feedback. *Review of educational research* 78, 1 (2008), 153–189.

[77] Dean Keith Simonton. 2000. Creativity: Cognitive, personal, developmental, and social aspects. *American psychologist* 55, 1 (2000), 151. doi:10.1037/0003-066X.55.1.151

[78] Kihoon Son, DaEun Choi, Tae Soo Kim, and Juho Kim. 2024. Demystifying Tacit Knowledge in Graphic Design: Characteristics, Instances, Approaches, and Guidelines. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing

Machinery, New York, NY, USA, Article 221, 18 pages. doi:10.1145/3613904.3642886

[79] Minhyang (Mia) Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 582, 11 pages. doi:10.1145/3411764.3445219

[80] Masaki Suwa and Barbara Tversky. 2002. External Representations Contribute to the Dynamic Construction of Ideas. In *Diagrammatic Representation and Inference*, Mary Hegarty, Bernd Meyer, and N. Hari Narayanan (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 341–343.

[81] Siddharth Swaroop, Zana Buçinca, Krzysztof Z. Gajos, and Finale Doshi-Velez. 2024. Accuracy-Time Tradeoffs in AI-Assisted Decision Making under Time Pressure. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (Greenville, SC, USA) *(IUI '24)*. Association for Computing Machinery, New York, NY, USA, 138–154. doi:10.1145/3640543.3645206

[82] Kesler Tanner and James Landay. 2018. "I Know It When I See It": How Experts and Novices Recognize Good Design. Springer, 249–266. doi:10.1007/978-3-319-97082-0_13

[83] Michael Terry, Elizabeth D Mynatt, Kumiyo Nakakoji, and Yasuhiro Yamamoto. 2004. Variation in element and action: supporting simultaneous development of alternative solutions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 711–718. doi:10.1145/985692.985782

[84] Barbara Tversky. 2002. What do sketches say about thinking. In *2002 AAAI Spring Symposium, Sketch Understanding Workshop, Stanford University, AAAI Technical Report SS-02-08*. Vol. 148. 151.

[85] Helena Vasconcelos, Matthew Jörke, Madeleine Grunde-McLaughlin, Tobias Gerstenberg, Michael S. Bernstein, and Ranjay Krishna. 2023. Explanations Can Reduce Overreliance on AI Systems During Decision-Making. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 129 (April 2023), 38 pages. doi:10.1145/3579605

[86] WebAIM. 2026. Contrast Checker. https://webaim.org/resources/contrastchecker/

[87] Andrew Williams. 2024. Delivering Effective Student Feedback in Higher Education: An Evaluation of the Challenges and Best Practice. *International Journal of Research in Education and Science* 10, 2 (May 2024), 473–501. doi:10.46328/ijres.3404

[88] Langdon Winner. 2017. Do artifacts have politics? In *Computer ethics*. Routledge, 177–192.

[89] Naomi E. Winstone, Robert A. Nash, James Rowntree, and Michael Parker. 2017. 'It'd be useful, but I wouldn't use it': barriers to university students' feedback seeking and recipience. *Studies in Higher Education* 42, 11 (2017), 2026–2041. doi:10.1080/03075079.2015.1130032

[90] Jason Wu, Yi-Hao Peng, Xin Yue Amanda Li, Amanda Swearngin, Jeffrey P Bigham, and Jeffrey Nichols. 2024. UIClip: A Data-driven Model for Assessing User Interface Design. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 45, 16 pages. doi:10.1145/3654777.3676408

[91] Yong Wu and Christian D. Schunn. 2020. From feedback to revisions: Effects of feedback features and perceptions. *Contemporary Educational Psychology* 60 (2020), 101826. doi:10.1016/j.cedpsych.2019.101826

[92] Xiaotong (Tone) Xu, Arina Konnova, Bianca Gao, Cindy Peng, Dave Vo, and Steven P. Dow. 2025. Productive vs. Reflective: How Different Ways of Integrating AI into Design Workflows Affect Cognition and Motivation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 24, 15 pages. doi:10.1145/3706598.3713649

[93] Alvin Yuan, Kurt Luther, Markus Krause, Sophie Isabel Vennix, Steven P Dow, and Bjorn Hartmann. 2016. Almost an Expert: The Effects of Rubrics and Expertise on Perceived Value of Crowdsourced Design Critiques. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1005–1017. doi:10.1145/2818048.2819953

[94] Cynthia Zastudil, Magdalena Rogalska, Christine Kapp, Jennifer Vaughn, and Stephen MacNeil. 2023. Generative ai in computing education: Perspectives of students and instructors. In *2023 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[95] Zhuohao (Jerry) Zhang, Ruiqi Chen, Mingyuan Zhong, and Jacob O. Wobbrock. 2025. SlideAudit: A Dataset and Taxonomy for Automated Evaluation of Presentation Slides. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. Association for Computing Machinery, New York, NY, USA, Article 13, 23 pages. doi:10.1145/3746059.3747736

[96] Chanjin Zheng, Zengyi Yu, Yilin Jiang, Mingzi Zhang, Xunuo Lu, Jing Jin, and Liteng Gao. 2025. ArtMentor: AI-Assisted Evaluation of Artworks to Explore Multimodal Large Language Models Capabilities. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for

A Appendix

A.1 Design Probe Co-Design

This was an example of design probe annotations for the whitespace principle. Co-design participants iterated on their preferred design for both awareness- and solution-centered annotations (Figure 10).

A.2 Annotation Explanations

Each annotation had corresponding text explanations in table form. These are the range of possible explanations in the system across awareness-centered and solution-centered annotations (Table 2).

A.3 Principle Details

These were the descriptions that could be surfaced in the interface when participants clicked on "show" to see more about the principle or how to apply the principle (Table 3).

A.4 System Implementation Details

We describe the implementation of how the server communicates with the user interface. The frontend interface is built on top of Polotno[9], an open-sourced web-based visual design tool. The server is a Python backend, using Pillow[10] to generate annotation images. A websocket server facilitates communication between them. The interface needs to provide updated data on the user's design to the server, and the server provides annotations and associated information back to the interface to surface to the user. Note that the tool runs locally in its current form.

Retrieving the Current On-Canvas Design. To analyze a user design in real-time, the interface monitors on-canvas design activities. Whenever there is a change on the canvas (with 4-sec debounce time to avoid over-requests), the interface sends the current design data to the server, including a canvas JSON object, full design image, and text-only image (for more close analysis of space occupied by text). Images were generated using Polotno Cloud Rendering API[11]. Algorithms primarily used the JSON description of the canvas, with a few requiring some level of image understanding directly from the full design or text-only images.

Generating Annotation-related Data. Once the server receives the design data, it runs algorithms to generate annotations across all principles. Updated annotations are saved in a location that is accessible by the interface and all other data is sent back via JSON to be used to populate the interface.

---

[9]https://polotno.com/
[10]https://pillow.readthedocs.io/en/stable/reference/ImageDraw.html
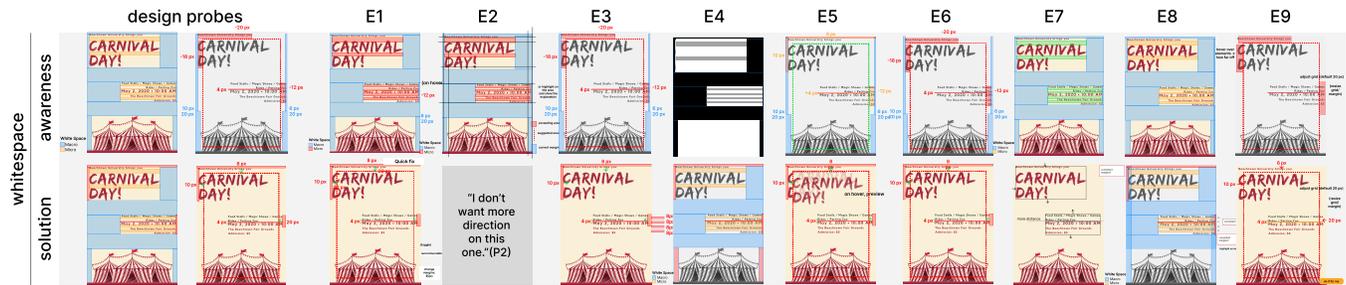[11]https://polotno.com/cloud-render

Figure 10: Iterative design of the whitespace design probes: our set of 4 annotations at the left for whitespace (2 awareness-centered top, 2 solution-centered bottom), along with the expert iterations of each. Most started with their preferred of our two designs and made revisions from there.

Table 2: Explanations for visual annotations: annotations on the canvas will have explanations in the text format on the design principles panel. The content in the square bracket indicates the specific fields that will appear in the text explanation table.

Principle	Awareness-centered Explanations	Solution-centered Explanations
Hierarchy	[Text element] in the design has [level of emphasis] visual emphasis.	The [title] does not seem particularly emphasized compared to [competing text element(s)], making it difficult to recognize what is most important. The [title] does not seem particularly emphasized.
Alignment	Text in the design forms x alignment groups with [text element(s)]. [Text element] is [internal alignment]-aligned.	Text in the design forms many separate alignment groups with [text element], which can make the design appear somewhat incohesive. [Text element(s)]'s relative positioning on the canvas skews [external alignment], which does not match with the text being [internal alignment]-aligned.
Whitespace	[Text element] is on the [edge] side of the canvas. Note other nearby elements and the lengths of the lines of text. [Text element] and its shape is on the [edge] side of the canvas with other nearby elements.	The line breaks in [text element] creates uneven text lengths, which can somewhat disrupt the visual flow of the text. [Text element] placed quite close to the [edge], which can make the design appear somewhat crowded. [Text elements] placed quite close to each other, which can make the design appear somewhat crowded.
Unity	[Text element] in the design has the following text properties: [font family], [style], [size], [color].	[Text elements] in the design use many different text properties: [font family], [style], [size], [color], which can make your design seem incohesive.

Table 3: Principle details: the full principle information after expanding "Learn More About" on the design principles panel.

Principle	Learn More About Principle	Learn More About Applying This Principle
Hierarchy	**Principle Definition:** Hierarchy is the visual level of emphasis of each element in your design—this should match the importance of the content. A common issue is having a weak point of entry for the viewer. **Common Issue:** The point of entry is the focal (most eye-catching) point of your design. If the point of entry is weak, it is unclear what element is the most important and viewers may not recognize the key message at first sight.	**Example:** On a website, a large, bold headline at the top of the page establishes that it's the most important element, while smaller subheadings and body text are visually less prominent, guiding the reader's eye through the content. **Related Actions:** Increase importance of title - make title text larger - make title different: e.g., bold, color, font, all caps/lowercase, etc. - use more prominent colors/fonts - use additional white space as frame Decrease importance of non-title text - make body text smaller - reduce brightness of body text - use neutral font for body text
Alignment	**Principle Definition:** Alignment refers to the relative positioning of elements in your design. A common issue is for elements to appear to have somewhat arbitrary alignment relative to each other. **Common Issue:** If several elements aren't aligned to common axes or boundaries, the design may appear to lack a sense of order or organization.	**Example:** In a webpage layout, aligning text blocks to the left or centering them on the page creates a structured flow that helps the viewer follow the content easily. **Related Actions:** Create and align common boundaries (external alignment) - align boundaries of nearby text boxes - align to closest margin (e.g. left page margin) - align to common shape (e.g. image) Match internal and external alignment - left/right align text that is positioned on towards the left/right margins - center position text that is center aligned - consider using the same internal alignment (e.g. left-aligned) for related groups of text
Whitespace	**Principle Definition:** Whitespace refers to the spacing between and around elements in your design. A common issue is to have insufficient or ragged spacing relative to other elements or the margins. **Common Issue:** If a block of text has too little margins relative to other elements or the edge of the canvas, or the text lengths are uneven, it can make the design appear somewhat crowded.	**Example:** The empty space around a logo or between paragraphs allows the content to be more digestible, making the design feel less crowded and more inviting. **Related Actions:** Elements should have breathing room/space - give elements space away from boundaries/margins - leave whitespace between elements - consider size of elements when determining how much whitespace should surround it (bigger elements might need more space) Avoid uneven lengths of text within an element - avoid widows, orphans, and singles in your text layout - consider adjusting text box width to make lines more even
Unity	**Principle Definition:** Unity refers to the cohesion in your design. A common issue is to have too many variations in fonts/styles across different elements such as text. **Common Issue:** Using too many variations in font families, sizes, styles, and colors can make a design look incoherent and unorganized.	**Example:** In branded materials, using consistent colors, logo styles, and fonts across all materials (like posters, brochures, and websites) creates a unified look and reinforces the brand identity. **Related Actions:** Limit number of font families, font sizes, font colors, font styles, etc. - reduce total number of fonts/sizes/colors/styles - make similar fonts/sizes/colors/styles the same - make different fonts/sizes/colors/styles more different Use text styles in a consistent manner - treat text of equal importance with same font/size/color/style