ProcessGallery: Contrasting Early and Late Iterations for Design Principle Learning

YU-CHUN GRACE YEN*, National Yang Ming Chiao Tung University, Taiwan
JANE L. E*, University of California, San Diego, USA
HYOUNGWOOK JIN, KAIST, Republic of Korea
MINGYI LI, University of California, San Diego, USA
GRACE LIN, University of California, San Diego, USA
ISABELLE YAN PAN, University of California, San Diego, USA
STEVEN P. DOW, University of California, San Diego, USA

Traditional design galleries enable users to search for examples based on surface attributes (e.g., color or style), and largely obscure underlying principles (e.g., hierarchy or readability). We conducted three studies to explore how galleries could be constructed to help novices learn key design principles. Study 1 revealed that novices gain perspective by observing how designs evolve throughout a process. Study 2 found that novices are better at identifying design issues when viewing iterations that show improvements for just one principle at a time, rather than multiple. Building on these insights, we created ProcessGallery, a tool that enables users to browse contrasting pairs of early-and-late iterations of designs that highlight key improvements organized by design principles. In Study 3, a within-subjects experiment, sixteen participants iterated on a seed design after viewing examples in ProcessGallery versus a traditional gallery. Using ProcessGallery, participants found more appropriate examples, assessed designs better, and preferred ProcessGallery for learning compared to a traditional gallery.

CCS Concepts: • **Human-centered computing** → **Empirical studies in interaction design**; **Interaction design**.

Additional Key Words and Phrases: Visual design, vicarious learning, learning from examples, design gallery, learner-sourcing, personalized learning

ACM Reference Format:
Yu-Chun Grace Yen, Jane L. E, Hyoungwook Jin, Mingyi Li, Grace Lin, Isabelle Yan Pan, and Steven P. Dow. 2024. ProcessGallery: Contrasting Early and Late Iterations for Design Principle Learning. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 112 (April 2024), 36 pages. https://doi.org/10.1145/3637389

1 INTRODUCTION

In creative domains, examples serve as invaluable resources for generating design inspiration, validating work, and learning new design skills and knowledge [51, 68]. With the rise of digital

*Both authors contributed equally to this research.

Authors' addresses: Yu-Chun Grace Yen, yyen@cs.nycu.edu.tw, National Yang Ming Chiao Tung University, Hsinchu, Taiwan; Jane L. E, University of California, San Diego, La Jolla, USA, je@ucsd.edu; Hyoungwook Jin, KAIST, Daejeon, Republic of Korea; Mingyi Li, University of California, San Diego, La Jolla, USA; Grace Lin, University of California, San Diego, La Jolla, USA; Isabelle Yan Pan, University of California, San Diego, La Jolla, USA; Steven P. Dow, University of California, San Diego, La Jolla, USA.

2573-0142/2024/4-ART112 $15.00
https://doi.org/10.1145/3637389

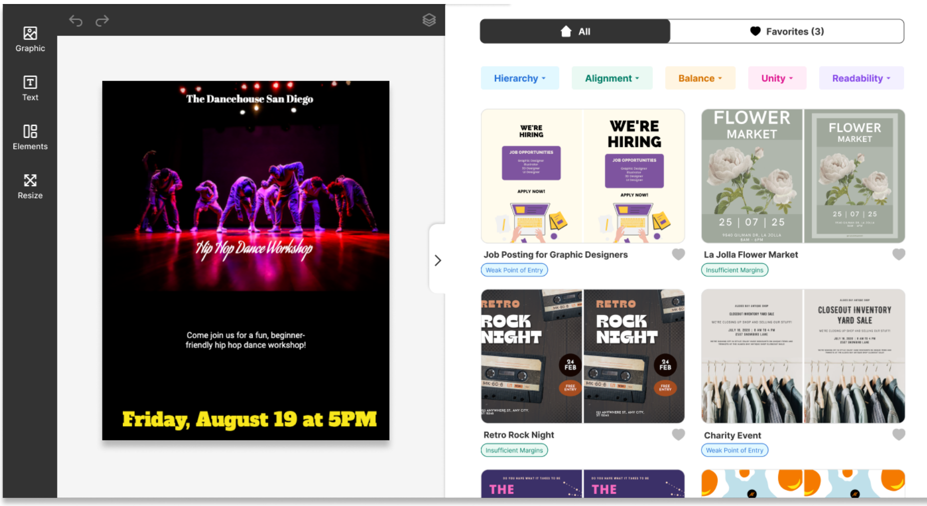


Fig. 1. ProcessGallery is a novel design gallery comprised of contrasting pairs of early and late iterations of poster designs. Each pair showcases potential fixes that resolve violations of a design principle. The poster on the left (Hip Hop Dance Workshop) is the design the current user is working on. Through the gallery, the user can search for examples that appear most applicable to what they intend to fix based on principles.

design tools and content sharing platforms, people increasingly share their own and browse others' work for feedback and inspiration [2, 6, 58]. The HCI research community has developed novel tools to help designers find and extract insights from examples, for instance, suggesting structurally or semantically similar examples to enable analogical innovation [22, 55] or recommending distant or "less popular" examples to boost creativity [24, 84]. With the advancement of generative AI, researchers have demonstrated the ability of machines to produce never-before-seen examples [72, 77]. For instance, DreamSketch captures a designer's definition of a design problem and synthesizes various design solutions that meet the objectives [57]. While the research community has given much attention to generating or finding examples to serve as inspiration, there has been less focus on how people can *learn* from examples.

Effective examples necessitate learners to identify the factors contributing to the success of a design, including the strategies employed [80, 94]. Instructors in educational settings often carefully select or produce examples that illustrate specific issues [94]. However, creating examples that address every possible problem for students to learn would be daunting for a single instructor. Many design communities curate examples to showcase exemplary work. However, novice designers often struggle to understand these examples and feel intimidated by the quality and overwhelmed by the quantity of such online galleries [70, 96]. Following a poor example or applying an example without knowing why it is effective can both be harmful to creative learning [29, 80]. Prior work suggests that learners are more likely to gain insight, stay motivated, and perform better in the future when viewing examples that are only slightly better than their own [21, 78]. This paper explores how visual design galleries can offer pedagogical benefits for novices, especially outside the classroom setting.

A learner-oriented design gallery would need to help novices find examples that are relevant to a learning situation and also present them in a way that optimally facilitates learning. Extensive research in educational psychology has delved into how sequencing examples in different orders may impact learning. However, these studies have mainly been conducted within structured settings (e.g. traditional classrooms or controlled laboratory environments [19, 20]. Moreover, most visual design galleries only allow users to browse and search examples based on surface attributes like colors or styles [25, 62, 63, 75]. Attending to surface features while browsing examples can cause fixation [52] and mimicry [49, 63, 67, 87], and might preclude learners from thinking about the deeper principles at play. According to Perceptual Learning Theory [86], presenting examples as contrasting pairs can help call attention to the deeper structure (i.e., design principle) more than showing examples individually [41, 42, 44, 81]. Similarly, tasting a flight of wines or viewing photos side-by-side can make subtle differences salient [86]. This research explores whether constructing a design gallery full of contrasting pairs of examples could both promote learning of deeper principles and help novices discover examples that are conceptually in their zone of proximal development [21, 78]. How might we design the user interface and interactions in a gallery so that novices can not only browse diverse ideas but also learn key principles instead of simply replicating or remixing surface features?

To explore this possibility, we performed two formative studies to guide the construction of a novel gallery system that leverages contrasting pairs of examples that show an early and later iteration of a design process. In Study 1, we created four gallery mock-ups and interviewed 12 novice designers to understand their perspective on a learner-oriented gallery. Each gallery mock-up emphasized different filtering mechanisms and presentations of design examples, such as filtering by key improvements or audience feedback. The mock-ups also probed two ways of presenting examples: showing a single image of the final version of a design (which is standard in most example galleries) versus presenting side-by-side image pairs that illustrate the early and late iterations of the same design. From the interviews, we found that novices strongly prefer viewing the early-and-late iteration pairs while browsing because it helped them find useful examples and highlighted improvements that are potentially applicable to their own work.

In Study 2, we investigated how to best construct examples of contrasting pairs to help novices recognize the key principles of visual design (e.g., hierarchy) Prior work on contrasting cases from the psychology literature suggests that the difference or similarity of contrasting examples could significantly affect whether novices extract insights [14, 40]. In an online experiment, 33 novice designers performed an assessment task to identify key issues on example pairs that either showed a single improvement or multiple improvements. The results showed that novices are better at identifying key insights (i.e., the underlying principles that led to improvements) when the example pairs highlight a single improvement rather than multiple. Our participants also expressed a desire for further explanation to aid their understanding of the examples.

Based on the insights from Study 1 and 2, we created ProcessGallery, a learning-oriented gallery that presents a collection of design examples using contrasting pairs of designs that highlight improvements tied to a single design principle. The gallery also enables users to filter examples by the types of improvements and principles they want to learn. To evaluate how ProcessGallery affects the process of finding and applying examples, in Study 3, we conducted a within-subjects evaluation ($N = 16$). Participants revised two seed designs (i.e., event flyers) after using both ProcessGallery and a baseline gallery. The baseline condition, modeled after traditional galleries, only shows final designs and only allows filtering by surface attributes. Our results showed that ProcessGallery guided novices to focus on the underlying principles in examples rather than surface attributes (e.g., color). Furthermore, after using ProcessGallery, participants were more accurate on a design assessment task, providing some indication of learning gains. Novices also significantly preferred interacting with ProcessGallery over the baseline tool because the new gallery made it easier for them to locate and apply useful examples for improving their work.

While this research initially explores the impact of ProcessGallery on individual participants, we hope the broader CSCW community gains insights about how to distill value from a collective resource (in our case, design examples created by online community members and augmented by our research team). Our work offers three contributions to the CSCW and learning communities. First, we contributed a novel gallery system that leverages contrasting examples and issue-based filters, instead of focusing on surface attributes. Second, we produced a collection of contrasting examples that could be used in a variety of instructional settings to teach core principles in visual design. Finally, our studies offer a deeper empirical understanding of how novices can learn from example galleries that illustrate early and late iterations of a design that improves over time. Our contributions provide implications for design tools and educational resources and point to several new avenues for future work.

2 RELATED WORK

We describe how our work extends the literature in the use of examples for inspiration exploration and learning. We also situate our contribution in the context of tools that support learning and applying examples in creative domains.

2.1 How Examples Generate Inspiration

Many design ideas are the outcome of surveying and applying existing knowledge such as past ideas and examples. Examples provide contextualized instances of how form and content integrate [63]. When facing a new challenge, people often search for similar situations in the past [74, 88] and replicate others' successful approaches as it is more efficient than reinventing them from scratch [63]. Examples also help designers evaluate the appropriateness and creativeness of their own solution by analyzing how the solution is situated in the existing space of designs. [13, 51].

While examples can give inspiration, they may also produce design fixation [11, 29, 76]. An informal survey from Stanford University [61] showed that designers' have mixed perspectives on the use of examples for getting inspiration; some designers think it helps expand their understanding of the topic, while others fear the potential conformity introduced by the exposure of examples [61]. Many researchers have examined how the timing of example delivery impacts creative outcomes [61, 85] and found that early and repeated exposure to examples led to more creative ideas than late exposure. Rather than presenting examples at regular intervals, proactively offering examples when users are stuck was more effective for improving the quantity and quality of the ideas generated for a product design task [85].

Another thread of work studies the effects of analogical distance between examples and target tasks on creative outcomes [23, 24, 39, 76]. A "near" example typically refers to instances in the same or similar domain and shares many surface features with the target design problem, while a "far" example is typically drawn from a different domain and shares little or no surface features. While most studies suggest that "far" analogical stimuli are more likely to prompt innovative solutions than near ones, examples that are too semantically distant from the target goal may hinder ideation [23, 24].

While much of the prior work on design examples seeks to avoid conformity or fixation to boost creativity, our work focuses on how novices learn deeper fundamental principles embedded within examples. In particular, our work adds to this literature by empirically testing how different example configurations affect the understanding of visual design principles.

2.2 How Examples Support Learning

Examples have also been shown to provide pedagogical values across domains as diverse as physics [26], design [93], and management [92]. Examples are concrete and more easily understood than abstract principles [41]. Learning from examples has been shown to be a more robust way to learn, and the knowledge gained through this approach can be better transferred to new situations, compared to learning from principles to specific examples (also called deductive learning) [7, 15, 83]. Comparing multiple examples can help knowledge transfer, as Gentner describes, "comparison processes can reveal common structure ... and thus promote transfer, even early in the learning when neither example is fully understood" [41]. Examples have also been used to supplement design feedback to help illustrate reviewers' points of view [54].

However, for examples to be effective, learners must be able to identify what to learn and how to apply the learned knowledge [49, 53, 70]. Unfortunately, many studies have found that quality alone is not enough to ensure effectiveness of an example [80]. Without guidance, students often fail to understand the intention behind the examples, even if they are good ones [43, 44, 80, 96]. In fact, replicating examples without discerning how and why it's effective can lead to design fixation, and sometimes, worse design [29, 50, 52]. In an engineering-design class, students were divided into example and control groups and tasked with a design problem; the example group received an example design and the control group did not. Analysis on the resulting ideas indicated a strong conformity effect that students borrowed many features in the examples, including even problematic ones [52]. This problem will only become more prevalent as learners are increasingly able to access examples of varying quality online.

Researchers in learning and cognitive psychology have investigated strategies for aiding in the discovery of and learning from relevant examples. A common approach is to organize examples into contrasting cases [81, 94]. Juxtaposing two examples and studying them simultaneously makes the contrasted features—that are tied to intended principles—more salient to learn [16, 45, 82]. Being exposed to multiple examples instead of one also mitigates design fixation [63] and helps form a more sophisticated mental model that expedites knowledge formation [41, 79]. We hypothesize that presenting examples that highlight contrasts can support the acquisition of relevant knowledge, helping novices relate examples to their own design situations for improvements.

However, incorporating contrasting pairs to teach iterative design presents a key challenge: How might we scale the production of contrasting pairs that illustrate various principles and obstacles learners may face during the creative process? In creative domains, people typically approach the same design problem in various ways; the distinct solutions usually contain different flaws that need to be addressed. An instructor may be able to produce examples illustrating common misconceptions, but it is nearly impossible to craft examples that cover all possible problems-to-be-solved for students. Juxtapeer has leveraged peer submissions to construct example pairs [17], but viewing two examples that were distant from each other (i.e., no shared surface features) did not guarantee knowledge development [41]. While the students performing a comparative review produced longer feedback that was more specific and included more expert terminology [17], the authors did not measure knowledge gains. Thus, it remains unclear how the comparative examples impacts learning and future design iterations.

Our work extends this thread of work by using iterative design as a resource for producing contrasting examples for novice learning. We hypothesize that presenting both the early and late iterations of the same design project not only better articulates the design mistake, but also offers inspiration for how to resolve it. Our work also provides insights on how to construct effective example pairs (single versus multiple improvements) and what information novices need to help their understanding of the examples.

2.3 Technologies for Searching and Learning from Examples

Most visual design galleries today only allow people to browse and search examples based on surface design attributes like colors or styles [25, 62, 63]. Lee *et al.* [63] developed Adaptive Ideas Web design tool that allows users to borrow elements from Web examples into their working canvas; the tool organizes examples by, for example, background color [63]. Bricolage enables example-based retargeting for web design by creating coherent mappings between pages based on its content structures [62]. GUIComp and other tools leverage computational power to analyze a work-in-progress design and provide real-time principle-based evaluations and examples [64], but they only present the final outcome of the design example.

Many online design communities encourage members to share their process in hopes of supporting learning [2, 58]. While designers may share intermediate drafts to illustrate their process [58], these example galleries do not effectively emphasize the insights or principles that led to improvements. Reddit (`r/design_critiques`) encourages their users to submit in-progress work to the community. However, viewing others' work to gain insights is tedious, as one has to manually click through each discussion thread and open multiple windows to view the designs [96]. Most galleries also only focus on and support searching and browsing around the final outcomes [25, 62, 63, 75]. Our work is unique in focusing on how to create a gallery that helps people find and learn from contrasting examples that illustrate process and emphasize the underlying principles.

Our ProcessGallery is distinctive in two ways: (1) it enables users to view curated early-and-late iterations of a design, making the key insights (e.g., the improvements) more salient, and (2) allows users to filter through examples based on fundamental principles, not just surface attributes. To our knowledge, our evaluation of the gallery is the first to measure both learning and performance (i.e., quality of the final revised design) on an iterative design task. Our results also shed light on the trade-offs between an improvement-based gallery (ProcessGallery) and an attribute-based gallery (most existing design galleries).

3 STUDY 1: EXPLORING GALLERY FEATURES THROUGH DESIGN PROBES

To investigate the potential for a gallery interface to support learning during the design exploration stage, we created a series of interactive mock-ups and conducted an exploratory probe with novice designers to understand their perspectives. While most galleries show only the final outcome of a design process, our probe explores how and whether a gallery could display examples as contrasting pairs using early and late iterations of the same design. Similarly, while typical UIs for browsing examples focus on surface features (e.g., color, genre, etc.), our exploratory probe illustrates how a gallery could leverage metadata (such as principle tags, expert ratings, and feedback) as filtering mechanisms to explain the differences between the contrasting pairs. The goal of our preliminary design probe study is to understand how novices think and feel about particular features aimed at supporting their learning process. We explored the following research questions:

- **RQ1**: How do novices prefer to view examples to support their learning process?
- **RQ2**: How do novices talk about finding and filtering examples to support learning?

3.1 Method

Below we describe our mock-up configurations and the study procedure.

3.1.1 Participants. Twelve participants (seven female, five male) were recruited through email distribution lists from an introductory design course at the University of California, San Diego. The participants were all between 18 and 24 years old and had no professional experience in visual design. The participants had all completed an assignment aimed at teaching basic principles in visual design; the assignment required the participants to use the Internet to find examples for learning, and then created and iterated on an event poster. The last author offered an optional extra opportunity to take part in the study, and participation was completely voluntary for the participants. The participants were explicitly informed that their feedback on the mock-ups would not affect the awarded participation credits. To prompt discussion, we asked the participants to

	Focus of Mock-ups	Example Presentation	Filtering Mechanisms
A	Attribute-focused	Single Image (Final Design)	Attribute Filters
B	Improvement-focused	Single Image (Final Design) + Metadata (Key Improvements)	Attribute Filters, Quality Filters, Improvement Filters
C	Iteration-focused	Early and Final Design	Attribute Filters, Quality Filters
D	Feedback-focused	Early and Final Design + Metadata (Feedback on Early Design)	Attribute Filters, Quality Filters, Feedback Filters

Table 1. Key features in the gallery mock-ups.

bring their posters to the study session. All study sessions occurred within a week of the submission of the poster assignment.

3.1.2 Interactive Mock-up Alternatives. Table 1 summarizes the key features in each gallery mock-up. The mock-ups were conceived by surveying existing research and commercial galleries for visual designs. All the examples displayed in the gallery mock-ups were collected from prior research projects. Figure 2 shows examples of two mock-ups, Gallery A and D.

Attribute-focused (Gallery A) presents the final outcome of each example using *single* images, organized in a grid view. Users can filter examples based on surface-level design *attributes*—genre, color, image usage, amount of text—which emulates existing portfolio sites such as Behance [2] and Dribble [6].

Improvement-focused (Gallery B) extends Gallery A by providing *text descriptions* about how each example improved upon an earlier draft. Users can filter examples based on its key *improvements*. Gallery B also introduces quality filters, inspired by Paragon [54], where users can filter examples based on the *quality* across different design aspects (e.g., the effectiveness of visual structure, visual focus, color contrast).

Iteration-focused (Gallery C) extends Gallery A and B by showing both the *early* and *final* iterations of the same example in pairs in a side-by-side view. In each example pair, the right image is the revision after addressing feedback given to the left poster. Users can filter examples using both surface-level design *attributes* and *quality ratings*. The filters can be used to search for either early versions, final versions, or both.

Feedback-focused (Gallery D) extends all the prior galleries further by enabling users to filter examples based on audience feedback. This is inspired by online design communities such as Reddit where users can search discussion threads by keywords based on their learning objectives [5].

3.1.3 Procedure. After getting informed consent, we presented participants with Gallery A (Attribute-focused) and asked them to browse the gallery interface and select two examples that were useful for learning. We started our exploration with Gallery A because it resembles popular design galleries such as Dribble and Behance, which may prompt participants to talk about their recent experience searching examples for improving their poster assignment. Then we prompted the participants to think aloud and talk about their impression of Gallery A's presentation and the filtering mechanisms.

Following the discussion of Gallery A, we presented participants with the three other gallery alternatives, which included novel features for filtering examples based on improvements or feedback (Gallery B and D respectively), quality ratings (Galleries B, C, and D), and for viewing examples as contrasting pairs (Galleries C and D). Since the "logic" of the mockups was built upon

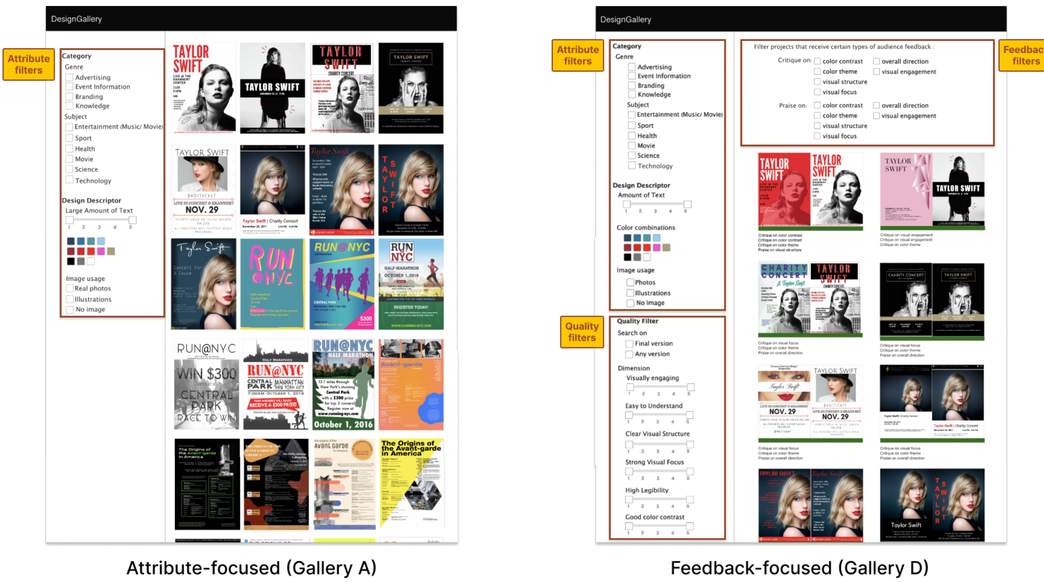


Fig. 2. Examples of two gallery mock-ups used in Study 1: Gallery A and D. Gallery A is attribute-focused, presenting single-image examples with attribute filters (e.g., genre, color). Gallery D presents both the early and late iterations of a design in a side-by-side view and enables filters by quality, and feedback given to the early draft, in addition to attributes.

each other, with later galleries incorporating features that were already introduced in the preceding ones, we decided to present these mockups in a fixed order. When introducing the later gallery, we focused participants' attention exclusively on the new features, without repeating those previously discussed. We asked participants to reflect on how the new features may help them browse, search, and learn from the examples.

We were aware of the challenge of participant response bias [31], as one of the authors was the class instructor. During Study 1, we took steps to ensure the participants were not asked to rank or pick only one of the mock-ups. Instead, we probed the value of the different features across all the mock-ups.

All study sessions were conducted remotely by the first author through video conferencing and lasted approximately 90 minutes. We offered assignment credits for completing the study. Participants shared their screens throughout the study. We informed the participants that their perceptions of the prototypes would not affect their course grades. All interviews were recorded and transcribed. Following Zhang and Cranshaw [99]'s method, the first author reviewed the transcripts and coded them for themes using an open coding approach [91]. We used inductive coding because there was no pre-existing theoretical framework regarding how and why a learner engages with an open gallery of examples for learning. Major themes were selected after multiple iterations and discussions with the rest of the research team.

3.2 Results

3.2.1 RQ1: How do novices prefer to view examples?

When browsing examples in Galleries A and B, six participants liked how the interfaces presented examples in a grid view rather than a slideshow view (seeing only one example at a time). Being able to view examples as a collection fosters visual comparisons and helps to spot the examples that interest them the most. One participant explained how she visually compared different examples and focused on what captured her attention:

> *"I first compare the posters because I'm like, what poster draws my attention first, and then figure out why this poster draws more than this one (point to another poster). Like this one, it has this yellow bold under the pink that makes the title really stand out and there's not a lot of information, but yet the important information gets across. And then this one I might not be the biggest fan because it kind of makes me lose focus. If you were to just give me one poster to look at, there's nothing for me to compare."* [P3, Female]

The participant described how Gallery A, like many traditional galleries, allows her to visually scan and get some sense of what is working. However, when we presented Gallery C or D that shows two iterations of a design, all 12 participants preferred to see the paired examples, rather than only a final version. Visually juxtaposing two design iterations helped participants draw connections and focus on the key strategies used to improve the design. One participant explained the value of the highlighting process:

> *"If I just saw the final poster, even if it was a super good one, I probably wouldn't notice a lot of things. For example, it was clear that the designer wanted us to really know that the event was at Central Park because they didn't bold the 'Central Park' text in the first one, and they changed '13-dash-one' to 13.1 miles because they realized that it was very, very unclear for the audience. Just seeing the past history of the poster shows me what are things that the authors themselves felt needed to be changed or they really want their audience to focus on."* [P1, Female]

While participants uniformly preferred viewing contrasting pairs that illustrate the process, they differed in terms of what caught their attention. One participant was struck by pairs that show extensive edits:

> *"This one is interesting, the final design is really good in my opinion, but the first one is...(laugh)...I definitely want to know what drives the designer to make such a huge improvement to their work. And for this one, the first draft is already really good, but the designer changed it from this blue with cursive style to this red, almost an entirely different mood. I would click on it and see why the designer thinks this mood or this color fits better than the initial one."* [P4, Female]

Two participants appreciated the pairs that isolated more singular changes from the early to later design. As one participant articulated:

> *"This pair is good because it shifts a singular change and I can see how that change affected the entire poster rather than a restructuring of the entire thing."* [P2, Male]

Our data raised the issue of whether contrasting examples should be more similar or more different when illustrating improvements over time, which we investigate further in Study 2. Independent of the degree of difference between pairs, three participants expressed the need for more explanation around what changed, why it was changed, and whether the changes were effective. One participant said:

> *"I love seeing the iterations at first, but since these are all really good posters, it's hard to differentiate which version is better because of what. I have to do a lot of work; like, I don't know what got improved in this pair, is it the visual focus, or legibility?"*[P4, Female]

In some of the gallery mock-ups, participants could see supporting text below example pairs describing its key improvements (in Gallery B) or the feedback that drove the iteration (in Gallery D). One participant talked about the value of seeing audience feedback on the first draft:

> *"I think it's unique to see from an audience point of view as what they feel like, and what draws their attention. If I made a poster like that, it's important to see what are the things I did similar to this designer that got praised on, and what are things that they're still getting critiques on and how would I be able to apply that to my own poster."* [P1, Female]

The participant described this process of relating to and mapping the audience feedback to their own design as a way to avoid similar issues. Three participants raised the point that seeing other designers struggle and receive critique helped to create a sense of psychological safety [37]. As one participant remarked:

> *"...it'd also help my ego. Sometimes seeing such great designs, I am like why am I not good enough? And I think it's a toxic environment: we all make mistakes, but people are only showing the best part of themselves. If I were able to see how people improved their work, it's a very positive environment and positive gallery: so there are iterations that designers do, they made mistakes and they just improved on it."* [P4, Female]

The experience of viewing someone else's process and feedback in the gallery is not the same as getting direct feedback, but it appears to facilitate some learning while creating emotional distance between the novice designer and potential critiques of their own work [38, 71].

3.2.2 RQ2: How do novices talk about finding and filtering examples?

After showing the traditional gallery configuration in the first mock-up, six participants anticipated using surface features (such as Genre, Color, Amount of text, etc.) to find examples. For example, one participant talked about using `Genre` to locate examples with similar subjects: *"I would select event posters to see how others allocate a lot of text. And also branding, I want to see some cool ideas to advertise a thing"* [P1, Female]. Three participants illustrated how they would use `Color` and `Amount of Text` as filters to find examples similar to their own design or align with their desired direction. One participant said,

> *"I would have a few selections of the colors I want to use, for example, I'm into black and white and red, I will definitely want to use this and see how others incorporate these colors, and maybe what other colors I should include."* [P5, Female]

Upon introducing the remaining mock-ups, five participants felt that the filters linked to specific improvements and feedback to be particularly useful. As one participant relayed:

> *"Searching by improved color contrast is more helpful because it shows how other designers solve specific problems within their design and the direction they took, given the challenges."* [P2, Male]

Participants did not want to just see good examples, but instead were seeking ways to relate their own work to others. One participant explained how using critique filters to explore examples might help her gain deeper insights into her own work:

> *"If I'm having problems with my own color contrast or visual structure, it's unique to see from an audience point of view what they feel like is good or poor visual structure, and what draws their attention. I think it's important to see what are the things I did similar to this designer that got praised on, and what are things that they're still getting critiques on."* [P1, Female]

The filtering mechanisms seemed to close the gap between the example gallery and one's own work. As P6 explained *"if I turn in my work and got critique on color contrast, I'll literally click the same one and see what it means or what I should improve on"* [P6, Male]. Likewise, P4 talked about how the filters *"seemed to take so much work off my shoulders"* [P4, Female].

Two participants appreciated seeing audience feedback and related those critiques to their own work, which might prevent them from making similar mistakes.One participant stated:

> *"I'd first find examples that are similar to my original poster and then based off of that look at how others improve their color contrast from their original poster. I think that would also help me figure out what problems I even have with my first poster. And how to improve it based on their revision."* [P1, Female]

Unlike the other filters that received nearly unanimous appreciation, the quality filter triggered mixed opinions. Five participants thought it would be useful for prioritizing and comparing examples with varying quality. One participant said:

> *"I would want to see something that's really visually engaging but also at the same time, I would want to look at something that is like low rating for visually engaging because that gives me ideas that I shouldn't go this route."* [P6, Male]

However, P1 shared concerns about the objectiveness of ratings: *"everyone's eyes are very different, I might see something that I feel like it's a such a good design, but someone else might like completely disagree with me"* [P1, Female]. Three participants strongly opposed the idea of including ratings. P4 said:

> *"Who's the one that determines this is good, and this is bad? If I was a user and I wanted to be better and then I got the label saying my design is bad, I'm not going to use this anymore. Putting quality labels on the designs created this kind of competition between people who get good and bad ratings. I want it to be a peaceful place that we just look at designs, improve, and learn."* [P4, Female]

While ratings provide a scaffold for browsing examples with differing quality, knowing the potential of being judged publicly seemed to decrease the learners' willingness to participate and contribute to the community.

3.3 Study 1 Discussion

This design probe study provided a number of insights that will inform the creation of a learning-oriented gallery of examples. First, our novice designers preferred seeing the evolution of an example rather than only the final outcome of the process, as that helped direct attention to the key improvements in each example [42, 45] and cultivate an inclusive environment for learners [58].

Second, most participants talked about looking for examples similar to their own work. While participants were familiar with surface-attribute filters in traditional galleries, they especially appreciated the ability to filter based on principled improvements and prior feedback. Participants talked about how they would relate to the same critiques encountered by other designers and how this narrowed the gap in finding useful examples. Viewing how others address a similar critique is an analogy for learning by observing others approaches [27]. It is however unclear how this might work in practice since the designer would need to be able to initially identify the problems in their own design, which can be difficult for novices [94].

Finally, we also found that novices need guidance on judging the quality of either design examples or their own work. While expert ratings could serve the purpose, many suspect adding quality ratings on design may lead to competition and that creative work is subjective and should not be labeled as good or bad. Showing a before-after view of the same design helped guide their judgment on what is relatively improved, instead of needing to reference their own knowledge (which is limited [94]).

We also learned that novice designers were divided on whether to include early-and-late example pairs with more extensive or isolated improvements. An open question is what makes a contrasting example effective for learning. We explore this question in Study 2. Another open question is how the ability to identify issues in a work-in-progress design impacts how novices select examples and how the examples actually benefit their design iteration. We introduce a gallery specifically for learning and compare it to a traditional gallery for an iterative design task in Study 3.

4 STUDY 2: INVESTIGATING THE EFFECTIVENESS OF CONTRASTING PAIRS

In Study 2, we build on our insights from Study 1 that novice designers benefit from seeing how a design improves from one iteration to the next. Prior work has shown that presenting contrasting examples that differ along key dimensions calls attention to the features being contrasted and this, in turn, can help people learn deeper principles [16, 42, 45]. For the goal of teaching visual design, contrasting pairs could be created such that salient differences highlight key principles, like hierarchy and alignment. While the Study 1 participants all agreed that contrasting early-and-late designs provide insights, they expressed varying preferences on whether the pairs should be near or far from each other.

A key question arises around *how many improvements* should be illustrated in one example pair. In practice, designers often have to change multiple aspects of a design within one design cycle. While example pairs that show multiple improvements might be more authentic, they might overwhelm and negatively impact novices' ability to recognize principled insights [94, 96]. On the other hand, providing two iterations that address only one principle may make the intended knowledge more salient [39, 63, 86].

To investigate this question, we manufactured two sets of example pairs that highlight either one improvement or several improvements — each improvement addresses one principle — and ran a between-subjects online experiment where novice designers were asked to identify design issues in the assigned example pairs. Each improvement addresses one principle issue. We were also curious if novices needed additional information about example pairs to aid their understanding. Therefore, Study 2 explored the following research questions:

- **RQ 1**: (How) does the degree of change within contrasting pairs of examples impact novice designers' ability to recognize the intended insights?
- **RQ 2**: What additional information could help novice designers make sense of contrasting examples?

Answers to these questions will not only inform how we populate a learning-oriented gallery with contrasting examples, but also will give us empirical data on how novices extract insights. The study also offers implications for how to source or manufacture such a collection of example pairs and how to supplement the examples to improve learning.

4.1 Method

We conducted an online between-subjects study with two conditions: `single-improvement` versus `multi-improvement`. Participants performed an assessment task where they identified the key insights they observed in a series of example pairs that show one improvement or multiple improvements.

4.1.1 Participants. 39 participants (16 female and 23 male) responded to our open call for participants through Amazon's Mechanical Turk [1]. Participation was limited to U.S. residents with some college education to ensure sufficient English writing ability for the open-ended questions. Data from six participants were removed due to failure to complete an attention verification task: under one of the example pairs we wrote "For this question, please select 'No improvements' from the options below." From the remaining 33 participants, all but one reported no academic training

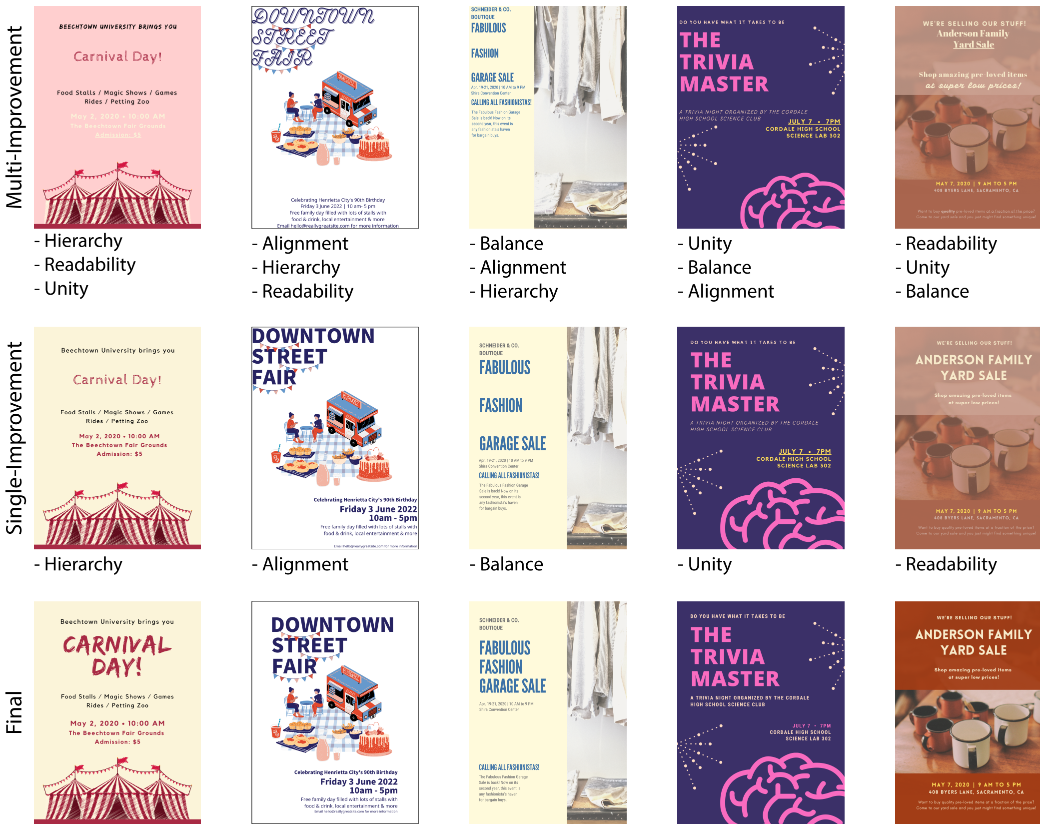


Fig. 3. Example stimuli for Study 2. Each column represents a set of designs for that stimuli. The stimuli were visually presented as pairs: either the `multi-improvement` (top) or the `single-improvement` (middle) design as the early iteration and the corresponding final design (bottom) as the later iteration. The `multi-improvement` earlier iterations exhibited multiple issues (corresponding to the principles listed below) that were improved upon to achieve the final designs; similarly for the `single-improvement` earlier iterations, but with just a single issue fixed for the same final designs.

or professional experience in design. Six participants (18%) had experience creating posters for their own personal projects.

4.1.2 Experiment Stimuli. Two design experts leveraged template designs from the popular online design editing tool (Canva [3]) to create examples as either `single-improvement` pairs or `multi-improvement` pairs (see Figure 3). All participants reviewed five example pairs that collectively covered key principles in visual design: Hierarchy, Alignment, Balance, Unity, and Readability. The principles were selected by surveying visual design textbooks and other resources [4, 32, 66, 95]. To understand common violations made by novices, the research team analyzed the feedback given to past student posters. Through a thematic coding process, we distilled a list of 37 common issues across the five principles. Table 3 displays the issues and how they map to the principles. The experts started by selecting high-quality templates and then editing them such that they include one (for `single-improvement`) or three issues (for `multi-improvement`); each issue was tied to one of the five principles.

4.1.3 Procedure. After digitally signing an informed consent form, the participants were asked to review the textbook definitions of the five design principles. Participants were then asked to confirm the statement, "I have read the descriptions and am ready to take the test." Prior work has shown that self-affirmation leads participants to follow instructions better and reduces incidents of cheating [89].

Participants were then shown five example pairs based on their assigned condition, one pair at a time. We utilized a feature called Randomizer in our survey tool (Qualtrics.com [10]) to randomly assign task respondents to one of the two experimental conditions (`single-improvement` versus `multi-improvements`). We chose an even distribution approach to ensure that our participants were randomized evenly. The survey tool also randomized the order of example pairs within conditions. Participants were informed that the poster on the right was a revision of the poster on the left, and their "Issue Identification Task" was to select all principles that showed improvements going from the left to the right poster. The principle definitions were available to participants throughout the study. Participants in either the `single-improvement` or `multi-improvements` conditions were instructed to select all the improvements they found in each pair. After participants completed the Issue Identification Task, they answered a post-task survey about their experience with the task (see Section 4.1.3).

4.2 Data Analysis

We measured the impact of example pairs in two ways: (1) the issue retrieval rate (i.e., how accurately the improved principles were identified by the participants), and (2) learning effort (i.e., how mentally demanding it is to find the improved principles). If exploring examples to learn principles requires too much cognitive workload, it is unlikely a learner would voluntarily browse a gallery full of example pairs.

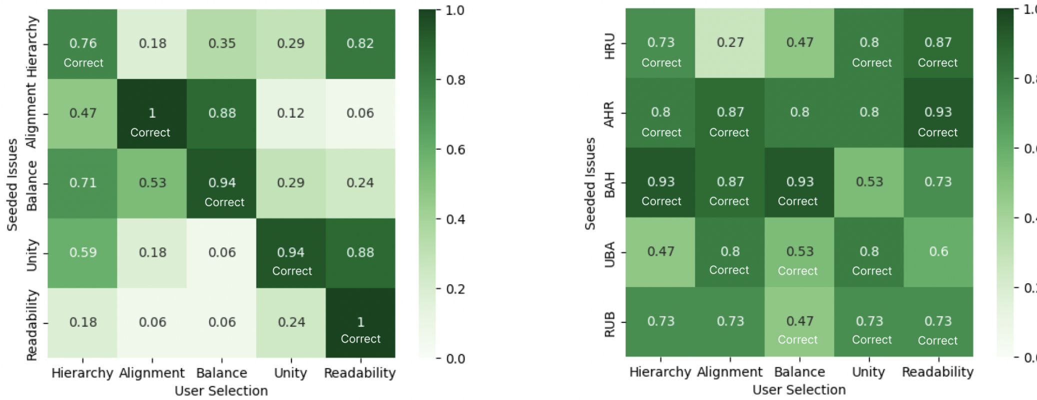
To calculate the issue retrieval rate, we compared the set of issues selected by participants with the set of issues intentionally embedded in the example pairs. In the `single-improvement` condition, there are a total of 85 items that were supposed to be identified (one principle per pair x 5 pairs x 17 participants), whereas, in the `multi-improvement` condition, there were 240 items (3 principles x 5 pairs x 16 participants). We computed the percentage of the retrieval rates in both conditions and ran a two-sampled z-test for proportions to check its significance.

To measure mental workload, we included the NASA-Task Load Index (TLX) questionnaire [48] in the post-task survey, which includes six Likert-scale questions about the task, pertaining to (1) mental demand, (2) physical demand, (3) temporal effort (how hurried or rushed it felt), (4) perceived success in accomplishing the task, (5) level of hard work required, and (6) frustration (how insecure, discouraged, irritated, or annoyed the task). We used the unweighted TLX for this analysis. Prior research has shown that unweighted TLX scores are highly correlated with weighted ones [18], and are gaining popularity due to the reduced amount of time needed to administer the task.

The post-task survey also asked the participants to explain: (1) how useful were the poster pairs for learning the five design principles? and (2) what other information could have helped you study the examples? The lead researcher analyzed the open-ended responses using an iterative open coding approach [91].

4.3 Results

4.3.1 RQ1: Single-improvement pairs helped novices identify principles more often than multi-improvement pairs.



(a) Single Improvement

(b) Multiple Improvements

Fig. 4. These heatmaps depict the percentage of intersection between participants' issue selections and the intended seeded issues by principle, indicating participants' performance in the Issue Identification Task. The cells marked as "Correct" indicate a correct issue selection. For example, the top row in (a) shows that 76% of the participants in `single-improvement` condition selected Hierarchy as the key improvement in the examples (correct), while 18% of the participants selected Alignment, which was incorrect for this example. Note that participants in both conditions were instructed to select all the key improvements they saw in the example pairs.

Figure 4 shows the summary of the percentage of intersection between participants' issue selections and the intended seeded issues. Participants in the `single-improvement` condition more consistently identified issues compared to the `multi-improvement` participants ($z = 1.95$, $p = 0.03$). Of the 85 issues embedded in the `single-improvement` pairs, 63 of them (74%) were successfully identified by the participants. In the `multi-improvement` condition, only 51% of the manufactured improvements were identified.

When viewing examples with multiple improvements, participants tended to miss key insights that were otherwise obvious in the `single-improvement` condition. For example, when reviewing the fifth example of the `multi-improvement` pairs (Readability + Unity + Balance), only 43% of the participants recognized that the overall unity of the design improved (i.e., too many variations in text treatment). In contrast, 93% of the participants in the `single-improvement` condition successfully identified the same issue (see Figure 3, fourth pair in the `single-improvement` condition).

In terms of learning effort, we did not measure a significant difference in the overall mental workload (the unweighted NASA-TLX scores) between conditions (`single-improvement` = 9.6, `multi-improvement` = 9.7, on 21 point scale, *n.s.*), according to a one-way ANOVA test. In both conditions, participants experienced medium to low effort in completing the task, as evidenced by comparing the scores to other similar tasks [47]. However, the ratings specifically for "frustration" were almost significantly higher for `multi-improvement` participants ($\mu = 7.53$) compared to `single-improvement` participants ($\mu = 4.12$, $p = 0.06$). As one participant reflected in the open-ended feedback:

> *"It would have been nice to see the examples spelled out clearly, instead of just having to figure it all out myself."* [P9, `multi-improvement`].

P9's response reiterated the challenge of identifying issues in work-in-progress work, especially when the number of edits increased.

4.3.2 RQ2: Participants asked for explanations and highlights of improvements.
Not only did participants in the `single-improvement` identify a higher percentage of issues, but several participants talked about how they learned from the example pairs.

> *"Coming from someone with beyond zero experience it really was pretty helpful to learn the absolute basics in a very fast manner."* [P14, `single-improvement`]

Towards the idea of using example pairs as a learning resource, a dozen participants across conditions commented that they still wanted more information about the issues and the rationale behind each design edit. As one participant explained:

> *"The designers' explanations would have been very useful. Understanding why the designer made those edits would help me learn more about each element and why it was incorrect in the original version."* [P6, `single-improvement`]

P6's response echoes some of the Study 1 participants who requested for explanations to help them learn from examples. When viewing examples with multiple improvements, several participants wanted to see visual indicators that highlight specific revisions between the example pairs.

> *"If there were posters that were drastically different from one another. Maybe some arrows pointing to the areas that were changed and the word next to the arrows."* [P3, `multi-improvement`]

Another participant wanted to see the entire revision history so they could see its progression:

> *"It would have been useful to see the process of editing, showing how text is moved, re-sized, fonts changed—a step-by-step slideshow that would help illuminate each principle."* [P10, `multi-improvement`]

The idea of a step-by-step slideshow of revisions aligns well with the scenario represented in our `single-improvement` condition where each pair only shows one key improvement, i.e., one step toward the final iteration.

4.4 Study 2 Discussion

Study 2 shows that isolating a single improvement in contrasting examples helped novices more consistently identify the principle tied to that improvement, compared to examples that show multiple sweeping improvements at once. When two instances are adjacent to each other, people tend to judge them relative to each other rather than on their own merit [17]. But too many differences within a pair can make it hard to draw connections [14].

Presenting contrasting images alone without offering explanation may not be sufficient for comprehending the key insights embodied in the example. This is especially true in creative domains where people with different backgrounds or expertise typically perceive the same creative work in different ways [98]. Our novice participants expressed the desire to read the designers' rationale for making those edits. Just as Schön characterized design as a reflective conversation with the design situation [80], externalizing such design rationale may help others learn vicariously from prior design experiences [33]. Future work can explore the effects of explanations or annotations [35, 36, 90] and whether they help to support comprehension of multi-improvement pairs.

Study 2 produced insights that informed the construction of our learning-oriented gallery for visual design. ProcessGallery, which we evaluate in Study 3, therefore included single-improvement pairs along with explanations of each improvement. Producing such example pairs, however, is non-trivial and time-consuming. Study 2 findings also point to angles for future work: how to mine examples from others' iterative process, and how to capture their step-by-step edits and rationale for changes in a design. Also, Study 2 focused on identifying issues, which is not the same as learning how to apply examples to one's own design work. Thus, we directly explore how this strategy impacts design performance in Study 3.

At the beginning of the task, we provided text-based definitions of the five principles and asked participants to self-validate their understanding. While self-affirmation may encourage adherence to instructions [89], it may be inadequate to gather evidence of learning. According to Bloom's taxonomy of learning [60], one's declarative knowledge differs from their ability to assess designs based on that knowledge or apply it to their own creative work. There is a possibility that our participant groups were unbalanced in these aspects. However, we believe that the random assignment across 39 participants could mitigate this probability. Additionally, as visual design examples were the primary stimuli in Study 2, we chose to present text-only definitions of the principles to establish key terminology that would be used in the interface. Introducing visuals during the instructions stage could have interfered with the controlled experiment. Future research is needed to develop instruments to test knowledge in design.

5 THE PROCESSGALLARY SYSTEM

To continue our line of inquiry, we take the result from Study 2 that single-improvement pairs help novices gain insight and set out to construct a gallery of diverse single-improvement examples.

Based on the insights from Studies 1 and 2, we designed and implemented ProcessGallery, a novel gallery interface that presents visual design examples as contrasting pairs that highlight revisions based on design principles. We hypothesize that the presentation and user interaction in the tool will assist novices in finding examples relevant to their design goals (e.g., by filtering based on intended revisions), extracting insights from examples (e.g., by reading explanations about what had improved), and applying what they learn to a work-in-progress design (e.g., by seeing the evolution of how others address a similar design issue). Below, we describe ProcessGallery's key features and how they may facilitate the discovery of relevant examples for learning and improving a visual design.

Illustrate Process in the Browsing Interface. ProcessGallery presents both the early and late stages of an example design side-by-side, each pair comparing a design project before and after addressing a specific issue. The poster on the right is a revision of the poster on the left. Figure 5 (right) shows an example of the detailed view of an example pair. Using this presentation, users can garner initial impressions and identify examples that appear most applicable to their design. Users may browse example pairs relating to common violations within the five visual design principles described in Section 3.3, see Figure 5 (left). To aid sense-making, ProcessGallery displays a tag under each example pair indicating the issue resolved through the iteration. The tags are color-coded to indicate the high-level principle (e.g., "Hierarchy") each issue falls under (e.g., "Weak Point of Entry").

Surface Principles and Issues in the Filtering Mechanisms. The gallery provides filter drop-down menus for each principle. Users can click on the principle to see common issues and then select any issues they want to learn about or address in their own work. Users can select one or more issues and the gallery will respond by displaying the subset of examples that match any of the selected filter criteria. Users can read the definitions of each principle by hovering over its corresponding information icon. Prior work has suggested that proactively presenting high-level domain knowledge (e.g., the meaning of the design principle) at specific points of time when most relevant to the user's process (e.g., while browsing related examples) can promote learning [70].

Highlight Improvements in the Detailed View. Based on our Studies 1 and 2, novices requested explicit explanations describing the improvements between iterations. In ProcessGallery, users can click on the example pair to see a detailed view with an explanation. The explanation page

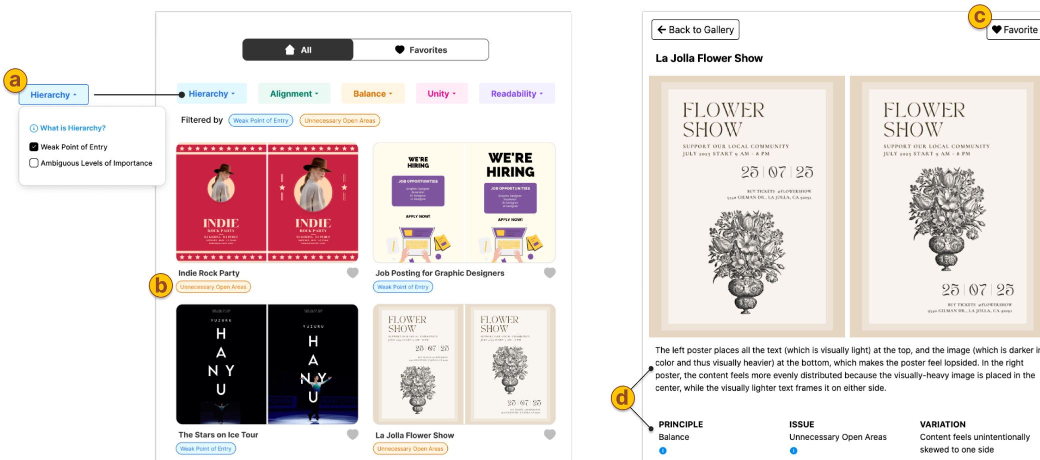


Fig. 5. ProcessGallery's browsing interface (left) and detailed view (right). While browsing, a user can interact with the principle drop-down menus (a) to filter examples based on related issues they intend to learn or address. In the drop-down, the user can hover over unfamiliar terms (e.g., "What is Hierarchy") to read their description. The tag under each example pair (b) indicates the issue addressed by the iterations. On the detailed project page, they can save useful examples to their "Favorites" collection (c). They can also read a written justification for the designer's iterations (d).

shows the enlarged design images along with a written description. Users can "Favorite" as many examples as they need and revisit their collection in the "Favorites" tab. For each favorited example, ProcessGallery provides a text box, prompting the user to write what they learn from each example and how they might apply it to their own work. Requiring self-explanation on an example increases a learner's engagement with its content, which can lead to greater understanding and knowledge transfer [26].

5.1 Constructing the Example Pairs

We repeated the same procedure in Section 4.1.2 to create 52 `single-improvement` contrasting pairs (104 poster) to populate ProcessGallery. Each example pair demonstrates one common violation, and collectively, the examples in ProcessGallery cover all 37 violations (identified in Section 4.1.2) across the five visual design principles: Hierarchy, Alignment, Balance, Unity, and Readability. The design experts also generated the metadata necessary to construct the ProcessGallery interface, including the issue information and detailed explanations of the example's improvements. They also labeled surface attributes of the examples including the topic, purpose, and types of graphics used in the examples. We also ran an open-source Python script to extract the dominant color for each example poster [8]. These surface attributes were used to construct a more traditional gallery that was used as a baseline condition in the tool evaluation in Study 3.

5.2 Implementation

To establish our evaluation environment, we used Polotno [9], an open-source web-based design editing tool, as the foundation for ProcessGallery. Polotno provides essential JavaScript libraries and React components that support the necessary functions for creating visual designs. The user interface comprises three major sections (See Figure 1). The left panel includes all the essential tools for creating or modifying a design, e.g., options for Background images and Text blocks. The central canvas area is where users can compose their visual designs. To streamline the design task, we removed unnecessary features such as templates. The gallery presentation and filtering mechanisms are located on the right side, enabling users to browse and search through examples while working on their design. A Firebase database[1] was used to store example pairs and user-provided data (e.g., filters selected, favorite examples, and the designer's text explanation for selected examples).

6 STUDY 3: PROCESSGALLERY EVALUATION

To investigate how ProcessGallery impacts how novices learn and apply examples for iterative design, we conducted a comparative evaluation where novice designers revised provided designs after browsing ProcessGallery and a baseline gallery. The baseline gallery resembles the attribute-focused gallery (Gallery A) in Study 1 and presents only the final outcome of each example using a single image. In the baseline gallery, participants can filter examples by genre, topic, color, and image use. Both galleries contain the same 52 examples, but ProcessGallery presents examples as pairs of early-and-late iterations, which results in 104 design images in total. The study was designed to answer the following research questions:

- **RQ 1:** How does ProcessGallery influence the way novices browse and select examples compared to a traditional gallery?
- **RQ 2:** How does ProcessGallery impact novices' ability to identify and improve issues in a given seed design compared to a traditional gallery?
- **RQ 3:** Which gallery do novices prefer for learning and which do they prefer for improving iterative design? and why?

6.1 Method

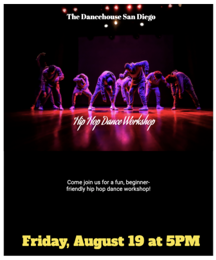
To evaluate the tool, participants completed iterative design tasks using both ProcessGallery and the baseline gallery. We used a within-subjects design to control for variation in participants' design knowledge, familiarity with existing design galleries (e.g., Dribble, Behance, etc.), and experience revising a poster design. We also wanted participants to be able to explicitly compare and reflect on the differences between the two gallery experiences. To eliminate individual differences when creating an initial draft design, all participants were given seed designs to assess and improve based on input from the galleries.

6.1.1 Participants. Sixteen participants (9 female, 7 male) between the ages of 18-34 were recruited via email, word of mouth, and by posting in multiple student forums. All but three participants had no educational or professional experience in visual design. None of the participants were part of our two formative studies or had advanced knowledge of the project. We focus on novices and beginner designers in this evaluation because we feel this population will benefit the most from a learner-oriented design gallery.

6.1.2 Task Materials. Participants were provided with two event marketing flyers to assess and iterate on (Figure 6). Each flyer exhibits a set of issues that violate multiple principles in visual design. The flyers were created by a design expert who has experience teaching visual design principles. Both flyers contain text and images and require little to no prior domain knowledge to comprehend. The flyers had a similar number of issues.

6.1.3 Procedure. Each participant completed two design scenarios, each using a different seed design and gallery condition. We counterbalanced both the order of the two seed designs and the tool conditions. The two scenarios followed the same structure. Each study session lasted roughly

[1]https://firebase.google.com/



Dance Workshop Flyer

Targeted issues:

- Hierarchy – Weak Point of Entry
- Hierarchy – Ambiguous Levels of Importance
- Alignment – Insufficient Margin
- Balance – Unnecessary Open Area
- Unity – Inconsistent/Too Many Variations in Text
- Unity – Inconsistent Color Choices
- Readability – Poor Text Legibility

Bakery Fundraising Flyer

Targeted issues:

- Hierarchy – Ambiguous Levels of Importance
- Alignment – Insufficient Margin
- Balance – Not Enough Space Between Content
- Unity – Unnecessary Design Elements
- Unity – Incohesive Graphic Choices
- Readability – Content Obscured

Fig. 6. The seed designs used in Study 3: a dance workshop flyer (top left) and a bake sale fundraising flyer (bottom left). Participants were asked to evaluate and iterate on the flyers after using ProcessGallery and the baseline tool.

2 hours, and we provided $30 Amazon gift cards as compensation. During the study session, the experiment proctor was available only for resolving technical issues. We describe the steps in the scenarios.

0. Practice Task. Before the main tasks, participants were guided to perform a set of actions on the canvas to familiarize themselves with the features in Polotno's canvas editor. For example, one task instructed participants to click on the Text section and create a header that says "Hello." All participants successfully completed the practice tasks within 3 minutes."

1. Review a Design Brief. Participants first reviewed a fictional design brief that communicates the event details, design goals, and requirements for the task. Participants were asked to help the client improve the draft flyers (see Figure 6) to deliver event details more clearly and in a visually appealing way. The experiment facilitator loaded the seed design into the tool before the study started. Participants were informed that they would be revising the design in the later stage.

2. Assess the Seed Design. After reviewing the design brief, participants completed a design assessment task to evaluate the seed design using 15 predefined issues organized by associated high-level principles (see Figure 7). Participants checked all the issues they observed in the seed design (in a Yes/No binary format). On the page, participants could hover over unfamiliar terms to view their explanations. Participants performed the same assessment task across conditions for the assigned seed designs. We use this score as a proxy for participants' existing knowledge, as they had not yet browsed examples or engaged in a design task. We refer to this score as a participant's `pre-gallery knowledge` in the rest of the paper.

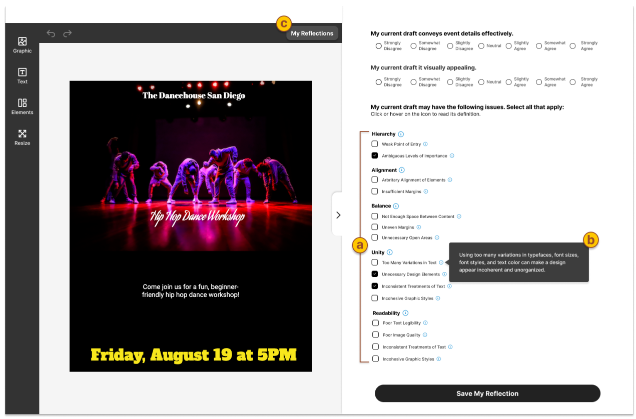


Fig. 7. The interface for assessing the seed design. Participants can select issues they observe in their seed design using the assessment checklist (a) and browse tool-tip definitions for each checklist item (b). All results are stored in a reflection summary that can be accessed throughout the task (c).

3. Browse a Gallery. After reviewing the seed design but before actually revising it, participants were asked to browse examples using the gallery corresponding to the current condition (ProcessGallery or baseline). Their task was to select at least three examples that they believed would be useful and inspirational for helping them address the issues they identified in the previous step. For each selected example, we asked the participants to write one or two sentences describing why this example was useful. Throughout this step, participants could click on "My Reflection" (see Figure 7-(a))to see their assessment of the seed design.

4. Re-assess the Seed Design. After browsing the example gallery specific to each condition, the participant repeated the assessment task (in Step 2). This step allows us to see if participants' assessment changed after viewing examples, and before they started editing the seed design. We refer to this score as a participant's `post-gallery knowledge` in the rest of the paper.

5. Revise the Seed Design. After re-assessing the seed design, participants were asked to use approximately ten minutes to revise the seed design. As basic requirements, the final flyer had to incorporate all the provided text and use at least one image from the media library provided by the client. During that time, they could keep browsing the gallery. Their goal was to address the issues they identified through the assessment task.

6. Complete a Survey. At the end of each design scenario, participants completed a questionnaire about their experience under the relevant condition. The questionnaire asked participants how the various tool features facilitated (or hindered) their tasks.

After both scenarios, participants filled out an exit survey in which they compared ProcessGallery and the baseline gallery on a 7-point scale (1 = baseline tool preferred, 7 = ProcessGallery preferred); a score above 4 would indicate a preference for ProcessGallery. The survey included questions about which gallery made it easier to: (1) find examples that were useful and inspirational for addressing the issues in the seed designs, (2) figure out what to learn from examples, and (3) apply the lessons from the examples to the seed design. We also asked which gallery: (4) do you believe can help you learn design principles better? (5) made it easiest to explore many different designs and ideas? (6) made you feel more creative while doing the activity? and (7) made you feel most satisfied with what you got out of the system, given the time spent. Finally, we also surveyed their perceived cognitive workload for browsing the examples: (8) which tool made you feel less overwhelmed due to the amount of examples available in the gallery?

At the end, the experimenter interviewed the participants, asking them to demonstrate their strategies for finding examples and explaining how these examples impacted their revision. In the survey and interview, ProcessGallery was referred to as `improvement-based` gallery, and the baseline tool was an `attribute-based` gallery. We avoided the use of our gallery name or "baseline gallery" to mitigate the bias toward either gallery. Figure 8 shows examples of the revised design by our participants and their respective favorited examples.

6.2 Data Analysis

The design assessment task (in Steps 2 and 4) was used to measure participants' knowledge of visual design. The task is worth 14 points; to receive one point, participants must correctly select if the issue exists or does not exist.

In order to compare how well the novices improved the seed designs, two experts who had experience teaching visual design collaboratively evaluated the revisions submitted by our participants. The experts first assessed the degree of change between the seed design and the revision on a 7-point scale (1 = minor change, 7 = extensive change), and identified the number of issues they observed using the same issue list from the assessment task. A low number (out of a total of 14 issues) indicates a high-quality design.

We used linear mixed-effects models to examine the effects of the condition, seed design, and `pre-gallery knowledge`, on our study measures. We chose mixed-effects models to account for the individual variability due to the use of a within-subjects study design. For the comparative ratings from the exit survey, we performed one-sample t-tests using the neutral rating (4) as the population mean. Open-ended responses in surveys and interview transcriptions were analyzed using an inductive coding approach to develop themes. We also analyzed the screen recordings from each study session to identify patterns of use.

6.3 Results

All participants (N = 16) successfully completed the two design scenarios using ProcessGallery and the baseline gallery. We collected 16 revised designs for each seed design.

6.3.1 RQ1: ProcessGallery helped people find examples related to intended revisions.

Participants in both gallery conditions actively used the drop-down filtering features while browsing and searching for useful examples. Table 4 shows the counts of the different gallery filters participants interacted with in the different scenarios. Analysis of logs for filter usage showed that ProcessGallery participants selected a similar number of filters ($M = 3.8$ filters selected per person) compared to baseline participants ($M = 4.6$ filters, *n.s.*).

When using ProcessGallery, participants primarily searched for examples that showcase the process of resolving issues they identified in the seed designs. For example, when using ProcessGallery to filter examples for improving the Dance Workshop flyer, six people used "Weak Point of Entry" and five used "Too Many Variation in Text", which reflected the key issues in the seed design. Some seed design issues—such as "Unnecessary Open Area"—were identified less often (2 out of 8 participants in this task scenario) and hence, only got used as a filter by two participants.

With the baseline tool, participants mostly searched for examples that shared similar surface features with the seed design. For example, in the Bakery Fundraising scenario, the most selected filters are "Food" ($n = 4$) and "Fundraising And Charity Ad" ($n = 3$).

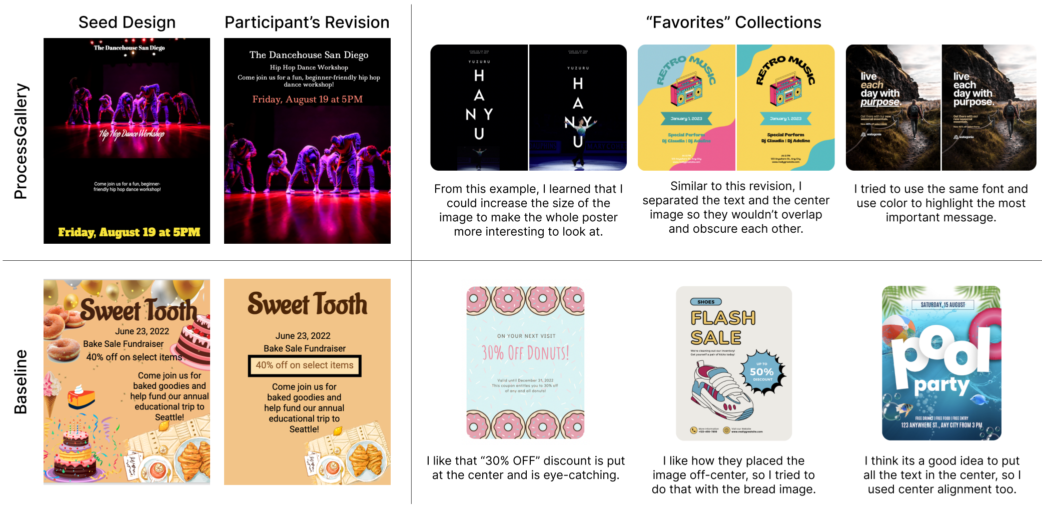


Fig. 8. Seed designs, participant revisions, and their respective "Favorites" collections for both conditions. Upon identifying salient differences between pairs in ProcessGallery, participants describe their intent to apply techniques similar to those observed in their "Favorites."

A mixed-effects model suggested that the condition did not impact the amount of time participants used to browse the galleries ($ProcessGallery$ = 6.9 minutes versus $baseline$ = 5.7 minutes, *n.s.*). However, our participants reported in the exit survey that they were more satisfied with what they found when using ProcessGallery than using the baseline gallery (Q7: $M = 5.53$ (ProcessGallery preferred), $SD = 1.46$; $t(15) = 4.07$, $p = 0.001$). One participant said:

> *"I think this (ProcessGallery) helps guide my search better, I can quickly filter examples for the issues I want to work on later. When I used the second gallery (baseline condition), I don't really know what I am looking at. All the posters look good to me, I end up just "guessing" what might be useful for me."* [P12, Female]

The issue-based filters provided a scaffold for our novice participants to explore examples by intended revisions.

6.3.2 RQ2: ProcessGallery improved participants' ability to assess and refine a design.

Participants in both gallery conditions could more accurately identify the issues in the seed design after browsing the examples. Although the improvements seem larger in the ProcessGallery condition ($Score_{before}$ = 9.1 (SD = 2.6) versus $Score_{after}$ = 9.9 (SD = 2.3)) compared to Baseline: ($Score_{before}$ = 9.1 (SD = 2) versus $Score_{after}$ = 9.6 (SD = 3)), our mixed-effects model suggested that the difference was not statistically significant. No ordering effect was present. In the post-interview, one participant articulated how ProcessGallery helped her realize new issues in the seed design:

> *"I've always known that it's important to leave enough space for printing but I did not see it (the "Insufficient Margin" issue in the bakery poster) until I saw this pair (pair #26, where we highlight the fix of insufficient margin problem) and was like oh right! The margin was really small. I think being able to see them side-by-side does help me catch these otherwise nuanced details that I might miss if only seeing a single image."* [P12, Female]

	Hierarchy	Alignment	Balance	Unity	Readability
ProcessGallery					
Bakery	0.5	0.5	0.63	0.5	0.5
Dance	0.63	0.38	0.25	0.75	0.5
Cross Scenarios	0.56	0.44	0.44	0.63	0.5
Baseline					
Bakery	0.75	0.75	0.88	0.75	0.5
Dance	0.5	0.63	0.63	0.5	0.5
Cross Scenarios	0.63	0.69	0.75	0.63	0.5

Table 2. Summary of the percentage of revised designs that still had issues in each principle.

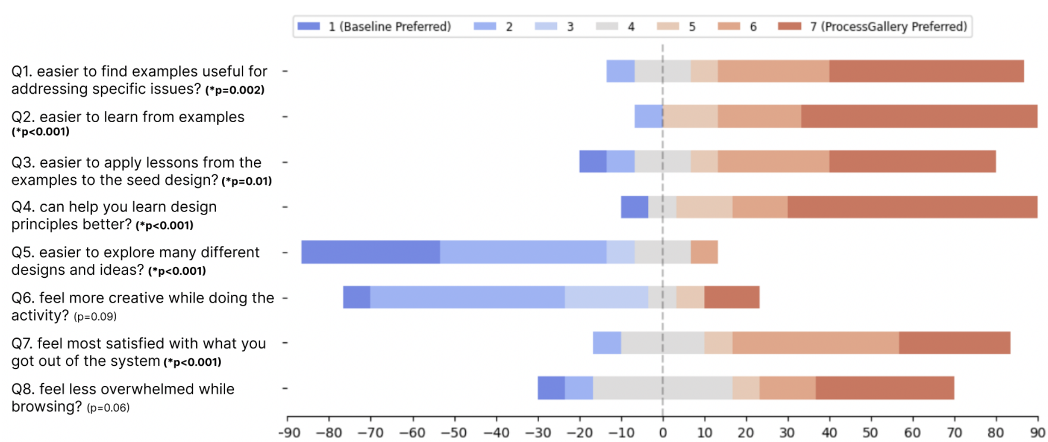


Fig. 9. Diverging stacked bar chart summarizing the survey responses from participants after completing the two iterative design tasks. The survey questions asked participants to compare their experience browsing, searching, and applying examples using ProcessGallery versus using the baseline tool. P-values for questions with significant results were marked in bold.

According to our expert ratings, the degree of change between the seed design and participants' final revisions was higher in ProcessGallery ($\mu = 4.06, SD = 1.6$) than in the baseline condition ($\mu = 3.6, SD = 2$), but a mixed-effects model suggested the difference was not statistically significant.

The revised designs in both conditions had fewer issues than in the seed design, indicating the revisions improved from the seed design in both conditions. Table 2 shows the summary of the percentage of revised designs that still had issues in each principle. We noticed that when using ProcessGallery, participants who worked on Bakery scenarios were more likely to fix issues across all principles. However, the differences in the number of remaining issues between conditions were not significant, according to our regression model.

6.3.3 RQ3: Participants preferred ProcessGallery for both learning and improving a visual design. Figure 9 shows the summary of the exit survey responses from participants. According to the ratings, participants preferred ProcessGallery over the baseline tool for finding examples for addressing design issues ($M = 5.8$, $SD = 1.5$; $t(15) = 4.8$, $p = 0.002$). In particular, participants reported that ProcessGallery made it easier for them to figure out what to learn from examples ($M = 6.2$,

$SD = 1.4$; $t(15) = 6.2$, $p = 0.00002$) and apply what they learned from the examples to the given design scenario ($M = 5.47$, $SD = 1.9$; $t(15) = 2.95$, $p = 0.01$).

All but three participants anticipated that ProcessGallery would better help them learn design principles compared to the baseline gallery ($M = 6$, $SD = 1.7$; $t(15) = 4.58$, $p = 0.0004$). One participant articulated how the example pairs and their associated explanation helped her learn key design knowledge and incorporate the insight during her revision:

> *"The explanation about the example pair is really helpful [for] learning how people who know more about design would think about this. Also, seeing the 'before' is also a big strength because it kind of helps me know what to avoid or what to watch out for, and also what to do about it too."* [P3, Female]

Our participants had mixed opinions about which tool made them feel more creative ($M = 3.13$; $SD = 1.5$; $t(15) = -1.8$, $p = 0.09$, *n.s.*). In particular, the baseline gallery was deemed more effective for exploring diverse design ideas ($M = 2.27$, $SD = 1.4$; $t(15) = -4.6$, $p = 0.0003$). In the post-interviews, many participants explained that viewing examples in ProcessGallery made them focus more on the *"technical errors"* [P7, Male] in an example rather than its high-level theme or ideas. Many participants mentioned that both galleries had their own strengths. One participant commented:

> *"I think it's (ProcessGallery) mostly helpful when I am revising a design. But if I was just trying to create something from scratch, I would probably use the first interface (baseline gallery), because it is still important to be able to see how others approach the same design prompt. Or maybe combine the two galleries? So like I can first select the posters for the same topic, and within that collection, you show me how to address issues using this before-after view."* [P4, Female]

This expresses that the best configuration of a design gallery may include concepts from both ProcessGallery and the baseline gallery and let users decide when to use what features.

7 DISCUSSION & FUTURE WORK

This paper presents ProcessGallery, a novel gallery that displays a collection of visual design examples as contrasting pairs that highlight revisions tied to design principles. Study 3 found that ProcessGallery helped novice designers more accurately assess a given design and pay attention to underlying principles—rather than surface features—when they browse, filter, and apply examples, compared to a traditional gallery that only showcases final designs. Using ProcessGallery, our novice participants tended to select more examples that showcased different solutions to the issues in seed designs, whereas in the baseline, the participants picked examples that looked similar or had the same design goals (e.g., subject or purpose). ProcessGallery was strongly preferred for learning, but not for inspiration. In this section, we discuss how our findings are contextualized in the broader example-based visual learning literature. We also address the tension between learning and creativity that we identified from Study 3, and explore the context for the use of ProcessGallery. Finally, we delve into the technological support for constructing such galleries.

7.1 Example-based Learning in Example Gallery

7.1.1 Single-improvement Pairs Facilitate Principle Learning. Our Study 2 results indicated that single-improvement pairs were more effective than multi-improvement pairs for identifying underlying design issues. Although not statistically significant, participants who browsed single-improvement example pairs (in Study 3) seemed to learn better compared to browsing the same collection but with only final outcomes. When browsing single-improvement pairs in the collection, the repeated exposure to the same category directs learners' attention to what is similar (and different) among those items [19]. This effect is likely diminished when presenting examples with only final outcomes or that show multiple improvements.

In practice, learners may select more than one principle at a time, our current mechanism presents one collection after another (e.g., Emphasis, Emphasis, ..., Readability, Readability, ..., Hierarchy, Hierarchy, ...), mimicking the "blocked sequencing" in which examples from the same category are presented as a group. Prior research in educational psychology, however, has suggested that a different ordering,interleave sequencing (e.g.,Emphasis, Readability, Hierarchy, Emphasis, Readability, Hierarchy), is more effective for inductive category learning [56, 59, 69]. Future investigation should look into how the different sequencing in the gallery may affect design learning. Other future work can study whether novices can learn adequately from multi-improvement pairs if they are labeled extensively or the edits are broken down by issue.

7.1.2 Learning from Examples Versus Getting Feedback. Notably, participants improved a seed design without receiving any external feedback. Being able to inspect how others approach a similar problem can be as powerful as receiving feedback, and showing evidence of the kinds of vicarious learning experiences that Albert Bandura promotes [12]. Future studies could explore the relative impacts of personalized feedback and self-directed learning through contrasting examples on a novice's ability to learn and perform visual design.

However, the effectiveness of such a learning experience depends largely on how well a learner recognize areas for improvement in the first place. In Study 3, most participants failed to identify all the key issues, thus missing opportunities to further improve the seed design. Another gallery configuration could integrate reflection activities into the interface. Reflection is a metacognitive process through which designers assess the design situation and its alignment with project goals [97]. We hypothesize that engaging in reflection may help designers synthesize insights from examples and their design experiences. Future research could explore how and when to scaffold reflection within an example gallery to support designers' growth and development.

7.2 Tension Between Learning and Creativity

In Study 3, our participants preferred ProcessGallery for learning purposes but considered the traditional gallery more suitable for seeking creative inspiration. Our study platforms were designed to isolate the impact of specific features on creative practices. However, future work is needed to explore the design of galleries that can simultaneously promote the learning of core principles in a design domain, while also encouraging exploration to stimulate creative ideas. One approach could involve providing users with the ability to toggle between viewing iterations and final outcomes, which allow them to switch between seeing a variety of design concepts, which can be valuable for early inspiration, and viewing single-improvement pairs of examples, which can help users refine their existing designs based on principles.

7.3 Motivation and Context for Tool Use

7.3.1 Improving Design Education. While ProcessGallery was initially designed to support personalized learning that often takes place outside the classroom, we are excited about its potential to enhance design classrooms. In particular, we acknowledge that instructors heavily rely on examples to teach [46, 80, 94], but creating representative examples can be a time-consuming task. The diverse approaches showcased in ProcessGallery for addressing similar issues enhance instructors' demonstration capabilities without requiring additional effort to generate a large number of examples. We also anticipate the possibilities of collaborative teaching through the use of ProcessGallery. Consider a junior instructor with limited experience teaching a beginner design class. By browsing example pairs in ProcessGallery, they can become familiar with common mistakes students may make and integrate effective strategies observed in the gallery into their teaching materials.

The current collection of example pairs in ProcessGallery is designed around learning design principles. However, the concept of showcasing and filtering design iterations based on specific learning goals can also be extended to learn more advanced techniques. As artists increasingly share their creative process, experienced designers can also observe and learn new techniques employed by others. An interesting future work would be to explore the difference between how novices and experts capture “meaningful” iterations.

7.3.2 Supporting Collaborative Learning. We envision that a tool like ProcessGallery can provide valuable support for collaborative design tasks. For example, learners could work together to explore and discuss example pairs and to help each other extract useful insights for a shared design or individual designs. This would be a fruitful way to build on Chi and Wylie’s ICAP framework which emphasizes a move from “active” to “constructive” to “interactive” learning materials where students are encouraged to take turns articulating verbal insights that build on each other [28]. Research has shown that students learn better when they observe tutorial dialogues with their peers as opposed to watching them alone [27]. ProcessGallery can help draw attention to the key insights within examples, thereby enabling groups of students to discuss and explore the diverse interpretations and applications of complex design knowledge. In a similar vein, Cook *et al.* [30] discovered that performing a team reflection following individual reflection on peer feedback more effectively facilitated team collaboration and project iterations. This potentially suggests future work that explores patterns of individual versus collaborative efforts during different stages of creating and learning from examples.

7.3.3 Facilitating Iterative Processes. While ProcessGallery was designed to support learning, we are interested in observing how our participants apply the techniques used in the examples to their own tasks, which reflects a higher level of learning according to Bloom’s learning pyramid [60]. Based on our findings, we recommend using the tool after having a draft, as the goal would likely be seeking input to refine their concept. An open question remains: at what stage of the creative process does a tool like ProcessGallery provide the most value? Our study scenarios focused on design iterations that occurred relatively late in the process (after a complete draft). Another question is: when is the optimal time to extract examples from the creative process? For instance, would viewing iterations of paper sketches be as useful as viewing more polished designs? Or whether individuals learn better from exemplary work with minimal to no changes over time or reasonably good quality work that illustrates strategies leading to significant improvements? Future work could also study whether ProcessGallery better stimulates parallel prototyping by exposing designers to diverse ideas for improvements. Prior work has suggested that exploring various alternatives early in the process leads to better outcomes compared to fixating on a single version [34]. Exploring this aspect could provide practical guidance on how to support the practice of parallel prototyping.

7.3.4 Standalone Application or In-tool Support. Our ProcessGallery prototype is currently incorporated into the experimental platform along with the design editing tool (Polotno). In practice, the same gallery configuration can be constructed as a standalone application, similar to Behance or Dribbble. The ability to view examples (directly embedded within the design tool) while working on the task may encourage the exploration of examples. Future work can explore how placing the gallery—–either within the design tool or as a stand-alone application—–influences the use of examples.

7.4 Approaches for Constructing Improvement-based Galleries

We manufactured the examples in ProcessGallery for the purposes of this investigative study. Below we discuss potential approaches for bootstrapping gallery collections.

7.4.1 Collecting and Distilling Insights on Examples from Learning Communities. One approach is to gather examples from iterations of class assignments, where students receive feedback and make revisions guided by specific rubrics developed by instructors. For instance, if a student's iteration from one draft to the next sees a big improvement in the score in "The Use of Hierarchy" for their visual design assignment, the before-after versions can likely serve as an effective pair for learning the principle of Hierarchy. This would require instructors to capture and carefully assess each stage of student work with detailed rubrics, but this could potentially enable the construction of example collections that encompass different learning goals (principles). Looking beyond visual design, such construction of paired examples could benefit other domains such as architecture or user interface design, but would need to leverage the domain-specific or even instructor-specific principles as the structural elements.

Another approach is to empower designers in online communities to contribute example pairs by highlighting the most meaningful iterations from their creative process and leaving reflections on those pivotal moments. Prior studies have shown that deliberate reflection on past iterations leads to significant improvements in design [30, 97]. Notably, participants in all three studies expressed a desire to access designers' reasoning behind their decision-making process during the revision phase. Mosaic has made steps towards this vision by centering the design of online communities around sharing and discussing the design process rather than focusing solely on outcomes [58]. Our work provides further insights into how to deconstruct the creative process into learnable units.

7.4.2 Technological Support for Curating and Searching through Examples. Currently, ProcessGallery comprises 52 pairs of examples. However, as the collection expands, it becomes essential to curate the collection to better suit learners' specific needs. Collaborative filtering techniques, such as collecting endorsements from peers or instructors on submissions that demonstrate effective strategies for problem-solving, can be employed to curate the example pool. Advanced computer vision and computational models can also be utilized to assess similarities between examples and the target design, providing further assistance in the curation process [20, 73].

Imagine a scenario where thousands of submissions are stored in a repository with data including initial designs, feedback, and associated reflection and revisions. Consider a new user who receives feedback regarding the layout of their design and wants to find examples that address similar layout issues. With recent advances in natural language processing and information retrieval, an extension to ProcessGallery could enable the user to search the repository based on the collected feedback and reflections.

7.4.3 Integrating Example Pairs into Computational Feedback Systems. As people increasingly collaborate with AI tools that provide real-time feedback on creative work, an alternative approach could utilize AI to automatically detect meaningful iterations for learning. For instance, GUIComp integrated computational feedback within a visual design tool by calculating visual complexity scores and attention maps on the current canvas [65]. ProcessGallery can connect with tools like GUIComp to automatically collect potential example pairs that have undergone improvements, as measured by the scores as a starting point. However, this approach may still require additional human input to select the snapshots that best represent the improvements. Audience feedback and designer reflections are also necessary to explain the revisions.

7.4.4 Meaningful Iterations Versus Exemplary Work? Showing the principled-based evolution side-by-side rather than showing only the final outcome seems to redirect novices' attention from surface-level features toward the underlying principles tied to the improvements. This finding aligns with previous research that highlights the value of comparing and contrasting good (i.e., revised design) and bad examples (i.e., initial design) for effective learning [94]. One significant advantage of using before-after versions as contrasting pairs, as opposed to distinct examples, is that learners can observe the impact of the strategies employed to improve the design (e.g., witnessing how changing the font color improves information hierarchy). Future work should study the comparative effectiveness of showcasing meaningful iterations versus presenting exemplary work that demonstrates good implementation of the same principle. Finally, creating example pairs that highlight only one change would require a step-by-step breakdown of the process. How can we get designers to not only share their process, but also annotate it such that it's a useful learning resource?

Another open question pertains to the importance of the overall quality of the example pairs. Our Study 1 participants value the emphasis on improvements, as it signifies a focus on learning rather than exemplary examples associated with mastery and experience, which could potentially undermine confidence and motivation. However, if the revised design remains of low quality despite demonstrating effective strategies for improving a particular principle, it is unclear whether the pair would still be effective for learning.

8 LIMITATIONS

One limitation of our work was that the designers were provided with seed designs rather than creating their own initial design. This study design allowed us to better isolate the tool effects and to investigate how people with different abilities and experiences address the issues in the same seed design. However, people may react differently when seeing examples that are similar to a solution they "own". They might also identify different issues or disagree with the revisions represented in the contrasting pair. It would be interesting to further investigate how ProcessGallery impacts the adoption of examples when working on one's own design. Another limitation was that the participants only used the gallery for revising a work-in-progress design. We imagine the patterns of use may be different, for example, during early brainstorming stages or after a designer realizes a mostly final design. Future work could monitor if and how the strategies for finding examples change over the entire creative process.

9 CONCLUSION

Novices often struggle to identify and understand useful examples, especially when they browse online galleries such as Behance or Dribble. These platforms contain millions of examples, but provide no guidance on how to find, learn from, and apply them. To better understand how to assist learning from examples, we conducted two formative studies with novice designers and found that novices preferred accessing both the earlier and later iterations of an example rather than only seeing the final outcome (Study 1) and that isolating a single improvement across an example pair helped novices better identify insights compared to examples that include multiple revisions (Study 2). Based on these insights, we created ProcessGallery, a novel gallery interface that helps learners browse, search, and learn from collections of examples. We hope to demonstrate that designers of creative platforms and tools can not only facilitate idea exploration, but can and should also create opportunities to promote self-directed learning in visual design.

10 ACKNOWLEDGMENTS

This work was supported in part by NSF awards 2009003 and 2030859. The second was awarded to the first author as a NSF Computing Innovation Fellow. We would also like to thank our undergraduate interns who contributed—Yi-Hung Chou, May-Ann Gray Wong, Cora Xing, and Gennifer Hom. Their early exploration of the project laid a valuable foundation for its progress.

REFERENCES

[1] 2022. Amazon's Mechanical Turk. https://www.mturk.com/
[2] 2022. Behance.net. https://www.behance.net/
[3] 2022. Canva.com. https://www.canva.com/
[4] 2022. Coreldraw Design Principles. https://www.coreldraw.com/en/tips/graphic-design-principles/
[5] 2022. Design Critiques community on Reddit. https://www.reddit.com/r/design_critiques/
[6] 2022. Dribble.com. https://dribbble.com/
[7] 2022. Learning through Examples: Inductive Learning. https://www.facultyfocus.com/articles/course-design-ideas/learning-through-examples-inductive-learning/
[8] 2022. Open source for extracting dominant colors in examples: Color Thief. https://www.npmjs.com/package/colorthief
[9] 2022. Polotno Studio. https://studio.polotno.com/
[10] 2022. Qualtrics. https://www.qualtrics.com/
[11] Marine Agogué, Akin Kazakçi, Benoit Weil, and Mathieu Cassotti. 2011. The impact of examples on creative design: explaining fixation and stimulation effects. In *DS 68-2: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 2: Design Theory and Research Methodology, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011.*
[12] Albert Bandura. 1965. Vicarious processes: A case of no-trial learning. In *Advances in experimental social psychology*. Vol. 2. Elsevier, 1–55.
[13] Nathalie Bonnardel. 1999. Creativity in design activities: The role of analogies in a constrained cognitive environment. In *Proceedings of the 3rd conference on Creativity & cognition*. 158–165.
[14] Lera Boroditsky. 2007. Comparison and the development of knowledge. *Cognition* 102, 1 (2007), 118–128.
[15] Renske Bouwer, Marije Lesterhuis, Pieterjan Bonne, and Sven De Maeyer. 2018. Applying criteria to examples or learning by comparison: Effects on students' evaluative judgment and performance in writing. In *Frontiers in Education*, Vol. 3. Frontiers Media SA, 86.
[16] John D Bransford, Jeffrey J Franks, Nancy J Vye, and Robert D Sherwood. 1989. New approaches to instruction: Because wisdom can't be told. *Similarity and analogical reasoning* 470 (1989), 497.
[17] Julia Cambre, Scott Klemmer, and Chinmay Kulkarni. 2018. Juxtapeer: Comparative Peer Review Yields Higher Quality Feedback and Promotes Deeper Reflection. 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. https://doi.org/10.1145/3173574.3173868
[18] Alex Cao, Keshav K Chintamani, Abhilash K Pandya, and R Darin Ellis. 2009. NASA TLX: Software for assessing subjective mental workload. *Behavior research methods* 41 (2009), 113–117.
[19] Paulo F Carvalho and Robert L Goldstone. 2015. What you learn is more than what you see: What can sequencing effects tell us about inductive category learning? *Frontiers in psychology* 6 (2015), 505.
[20] Paulo F Carvalho and Robert L Goldstone. 2022. A computational model of context-dependent encodings during category learning. *Cognitive Science* 46, 4 (2022), e13128.
[21] Seth Chaiklin et al. 2003. The zone of proximal development in Vygotsky's analysis of learning and instruction. *Vygotsky's educational theory in cultural context* 1, 2 (2003), 39–64.
[22] Joel Chan, Joseph Chee Chang, Tom Hope, Dafna Shahaf, and Aniket Kittur. 2018. SOLVENT: A Mixed Initiative System for Finding Analogies between Research Papers. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 31 (nov 2018), 21 pages. https://doi.org/10.1145/3274300
[23] Joel Chan, Katherine Fu, Christian Schunn, Jonathan Cagan, Kristin Wood, and Kenneth Kotovsky. 2011. On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples. (2011).
[24] Joel Chan, Pao Siangliulue, Denisa Qori McDonald, Ruixue Liu, Reza Moradinezhad, Safa Aman, Erin T Solovey, Krzysztof Z Gajos, and Steven P Dow. 2017. Semantically far inspirations considered harmful? accounting for cognitive states in collaborative ideation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. 93–105.
[25] Siddhartha Chaudhuri, Evangelos Kalogerakis, Stephen Giguere, and Thomas Funkhouser. 2013. Attribit: content creation with semantic attributes. In *Proceedings of the 26th annual ACM symposium on User interface software and*

technology. 193–202.
[26] Michelene TH Chi. 2018. Learning from examples via self-explanations. In *Knowing, learning, and instruction*. Routledge, 251–282.
[27] Michelene TH Chi, Marguerite Roy, and Robert GM Hausmann. 2008. Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive science* 32, 2 (2008), 301–341.
[28] Michelene TH Chi and Ruth Wylie. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist* 49, 4 (2014), 219–243.
[29] Evangelia G Chrysikou and Robert W Weisberg. 2005. Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, 5 (2005), 1134.
[30] Amy Cook, Steven Dow, and Jessica Hammer. 2020. Designing interactive scaffolds to encourage reflection on peer feedback. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1143–1153.
[31] Nicola Dell, Vidya Vaidyanathan, Indrani Medhi, Edward Cutrell, and William Thies. 2012. " Yours is better!" participant response bias in HCI. In *Proceedings of the sigchi conference on human factors in computing systems*. 1321–1330.
[32] Donis A Dondis. 1974. *A primer of visual literacy*. Mit Press.
[33] Kees Dorst and Judith Dijkhuis. 1995. Comparing paradigms for describing design activity. *Design studies* 16, 2 (1995), 261–274.
[34] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 4 (2010), 1–24.
[35] 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* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376635
[36] Jane L. E, Kevin Y. Zhai, Jose Echevarria, Ohad Fried, Pat Hanrahan, and James A. Landay. 2021. Dynamic Guidance for Decluttering Photographic Compositions. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 359–371. https://doi.org/10.1145/3472749.3474755
[37] Leon Festinger. 1954. A theory of social comparison processes. *Human relations* 7, 2 (1954), 117–140.
[38] Eureka Foong, Steven P. Dow, Brian P. Bailey, and Elizabeth M. Gerber. 2017. Online Feedback Exchange: A Framework for Understanding the Socio-Psychological Factors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4454–4467. https://doi.org/10.1145/3025453.3025791
[39] Katherine Fu, Joel Chan, Jonathan Cagan, Kenneth Kotovsky, Christian Schunn, and Kristin Wood. 2013. The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design* 135, 2 (2013).
[40] Dedre Gentner, Keith J Holyoak, and Boicho N Kokinov. 2001. *The analogical mind: Perspectives from cognitive science*. MIT press.
[41] Dedre Gentner, Jeffrey Loewenstein, and Leigh Thompson. 2003. Learning and transfer: A general role for analogical encoding. *Journal of educational psychology* 95, 2 (2003), 393.
[42] Dedre Gentner and Arthur B Markman. 1997. Structure mapping in analogy and similarity. *American psychologist* 52, 1 (1997), 45.
[43] Mary L Gick and Keith J Holyoak. 1980. Analogical problem solving. *Cognitive psychology* 12, 3 (1980), 306–355.
[44] Mary L Gick and Keith J Holyoak. 1983. Schema induction and analogical transfer. *Cognitive psychology* 15, 1 (1983), 1–38.
[45] Mary L Gick and Katherine Paterson. 1992. Do contrasting examples facilitate schema acquisition and analogical transfer? *Canadian Journal of Psychology/Revue canadienne de psychologie* 46, 4 (1992), 539.
[46] Elena L. Glassman, Jeremy Scott, Rishabh Singh, Philip J. Guo, and Robert C. Miller. 2015. OverCode: Visualizing Variation in Student Solutions to Programming Problems at Scale. *ACM Trans. Comput.-Hum. Interact.* 22, 2, Article 7 (mar 2015), 35 pages. https://doi.org/10.1145/2699751
[47] Rebecca A Grier. 2015. How high is high? A meta-analysis of NASA-TLX global workload scores. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 59. SAGE Publications Sage CA: Los Angeles, CA, 1727–1731.
[48] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.
[49] Björn Hartmann, Scott Doorley, and Scott R Klemmer. 2006. Hacking, mashing, gluing: a study of opportunistic design and development. *Pervasive Computing* 7, 3 (2006), 46–54.

[50] Marti Hearst. 2009. *Search user interfaces*. Cambridge university press.

[51] Scarlett R. Herring, Chia-Chen Chang, Jesse Krantzler, and Brian P. Bailey. 2009. Getting Inspired! Understanding How and Why Examples Are Used in Creative Design Practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 87–96. https://doi.org/10.1145/1518701.1518717

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

[53] Elahe Javadi. 2012. *The impact of user interface design on idea integration in electronic brainstorming: an attention-based view*. Ph. D. Dissertation. University of Illinois at Urbana-Champaign.

[54] Hyeonsu B Kang, Gabriel Amoako, Neil Sengupta, and Steven P Dow. 2018. Paragon: An online gallery for enhancing design feedback with visual examples. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[55] Hyeonsu B. Kang, Xin Qian, Tom Hope, Dafna Shahaf, Joel Chan, and Aniket Kittur. 2022. Augmenting Scientific Creativity with an Analogical Search Engine. *ACM Trans. Comput.-Hum. Interact.* (mar 2022). https://doi.org/10.1145/3530013 Just Accepted.

[56] Sean HK Kang and Harold Pashler. 2012. Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology* 26, 1 (2012), 97–103.

[57] Rubaiat Habib Kazi, Tovi Grossman, Hyunmin Cheong, Ali Hashemi, and George W Fitzmaurice. 2017. DreamSketch: Early Stage 3D Design Explorations with Sketching and Generative Design.. In *UIST*, Vol. 14. 401–414.

[58] Joy Kim, Maneesh Agrawala, and Michael S. Bernstein. 2017. Mosaic: Designing Online Creative Communities for Sharing Works-in-Progress. 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, 246–258. https://doi.org/10.1145/2998181.2998195

[59] Nate Kornell and Robert A Bjork. 2008. Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological science* 19, 6 (2008), 585–592.

[60] David R Krathwohl. 2002. A revision of Bloom's taxonomy: An overview. *Theory into practice* 41, 4 (2002), 212–218.

[61] Chinmay Kulkarni, Steven P Dow, and Scott R Klemmer. 2014. Early and repeated exposure to examples improves creative work. In *Design thinking research*. Springer, 49–62.

[62] Ranjitha Kumar, Jerry O. Talton, Salman Ahmad, and Scott R. Klemmer. 2011. Bricolage: Example-Based Retargeting for Web Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2197–2206. https://doi.org/10.1145/1978942.1979262

[63] Brian Lee, Savil Srivastava, Ranjitha Kumar, Ronen Brafman, and Scott R. Klemmer. 2010. Designing with Interactive Example Galleries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2257–2266. https://doi.org/10.1145/1753326.1753667

[64] 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. https://doi.org/10.1145/3313831.3376327

[65] 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*. 1–13. https://doi.org/10.1145/3313831.3376327

[66] Ellen Lupton. 2014. *Thinking with type: A critical guide for designers, writers, editors, & students*. Chronicle Books.

[67] Richard L Marsh, Joshua D Landau, and Jason L Hicks. 1996. How examples may (and may not) constrain creativity. *Memory & cognition* 24, 5 (1996), 669–680.

[68] Scarlett R Miller and Brian P Bailey. 2014. Searching for inspiration: An in-depth look at designers example finding practices. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 46407. American Society of Mechanical Engineers, V007T07A035.

[69] Matthew E Mundy, Robert C Honey, and Dominic M Dwyer. 2009. Short Article: Superior Discrimination between Similar Stimuli after Simultaneous Exposure. *Quarterly Journal of Experimental Psychology* 62, 1 (2009), 18–25.

[70] Tricia J Ngoon, Joy O Kim, and Scott Klemmer. 2021. Shöwn: Adaptive Conceptual Guidance Aids Example Use in Creative Tasks. In *Designing Interactive Systems Conference 2021*. 1834–1845.

[71] 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. https://doi.org/10.1145/2998181.2998319

[72] Amin Heyrani Nobari, Muhammad Fathy Rashad, and Faez Ahmed. 2021. Creativegan: Editing generative adversarial networks for creative design synthesis. *arXiv preprint arXiv:2103.06242* (2021).
[73] Robert M Nosofsky. 2011. The generalized context model: An exemplar model of classification. *Formal approaches in categorization* (2011), 18–39.
[74] Laura R Novick. 1988. Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, memory, and cognition* 14, 3 (1988), 510.
[75] 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*. 1221–1224.
[76] A Terry Purcell and John S Gero. 1991. The effects of examples on the results of a design activity. In *Artificial Intelligence in Design'91*. Elsevier, 525–542.
[77] Lyle Regenwetter, Brent Curry, and Faez Ahmed. 2022. BIKED: A dataset for computational bicycle design with machine learning benchmarks. *Journal of Mechanical Design* 144, 3 (2022).
[78] Todd Rogers and Avi Feller. 2016. Discouraged by peer excellence: Exposure to exemplary peer performance causes quitting. *Psychological science* 27, 3 (2016), 365–374.
[79] D Royce Sadler. 1989. Formative assessment and the design of instructional systems. *Instructional science* 18, 2 (1989), 119–144.
[80] Donald A Schön. 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* Jossey-Bass.
[81] 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.
[82] Daniel L Schwartz, Jessica M Tsang, and Kristen P Blair. 2016. *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them.* WW Norton & Company.
[83] Jonathan T Shemwell, Catherine C Chase, and Daniel L Schwartz. 2015. Seeking the general explanation: A test of inductive activities for learning and transfer. *Journal of Research in Science Teaching* 52, 1 (2015), 58–83.
[84] Pao Siangliulue, Joel Chan, Steven P. Dow, and Krzysztof Z. Gajos. 2016. IdeaHound: Improving Large-Scale Collaborative Ideation with Crowd-Powered Real-Time Semantic Modeling. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 609–624. https://doi.org/10.1145/2984511.2984578
[85] Pao Siangliulue, Joel Chan, Krzysztof Z. Gajos, and Steven P. Dow. 2015. Providing Timely Examples Improves the Quantity and Quality of Generated Ideas. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (Glasgow, United Kingdom) *(C&C '15)*. Association for Computing Machinery, New York, NY, USA, 83–92. https://doi.org/10.1145/2757226.2757230
[86] Pooja G Sidney, Shanta Hattikudur, and Martha W Alibali. 2015. How do contrasting cases and self-explanation promote learning? Evidence from fraction division. *Learning and Instruction* 40 (2015), 29–38.
[87] Steven M Smith, Thomas B Ward, and Jay S Schumacher. 1993. Constraining effects of examples in a creative generation task. *Memory & cognition* 21, 6 (1993), 837–845.
[88] Charles Spearman. 1923. *The nature of" intelligence" and the principles of cognition.* Macmillan.
[89] Claude M Steele. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in experimental social psychology*. Vol. 21. Elsevier, 261–302.
[90] Masaki Suwa and Barbara Tversky. 2002. External representations contribute to the dynamic construction of ideas. In *Diagrammatic Representation and Inference: Second International Conference, Diagrams 2002 Callaway Gardens, GA, USA, April 18–20, 2002 Proceedings 2*. Springer, 341–343. https://doi.org/10.1007/3-540-46037-3_33
[91] David R Thomas. 2003. A general inductive approach for qualitative data analysis. (2003).
[92] Leigh Thompson, Dedre Gentner, and Jeffrey Loewenstein. 2000. Avoiding missed opportunities in managerial life: Analogical training more powerful than individual case training. *Organizational Behavior and Human Decision Processes* 82, 1 (2000), 60–75.
[93] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. Getting the right design and the design right. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 1243–1252.
[94] Joan Walker and Paul King. 2006. Contrasting cases: What we can learn from students' perceptions of "good" design. (2006).
[95] Robin Williams. 2015. *The non-designer's design book: Design and typographic principles for the visual novice.* Pearson Education.
[96] Yu-Chun Grace Yen and Steven P. Dow. 2022. Seeking Exemplars in the Wild: Exploring How Students Find Design Examples to Support Personalized Learning. In *Proceedings of the Ninth ACM Conference on Learning @ Scale* (New York City, NY, USA) *(L@S '22)*. Association for Computing Machinery, New York, NY, USA, 418–421. https://doi.org/10.1145/3491140.3528303

[97] Yu-Chun Grace Yen, Steven P Dow, Elizabeth Gerber, and Brian P Bailey. 2017. Listen to others, listen to yourself: Combining feedback review and reflection to improve iterative design. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. 158–170.

[98] Yu-Chun Grace Yen, Joy O Kim, and Brian P Bailey. 2020. Decipher: an interactive visualization tool for interpreting unstructured design feedback from multiple providers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[99] Amy X Zhang and Justin Cranshaw. 2018. Making sense of group chat through collaborative tagging and summarization. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–27.

A DESIGN PRINCIPLES

Principles	Violations	Issues
Hierarchy (H)	H.1 Weak Point of Entry	H.1.1 Title does not stand out.
		H.1.2 Key image does not stand out.
	H.2 Ambiguous Levels of Importance	H.2.1 Unclear relative importance of design elements (text, images, etc.).
		H.2.2 Unclear grouping of content.
		H.2.3 Illogical prioritization of design elements.
Alignment (A)	A.1 Arbitrary Alignment of Elements	A.1.1 Lack of text alignment.
		A.1.2 Mismatch alignment between text and image.
		A.1.3 Inappropriate background image cropping.
		A.1.4 Lack of alignment between graphics (e.g., images, icons, etc).
	A.2 Insufficient Margins	A.2.1 Content too close to the edge (no margins).
Balance (B)	B.1 Not Enough Space Between Content	B.1.1 Unclear grouping or spatial arrangement of content.
	B.2 Unnecessary Open Areas	B.2.1 Disjoint elements (e.g., too much unused space).
	B.3 Uneven Margins	B.3.1 Content feels unintentionally skewed to one side.
Unity (U)	U.1 Too Many Variations in Text	U.1.1 Using too many different typefaces.
		U.1.2 Overusing content emphasis (bold, italic, underline etc.).
	U.2 Unnecessary Design Elements	U.2.1 Noisy imagery or too many images.
		U.2.2 Too many lines, graphs, and other visual elements.
	U.3 Inconsistent Color Choices	U.3.1 Too any variations in text color.
		U.3.2 Too many variations in color.
		U.3.3 Inconsistent color palettes.
	U.4 Incohesive Graphic Choices	U.4.1 Inconsistent art styles.
Readability (R)	R.1 Poor Text Legibility	R.1.1 Inappropriate text size.
		R.1.2 Inappropriate line length.
		R.1.3 Distracting font effects or font.
		R.1.4 Unclear text treatments (kerning, leading, etc.).
		R.1.5 Ineffective text direction.
		R.1.6 Inappropriate line-breaks (widows, orphans).
		R.1.7 Inaccessible typefaces.
	R.2 Unsuitable Image Manipulation	R.2.1 Low resolution/blurry image.
		R.2.2 Image subject matter unclear.
		R.2.3 Low contrast image (image unclear).
		R.2.4 Inappropriate image size.
		R.2.5 Image has distracting border or background.
		R.2.6 Distorted/warped image.
	R.3 Content Obscured	R.3.1 Low contrast between text and background.
		R.3.2 Inappropriate occlusion of imagery (text over image, image over image).
		R.3.3 Inappropriate occlusion of text (image over text, text over text).

Table 3. Common issues made by novice designers, organized by the five principles included in the knowledge assessment tests and ProcessGallery interface.

B SUMMARY OF THE FILTER USAGE

Dance Workshop			
ProcessGallery	**Count**	**Baseline**	**Count**
Weak Point of Entry	6	Performing Arts	6
Inconsistent/Too Many Variations in Text	5	Public Event Ad	6
Poor Text Legibility	3	Photo	5
Insufficient Margins	2	Mixed	2
Ambiguous Levels of Importance	2	Typographic	2
Uneven Margins	2	Illustration And Vector	2
Unnecessary Open Areas	2	Visual And Literary Arts	2
Inconsistent/Too Many Variations in Text	2	Festivals And Fairs	2
Viewed Posters	1	Black	2
N/A	1	School And Education	1
Unsuitable Image Manipulation	1	Informational	1
Arbitrary Alignment of Elements	1	Recruitment	1
Not Enough Space Between Content	1	Private Event Ad	1
Incohesive Graphic Choices	1	Nonprofit Event	1
Inconsistent Color Choices	1	Fashion And Beauty	1
		Food	1
		Health And Medicine	1
		Sports	1
		Red	1
		Purple	1
		N/A	1
Bakery Flyer			
ProcessGallery	**Count**	**Baseline**	**Count**
Arbitrary Alignment of Elements	6	Food	4
Incohesive Graphic Choices	5	Fundraising And Charity Ad	3
Unnecessary Design Elements	5	Illustration And Vector	3
Insufficient Margins	4	Product And Services Ad	2
Uneven Margins	4	Typographic	2
Not Enough Space Between Content	3	Mixed	2
Ambiguous Levels of Importance	3	Pink	2
Poor Text Legibility	3	Photo	2
Content Obscured	2	N/A	1
Unsuitable Image Manipulation	2	N/A	1
Inconsistent/Too Many Variations in Text	2	Nonprofit Event	1
Unnecessary Open Areas	2	Graphic-Mixed	1
Weak Point of Entry	2	Orange	1
Viewed Posters	1	Sale	1
Inconsistent Color Choices	1	Recuitment	1
		Red	1
		Parties And Celebration	1
		Purple	1
		Public Event Ad	1
		Private Event Ad	1

Table 4. Dance Workshop and Bakery Flyer Filter Usage Counts

Received January 2023; revised July 2023; accepted November 2023