research statement | jane l. e

Computer scientists have long worked towards the vision of human-AI collaboration for augmenting human capabilities and intellect. Modern-day AI largely focuses on improving the combined capabilities of humans and machines through aiding users in performing tasks better and faster. This has shown to be incredibly powerful. For example, users can leverage translation tools to communicate in a foreign language without needing to learn it first. However, my work argues that this can overlook an important opportunity to help the user grow—specifically, for **computational tools to augment human intellect through helping them develop domain expertise for a given task**. Compare the first interaction to being fluent in the language and able to determine if nuances of a message are being communicated accurately. To support users in developing domain expertise, **I design tools that interactively guide users through key intermediate steps of a process**, rather than tools that immediately jump to the end of the process. To do so, my research leverages algorithmic techniques (e.g., computer graphics, computer vision, and artificial intelligence) to realize novel design interactions that consider insights from education and cognitive science.

A COMPUTATIONAL APPROACH TO SUPPORTING ARTISTIC VISION

Specifically, I leverage the insight that experts "see" in a different way [1]. Expert artists, for instance, spend years (if not decades) training their "artist's eye" [2] to perceive in ways that embed their expert domain knowledge—in this case, core artistic concepts. I call this "artistic vision". For example, a painter will look at a scene and instantly see background details to simplify to better focus their painting on the core subject, whereas a novice might see the whole scene in equal detail, and paint it as such. Combining technical and design methods, I introduce **novice-interpretable, real-time guidance that embeds expertise of important artistic concepts to scaffold novices' design processes**. By bringing this feedback directly into the design process, we can help novices train their own artistic vision (Figure 1).

We take inspiration from education research and traditional pedagogical or apprenticeship approaches for art practice to explore how computational tools can support developing artistic vision. My work explores 3 approaches—**(1) providing contextual exemplars as inspiration** [3, 4]: instructors often first introduce concepts to novices through related expert examples, **(2) supporting awareness of domain concepts through visual annotations** [5, 6]: instructors will often annotate directly on the students' work to help them consider principles, and **(3) encouraging reflection through principled feedback** [7]: critiques are core to such classes; instructors and peers will give feedback based on their interpretation of the work, identifying potential violations of core principles.

(1) Providing Contextual Exemplars: Concept-based Galleries Provide Actionable Guidance

Studio photographers control the distribution of bright and dark regions on their subject through careful light placement. In natural environments, photographers approximate this by rotating the subject. Humans' extraordinary visual system enables us to see 'through' lighting when we're there—while looking at the physical scene, it is easy to miss even distracting shadows on a person's face. The bad lighting is often only apparent later in the captured image. To give aspiring photographers artistic vision, we introduce a capture-time lighting awareness technique (Figure 2) [3]. Users select a target appearance from a gallery of lighting styles. Our technique computes the optimal subject orientation for the target style by capturing lighting through an HDR environment map with a 360 camera. Guided by an efficient, precomputed radiance transfer approach, it then tells the photographer how far to rotate the subject. Technical evaluations showed that our technique is robust across several indoor and outdoor scenes—using many different subjects—to achieve a variety of looks. User evaluations showed that the tool helps novices produce well-lit portraits by reducing cognitive load. We saw evidence of participants shifting their perspective and developing confidence: "*it made me think beyond just the content of*

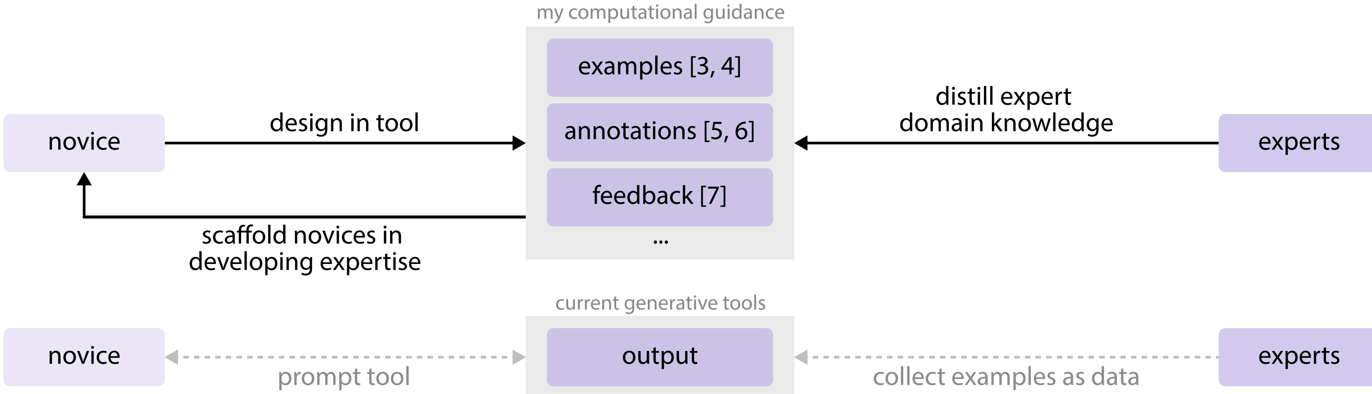

Figure 1. High-level framework of the components of my computational tools (top) for supporting developing expertise. *Computational guidance* (middle) embeds the *domain knowledge distilled from experts* (right). *Novices develop expertise scaffolded* by the guidance provided based on their designs (left). This contrasts with the common existing pattern (bottom, in gray) of connecting novices with experts through *generated output*.

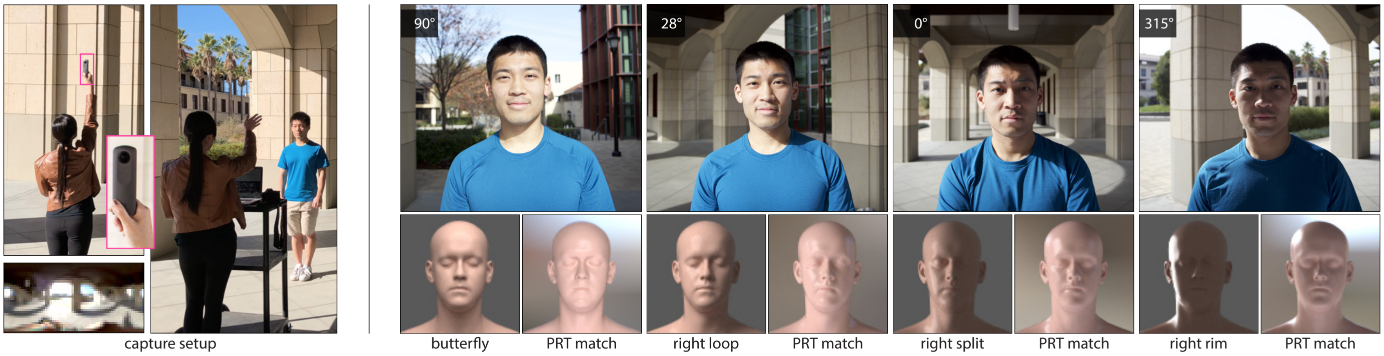

Figure 2. In a fixed lighting environment, photographers can produce many different lighting styles just by rotating the subject in place. The *capture setup* (left) shows a photographer capturing an HDR environment map and directing the subject to reorient based on our reorientation guidance. On the right are results for several target lighting styles in our contextual lighting gallery (*butterfly*, *right loop*, *right split*, and *right rim*)—including the final photo, the target style, and the preview of the closest *match* at the optimal angle as determined by our precomputed radiance transfer-based approach.

the photo, and also pay attention to where the light sources are... makes me feel like I can take much more dramatic and varied photos in a limited space." This additionally demonstrates the creative ownership that can come with developing expertise.

We also explore how galleries of visual design examples can be structured to better support novice designers. Expert designers know to evaluate the communication effectiveness of their designs based on high-level design principles, such as hierarchy or readability. Typical design galleries, instead, tend to be based on surface attributes (e.g., color or style). Observing that novice designers can develop artistic vision through examining others' design processes, we introduce 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 [4]. Our insight is that we can help novices develop artistic vision by learning from how prior designers fixed issues in their designs. Participants significantly preferred ProcessGallery for learning. They were able to find more relevant examples and were better at assessing designs based on design principles.

Examples are crucial to learning in almost any domain, and are especially helpful as inspiration in creative domains. These projects build on the insight that being able to interact with design galleries [8] based on different conceptual dimensions can further their effectiveness. We note that instructors do a lot of work to curate such examples that best illustrate important concepts; and therefore aim to embed such context into our galleries. In the photography lighting project, the context considered is the light in the environment, and in the second project with paired examples, we consider the context of how prior designers fixed common design issues relative to underlying principles.

(2) Supporting Awareness: Dynamic Annotations Encourage Conceptual Awareness

Photographic composition is often taught as alignment with composition grids—most commonly, the rule of thirds. Professionals create dynamic compositions through more complex grids, like the harmonic armature. Professionals see and select appropriate grids in their mind's eye. Introducing complex grids to novices can be overwhelming—experts often selectively highlight lines per image step-by-step, while explaining the relevant lines to the composition. To give novices artistic vision, we introduce a saliency-based algorithm that adaptively highlights relevant lines for the current scene and composition (Figure 4) [5]. Participants found our adaptive armatures helpful in capturing more well-composed images.

A third example of photographers' artistic vision is being able to see beyond an image's primary focus to identify clutter and other distractions that take away from the photograph. As with grid composition, novice photographers' attentional

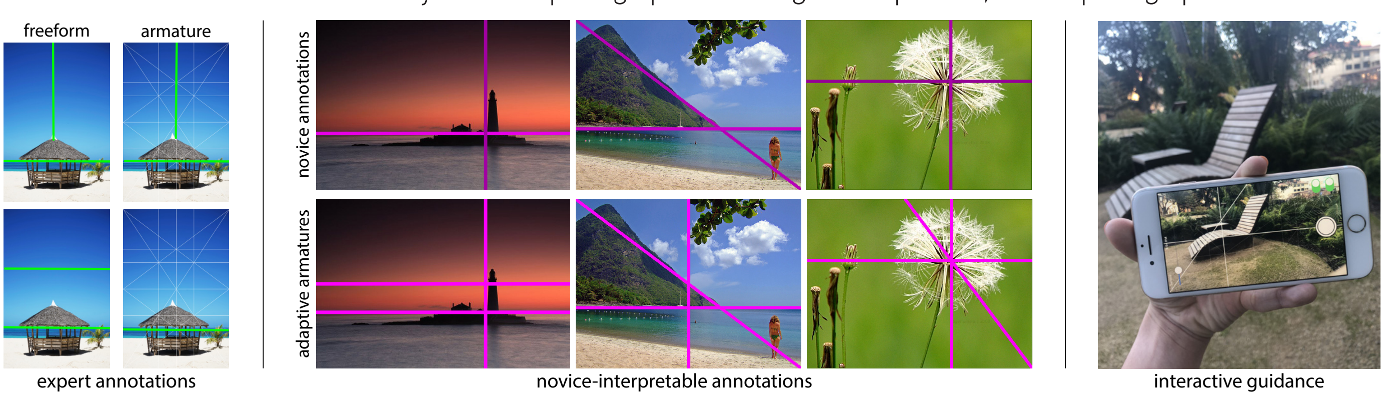

Figure 3. This illustrates the design process of our composition guidance. At the left are *expert annotations*—we asked 9 expert photographers to annotate a set of photos freeform and with the harmonic armature overlaid. Next, we wanted to understand if crowdsources *novices* (middle top) were able to similarly annotate images using the armature in a meaningful way. Based on the results, we designed our *adaptive armatures* (middle bottom) to select lines that intersected salient regions of an image. These adaptive armatures are provided as *interactive guidance* in the camera at capture time.

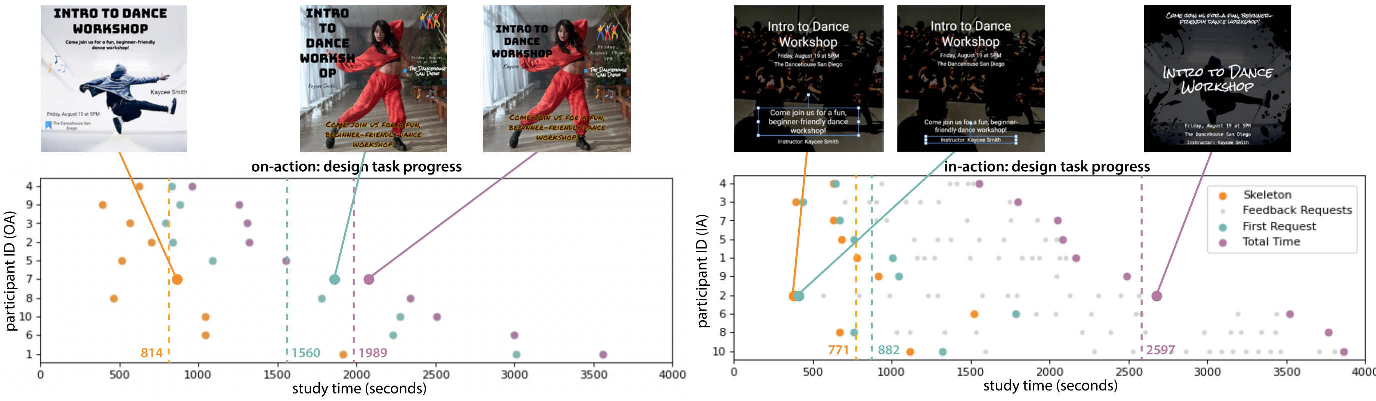

Figure 4. Timing of participants' feedback interaction in the *on-action condition* (left) and *in-action condition* (right). This shows when they had an *initial skeleton* (orange), made their *first feedback request* (teal), made further *feedback requests* (gray), and completed their *final design* (purple). Above we show a corresponding set of designs for two participants. Notice the degree of change is drastically higher before the first request on the left, and after on the right.

spotlight often misses clutter—until it's too late. Here, our work promotes artistic vision through a visual abstraction that makes clutter easier to see—while the novice is still in the moment trying to frame the scene [6]. We introduce an object-based saliency and edge detection technique to highlight contrast along subject and image borders, outlining potential distractors. Our capture-time tool interactively displays these overlays in the camera. An evaluation found that our overlays boosted participants' confidence in taking photographs that convey their story without distracting clutter.

This awareness approach builds on the insight that external representations, such as architecture sketches, can promote iteration through new interpretations and discoveries [9]. Our work introduces dynamic annotations as visual overlays that aim to embed expert knowledge to increase novice awareness of underlying concepts without being (too) prescriptive.

(3) Encouraging Reflection: Principled Computational Feedback Improves Performance But Risks Overreliance

Advances in AI have opened up the potential for creativity tools to computationally generate design feedback on-demand. However in traditional classroom settings, instructors not only have control over the type of feedback they provide, but also the timing of the feedback—has the student had enough time to self-reflect and struggle with the problem on their own? This often occurs at a slower pace than if controlled by the student. In a future where student designers have constant access to feedback, how would the timing of these requests impact their creative learning processes? To start understanding the tradeoffs between the potential future of on-demand feedback versus the slower-paced classroom practice of having feedback after completed drafts, we designed an interactive design probe varying the timing at which participants have access to (wizard-provided) feedback (Figure 4) [7]. Here, we extended guidance to feedback that directly provided interpretations based on design principles—where are the issue areas and why do they violate the principle.

We found that when given access, novice designers do tend to frequently request feedback. They noted feedback reminded them to consider other principles when hyper-focused on a single issue. While having access to feedback resulted in better performance (reducing more design issues), this was at the cost of novices feeling like they were overly relying on feedback instead of engaging in more holistic self-evaluation. On the other hand, when requiring a draft prior to feedback, participants were more reluctant to address feedback (as predicted by the concept of design fixation). This highlights the importance of finding the right tradeoffs between these timing patterns in order to balance ownership and learning with efficiency and performance—how can we provide novices with the necessary perspectives to allow them to sit with the discomfort and trust their own instincts and self-reflection?

FUTURE RESEARCH AGENDA

I envision a future where computational tools can complement traditional pedagogy to support people in developing expertise in domains they are excited to pursue. My interdisciplinary background prepares me to collaborate with system builders, algorithm designers, learning researchers, social scientists, and designers towards this goal. Based on my research framework (Figure 1), I group my future work into 3 high-level directions: (1) understanding how to **scaffold novices** for developing expertise, (2) algorithmic components for **computational guidance**, and (3) **distilling expertise** in a domain.

Providing Visual Explanations to Reduce Overreliance on Computational Feedback. As mentioned above, a core research question my ongoing work [7] has highlighted is the need to understand how to balance tradeoffs between benefits of earlier feedback (enabled by computational tools) and the potential downsides of overreliance, such as a reduced sense of ownership. Taking inspiration from prior work showing that explanations can help reduce overreliance on AI [10], I've taken initial steps towards trying to reduce overreliance via designing visual annotations as a source of explanation for novices to ground their self-reflection. To do so, we integrate insights from our awareness-focused annotations [5, 6]—by providing visual explanations of underlying principles, will novices have more confidence in their own abilities to interpret design quality rather than overly trusting computational feedback?

Integrating Generative AI Interactions into Design Process. Many recent tools in generative AI operate through a pipeline where a high-fidelity output is generated (e.g., text-to-image, wireframe-to-interface). These pipelines can save a lot of tedious work. They also have natural alignment with several aspects of the design such as brainstorming, rapid prototyping, or parallel prototyping. However, they skip many other intermediate stages of the design process. Can we adapt these tools to better support more steps of the design process such as moodboarding, iterative design, or low-fidelity prototyping? For instance, what if we instead envisioned pipelines that abstract high-fidelity input to low-fidelity interpretations? What would those look like and would they be helpful as inspiration or guidance for aspiring designers?

computatoinal guidance

Extending External Representations to Non-Visual Domains. Cognitive science emphasizes that the strength of external representations (such as the visual annotations in my work) comes significantly from their ability to highlight spatial relationships [9]. How might we extend some of these benefits beyond visual domains? For instance, to domains without a visual component (e.g., coming up with new research ideas), or to domains that have non-visual constraints (e.g., furniture that needs to support weight). Additionally, how might we extend spatial understanding beyond visuals? This could potentially mean instead leveraging other modalities—such as using audio, or converting them to text. For artistic domains, this could mean providing image captions that not only describe content, but also artistic aspects of the image.

Supporting Collaboration Across Mixed Expertise. Interdisciplinary teams include members across varying expertise: architecture, engineering, art, etc. Individual team members are experts in their own domain, but novices in each others'—can computational tools help to support communicating across their different domain-specific knowledge and language to collaboratively pursue higher level intentions? My work starts to explore visualizing parallels in computer algorithms (e.g., saliency) and artistic principles (e.g., composition); can we further support externalization of these parallels across expertise to support teamwork? For instance, many "aha moments" require being in collocated spaces, e.g., the architect observes the electrical engineer debugging, but actually likes the current "buggy" result. Can we better support these moments across time and space?

distill expertise

Developing Techniques for Capturing Experts' Tacit Knowledge. Most methods for capturing (and especially representing) expert knowledge have been demonstrated as bespoke approaches that rely on experts predetermining methods to communicate their knowledge to novices. HCI leverages interview methods like contextual inquiry to discover insights on practices that are so second-nature to experts that they themselves might not notice them. Researchers have also used gaze paths to represent radiologists' vision with more accuracy than they could describe [11]. How can we develop generalized methods for capturing expert knowledge, especially tacit knowledge that experts themselves might not know how to communicate? Additionally, these representations consider "experts" as a population as a whole, just as machine learning models capture domain information across all humans. How might we capture the necessary information to be able to modify representations for differing "lenses" for experts with differing preferences and perspectives?

REFERENCES

[1] Charles Goodwin. Professional Vision. *American Anthropologist 1994*.
[2] Michael F. Marmor and James G. Ravin. The Artist's Eyes: Vision and the History of Art. *Abrams 2009*.
[3] **Jane L. E**, Ohad Fried, and Maneesh Agrawala. Optimizing Portrait Lighting at Capture-Time Using a 360 Camera as a Light Probe. *UIST 2019*.
[4] **Jane L. E***, Grace Yen*, Hyoungwook Jin, Mingyi Li, Grace Lin, Isabelle Pan, and Steven P. Dow. ProcessGallery: Illustrating Design Principles in Online Example Galleries through Early and Late Iterations of Visual Designs. *To Appear CSCW 2024*.
[5] **Jane L. E**, Ohad Fried, Jingwan Liu, Jianming Zhang, Radomír Měch, Jose Echevarria, Pat Hanrahan, and James A. Landay. Adaptive Photographic Composition Guidance. *CHI 2020*.
[6] **Jane L. E**, Kevin Y. Zhai, Jose Echevarria, Ohad Fried, Pat Hanrahan, and James A. Landay. Dynamic Guidance for Decluttering Photographic Compositions. *UIST 2021*.
[7] **Jane L. E***, Grace Yen*, Isabelle Pan, Grace Lin, Mingyi Li, Hyoungwook Jin, Haijun Xia, and Steven P. Dow. When to Give Feedback? Exploring In-Action Feedback in a Visual Design Tool. *In Preparation*.
[8] Brian Lee, Savil Srivastava, Ranjitha Kumar, Ronen Brafman, and Scott R. Klemmer. Designing with Interactive Example Galleries." *CHI 2010*.
[9] Masaki Suwa and Barbara Tversky. External Representations Contribute to the Dynamic Construction of Ideas. *Diagrammatic Representation and Inference 2002*.
[10] Helena Vasconcelos, Matthew Jörke, Madeleine Grunde-McLaughlin, Tobias Gerstenberg, Michael S. Bernstein, and Ranjay Krishna. Explanations Can Reduce Overreliance on AI Systems During Decision-Making. *CSCW 2023*.
[11] Srinivas Sridharan, Reynold Bailey, Ann McNamara, and Cindy Grimm. Subtle Gaze Manipulation for Improved Mammography Training. *Eye Tracking Research and Applications 2012*.

*Authors contributed equally