VisiBlends: A Collaborative Design Tool for Visual Blends

Lydia B. Chilton
Columbia University
chilton@cs.columbia.edu

Savvas Petridis
Columbia University
sdp2137@columbia.edu

Maneesh Agrawala
Stanford University
maneesh@cs.stanford.edu

ABSTRACT
Visual blends are an advanced graphic design technique to draw attention to a message. They combine two objects in a way that is novel and useful in conveying a message symbolically. This paper presents VisiBlends, a system for creating visual blends that follows the iterative design process. We introduce a design pattern for blending symbols based on principles of the human visual object recognition system. Our pipeline decomposes the process into both computational techniques and human microtasks. It allows users to collaboratively generate visual blends with steps involving brainstorming, synthesis, and iteration. An evaluation of the pipeline shows that users can generate blends through independent microtasks, groups can collaboratively make visual blends for their own messages, and VisiBlends improves novices' ability to make visual blends.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Design, collaboration, workflow, visual blends.

INTRODUCTION
Visual blends are an advanced graphic design technique used in journalism, advertising and public service announcements to draw users' attention to a message. They blend two objects together in a way that is novel and useful in conveying a message symbolically. For example, in Figure 1, the Starbucks logo is blended with a sun to convey that "Starbucks is here for summer." Visual blends are widely considered to be creative, and many of the top image search results for "creative ads" are visual blends. (See Figure 2.)

Novices have never heard of visual blends and thus do not consider them as a way to attract attention to their news and announcements. Even when introduced to the concept of visual blends and seeing professional examples like those in Figure 2, it is still difficult to make one because there are two opposing goals: combining two objects into one while ensuring both objects are still recognizable. There are no

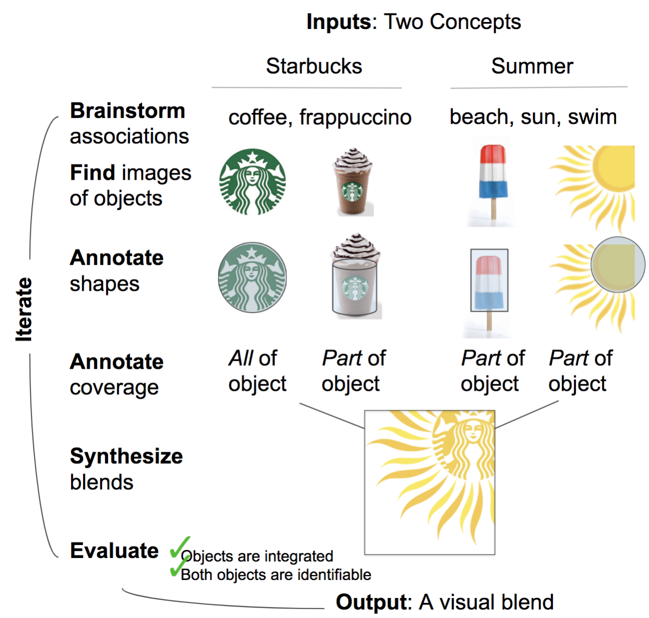

Figure 1: An example of the VisiBlends pipeline used to create a visual blend for the concepts *Starbucks* and *summer*.

obvious characteristics that visual blends share. They all use different objects and combine them in unique ways. It seems that every one requires creative inspiration and that there is no exact procedure.

To enable novices to make visual blends, we decompose the process of constructing them. Although there is no obvious *surface-level structure* to visual blends, there is a common *abstract structure* to many visual blends: they combine two objects with a similar shape. For example, in Figure 1 the Starbucks logo and the sun are both circles and in Figure 2, the Tabasco Bottle and fire extinguisher are both cylinders. In the design literature, abstract structure used to guide the design process is referred to as a design pattern [4].

We present a system that decomposes the process of making visual blends into a pipeline of human microtasks and computational techniques. The pipeline follows the iterative design process with steps involving brainstorming, synthesis, and iteration. The input to the system is two concepts such as *Starbucks* and *summer*. First, the system finds many possible symbols for both concepts by having people brainstorm related concepts and then find images for them. Next, we want to find pairs of images that can be blended. Because

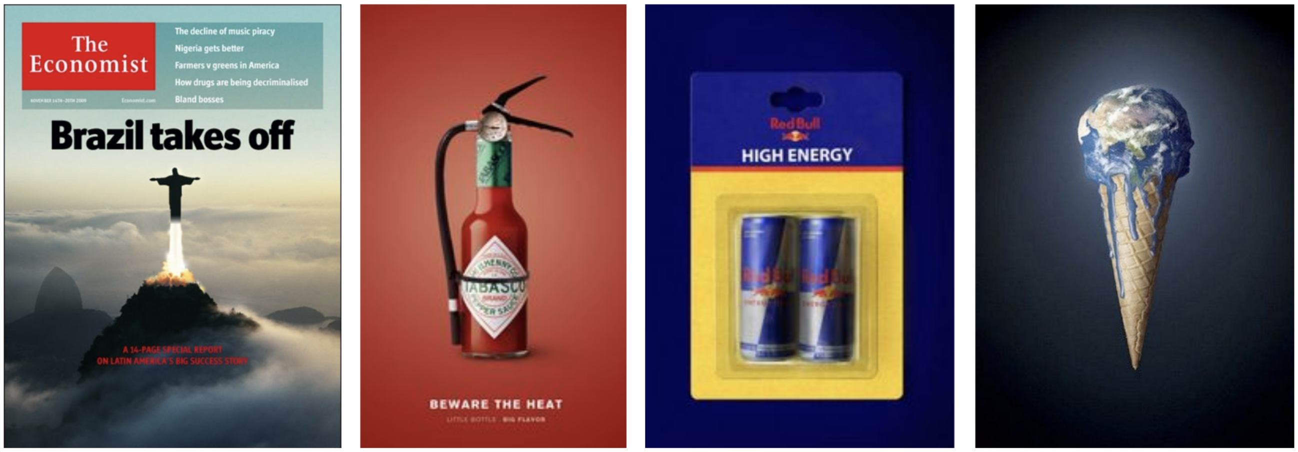

Figure 2: Examples of professional visual blends in print media. One magazine cover, two ads, and one PSA for global warming

there are too many possible pairs of images to look at them all, the system automatically synthesizes appropriate images into blends using the design pattern. It finds pairs of images that share a main shape, and then moves, scales, and crops them to create a mock-up of the blend. A human then evaluates each blend. If there are no blends, iteration is needed to refocus the brainstorming to find more symbols.

The VisiBlends system makes it possible for groups of people to collaborate towards creating a visual blend to convey their messages of news, events, and public service announcements. This paper makes the following contributions:

- Introducing and defining the problem of visual blends - a creative visual technique useful for conveying a message.
- Decomposing the process of creating visual blends into microtasks that follow the iterative design process including brainstorming, synthesis, and iteration.
- Introducing a design pattern which specifies the abstract structure for synthesizing images into visual blends.
- VisiBlends: a system for users to collaboratively generate visual blends for their own messages.
- Three studies showing users can generate blends through independent microtasks, groups can collaboratively make visual blends for their own messages, and VisiBlends improves novices' ability to make visual blends.

The discussion addresses lessons learned about decomposing the design process and how to generalize this approach to other creative design problems.

RELATED WORK

Linguistics and Visual Metaphors

Visual metaphors are a media technique studied in linguistics and psychology because they are related to the field of pragmatics - the study of how context and implicature are used to create meaning. Charles Forceville, a prominent researcher in visual metaphors, proposes a theory of the mechanics underlying visual metaphors [11]. He presents case studies of ads similar to those in Figure 2 and theorizes that when viewers encounter a visual metaphor, they recognize an object in an image, but also notice something is odd about it. This deviates from the viewers' expectations and causes them to seek a meaning. Often the meaning is not entirely clear from the image alone. Supporting text on the image is needed to indicate that the image is an advertisement for a product. Then the user understands that one of the objects is meant to be interpreted literally (like the Tabasco bottle) and one of the objects is meant to be interpreted figuratively (like the fire extinguisher, implying a meaning about the literal object ("Tabasco is hot.").

In his studies of the impact of visual metaphors on viewers across cultures, he identifies three types of visual metaphors: *Similies* where objects are "visually separate", *Hybrids* where objects are "fused together" and *Contextual Metaphors* where one of the objects is not visible but inferred from context or environment. He finds that Hybrids that fuse together objects have the highest positive impact on viewer appreciation. Additionally, he finds that highly complex visual metaphors are negatively correlated with appreciation. Simpler blends with fewer objects are easier to perceive and interpret [25]. In this paper we create the *Hybrid* blends Forceville describes: images that fuse two simple objects into one.

Design Patterns

Design patterns are high-level solutions to recurring engineering and design problems. This includes architectural patterns [4], software engineering patterns [12], and web design patterns [10]. Design patterns are reusable solutions, but because they are abstract, effort must be put into understanding when to apply them and how to adapt them to a new problem.

Design patterns can be used to automatically solve some design challenges such as laying out furniture in a room [19], generating usable maps [3], illustrating furniture assembly instructions [2], or sequencing cuts in film [16]. However, when the whole problem can't be computed automatically, human intelligence can also be used to complete design patterns. Motif [15] used design patterns to help novice film makers structure their videos. Human-robot interaction programming can be facilitated by design patterns [21, 14]. Crowd innovation techniques [28, 27] used schemas (which can be viewed as design patterns) to propose innovative products through analogies. In general, design patterns are abstract solutions that can provide high-level structure to solving problems.

Decomposing Design

Many web-based systems have made progress toward scaffolding the design process in online environments that enable collaboration. Yu and Nickerson [29] crowdsourced the design of chairs by mixing ideas across users to spur innovation. Yu and Kittur [28, 27] used the crowd in a two-stage, analogy-based product idea generation mentioned previously. BlueSky [13] and IdeaHound [22] both use brainstorming and crowd ideation to solicit a diverse set of ideas that span a design space. IdeaHound then uses a hybrid of human intelligence and computational techniques to cluster the ideas into an idea map to make sense of the design space. Voyant [26] and CrowdCrit [18] allow creators to solicit feedback and structured evaluation from crowds to enable them to get multiple perspectives on their work and iterate. In general, scaffolding the design process leads to better outcomes [6, 9, 20]. Recent systems have had success decomposing the design process into large tasks which have experts [20] or dedicated students [24, 30] do the tasks with the benefit of a manager who coordinates the workers. There is still an open problem of how to decompose the design process into independent microtasks so that people can work in a decentralized manner.

Microtask workflows are a common way to structure crowdwork [17, 5, 8]. However, open-ended tasks like planning conferences or vacations [31, 7] often require the collaborators to work together towards a common goal. In these systems the workflow is not static. The system provides feedback on progress as users iterate towards the goal. Creating collaborative pipelines with feedback and iteration are a promising approach to coordinating online collaboration to achieve the goal of a high-quality output that meets the constraints of the problem. VisiBlends takes a collaborative pipeline approach towards solving a creative design problem and explicitly leverages the iterative design process.

DEFINITION OF VISUAL BLENDS

Visual blends associate two concepts by blending objects related to each of the concepts. We define a visual blend as having the following properties:

1. **Two concepts.** The input to a visual blend is two concepts. These concepts can be concrete, such as "RedBull" or "Brazil" or more abstract like "energy" or "taking off."
2. **Two objects.** For each of the two concepts, the visual blend has an object that is a visually recognizable symbol of that concept.
3. **Two objects are integrated into one object.** In order for the two objects to appear blended, they must be integrated into one object. They cannot simply be next to each other or in the same scene as each other.
4. **Both objects are recognizable.** Both objects must be individually recognizable to the viewer so they can see what concepts they symbolize, and infer the association between the concepts.

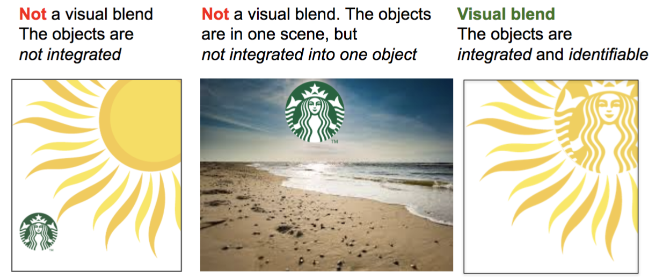

Figure 3: Three ways of combining symbols for *Starbucks* and *summer*. The first 2 do not meet the definition of a visual blend, but the third does by integrating the two symbols into one object.

Figure 3 shows examples of what is and is not a visual blend. All three images contain two objects, one of which is symbolic of *Starbucks*, and one which is symbolic of *summer*. The first image places the two objects near each other. This is not a visual blend. Although both objects are identifiable, they are not integrated. The middle image places a Starbucks logo on a beach scene and fully covers the sun. This is also not a visual blend. Since the Starbucks logo fully covers the sun, the two objects are not integrated into one object. Instead, the logo looks integrated into a beach scene. The image on the right is a visual blend. It integrates the Starbucks logo into the sun, and both objects are individually recognizable.

BLENDING DESIGN PATTERN

The main challenge in creating a visual blend is finding a way to blend two objects into one object and yet have both objects be recognizable. Our approach to satisfy these two requirements is based on theories of the human visual perceptual system and on an analysis of hundreds of visual blends.

The human visual system uses many different visual features at different stages to recognize an object including the object's simple 3D shape, silhouette, depth, color and details [23]. Based on this cognitive model, our approach to creating a visual blend is to make one object that borrows visual features from two objects. In the blend, some of the features will indicate one object, other features will indicate another object, and some features may indicate both objects.

The approach we explore here are blends of two objects that share a simple 3D shape but the silhouette indicates one object and the colors and details indicate the other object. For example, Figure 4 shows the *Earth + ice cream cone* blend. The *Earth* and *ice cream cone* share a simple spherical shape. The ice cream cone is identifiable by the silhouette created by cone and the Earth is identifiable by the blue and green color and pattern details inside its main sphere. We call this design pattern for blending objects *Single Shape Mapping* because it matches two objects based on a single shared shape, then blends them by mapping *all* of one object into *part* of the other object.

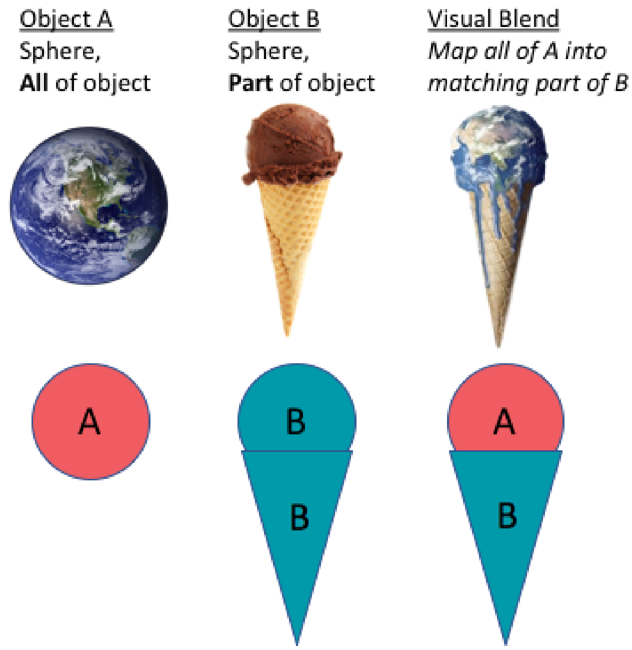

Figure 4: An illustration and examples of the *Single Shape Mapping* design pattern.

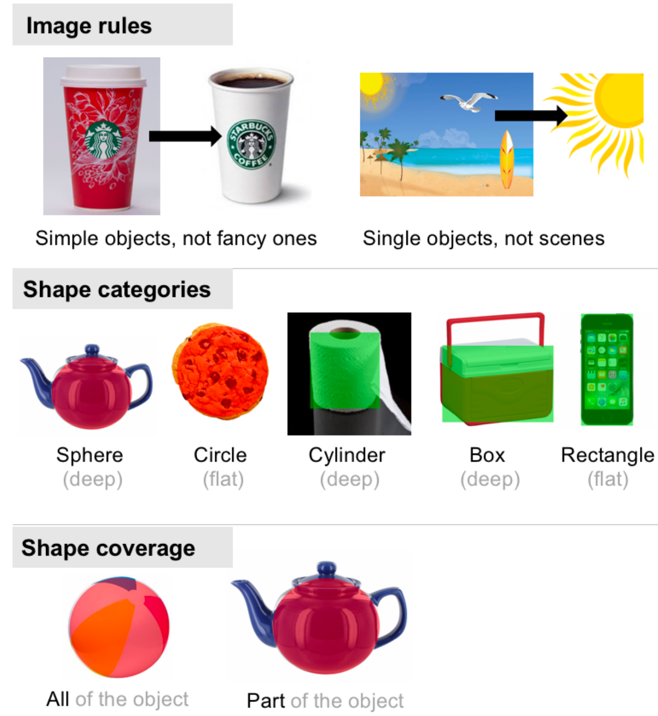

Figure 5: Guidelines for 3 pipeline microtasks: finding image, shape categories, and shape coverage: whether the shape covers all of the object or part of the object.

VISIBLENDS SYSTEM

The VisiBlends system is a pipeline that uses a hybrid of human intelligence and computational techniques to create visual blends. The pipeline scaffolds the design process with separate interfaces for microtasks like brainstorming, annotation, and evaluation. The pipeline is a website implemented in the Meteor web framework, using a Fabric.js drawing canvas to annotate shapes and mock up blends.

The target users are groups of 2 or 3 people with a message they want to convey. The group could be a company, student organization, news outlet, or any other community. These messages can be news headlines, advertisements or public service announcements. To begin, the users must find two important concepts from the message that they want to associate in the blend. For the headline "Football Dangerous to Youth Development," the users could pick *football* + *dangerous* as the two concepts to blend. The concepts must be broad enough so that there is enough variety in the symbols to find matches. If the concepts are not broad enough, the users may need to brainstorm to broaden them.

The pipeline has six steps inspired by the iterative design process. Before users can participate in the pipeline, they must complete a 15-minute training session to learn what visual blends are and the steps to make them.

For both of the two concepts,

1. Users **brainstorm** associations with the concept.
2. Users **find images** of objects that visually represent the concept in simple, iconic ways.
3. Users **annotate images** for shape and coverage.

With the collection of annotated images for both concepts,

4. The system automatically **detects** which images to blend.
5. The system automatically **synthesizes** the blends.
6. Users **evaluate** each blend.

If the first iteration does not yield a satisfactory blend, then users can repeat the process: brainstorm more concepts, find more images, and evaluate the new blends.

Pipeline Steps

Once users have received training on the entire pipeline, they are ready to participate in any of the microtasks. The input to the pipeline are two concepts such as *Starbucks* and *summer*.

1. Brainstorming

For each concept, we want to find multiple objects that represent it. Although most people can easily think of a few objects that represent a concept, we need many options to increase the likelihood of finding a blend. Thus, we broaden their brainstorm by having them think of associated people, activities and settings. One such brainstorm for *summer* includes the following:

- **Objects**: beach, pool, sunglasses, watering can, lawnmower
- **People and actions**: lifeguard: saving people from drowning, students: taking vacation from school, baseball manager: coaching players

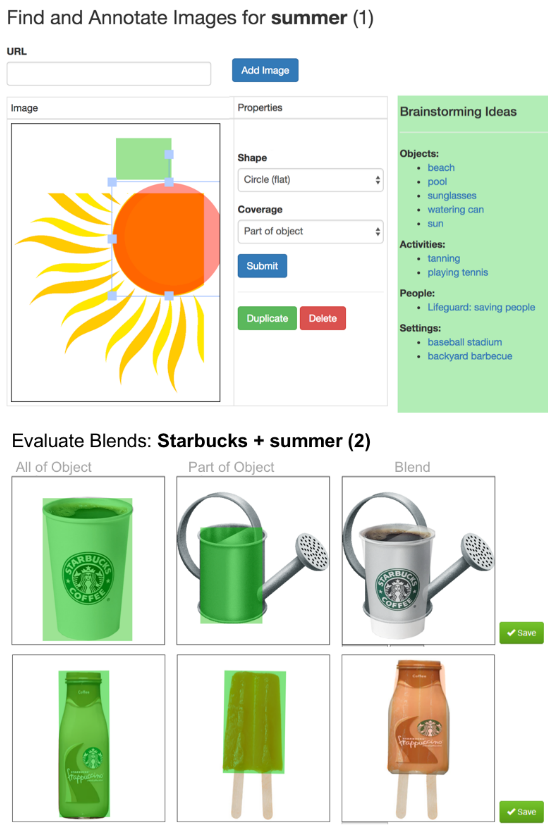

Figure 6: Interfaces for finding and annotating images (top), and seeing and evaluating blends (bottom).

- **Activities**: tanning at the beach, playing tennis, Celebrating 4th of July
- **Settings**: baseball stadium, backyard barbecue, pool, beach

The objects brainstormed for summer did not include “barbecue,” but one of the settings brainstormed was “backyard barbecue,” which led to the object “barbecue.” Brainstorming people, activities and settings indirectly helps users think of more objects. This technique is adapted from the an ethnographic observation technique called the AEIOU framework [1].

2. Finding Images

For a concept such as *summer*, a Google Image search will return many images that depict summer. However, very few of these images are actually useful for making blends. In the training session, we emphasize that the images users find must be simple, iconic objects with a single main shape. They can’t be people, animals, complicated scenes, or special versions of objects. To represent *Starbucks*, we want the iconic Starbucks cup of coffee, not the Christmas version, which clashes with the concept of summer. The top panel of Figure 5 shows good and bad examples of simple, iconic objects. Using the brainstorm results from the previous step, users find and input ten image URLS from Google Image Search. See Figure 6.

3. Annotate Images for Shape and Coverage

The *Single Shape Mapping* design pattern blends two objects based on their shape and how much of the object is covered by the shape. Thus we need users to annotate each image’s main shape and whether the shape covers the whole object in the image or only part of it.

In the interface, each image is presented in an HTML5 canvas where users can move and scale shapes to cover the main part of the object. They can then input what 3D shape best represents the object: circle (2D), sphere (3D), rectangle (2D), box (3D), or cylinder (3D) and whether that shape covers *all of the object* or *part of the object* (Figure 6). In the tutorial, users saw examples the shape category and shape coverage annotations and practiced annotating them (see Figure 5).

4. Matching Algorithm

Once users have found and annotated images for both concepts, the system can automatically detect which pairs of images should be blended according to the *Single Shape Mapping* design pattern. The algorithms takes in two sets of annotated image objects and finds all possible pairs across the two sets that meet the following criteria:

- **Shape match**: Both objects have the same main shape.
- **All-to-part match**: in one object the shape covers *all of the object* and in the other object, the shape covers *part of the object*.
- **Similar aspect ratio**: The height-to-width ratio of one object’s shape is within 50% of the other object’s shape.

5. Automatic Blend Synthesis

Once the matching algorithm has found a pair of annotated images, the synthesis tool produces a mock-up of the blend. We define Object A as the object with the shape mask covering *all of the object* and Object B as the object with the shape mask covering *part of the object*. The goal is to replace the masked part of Object B with all of Object A, as seen in Figure 1 and the “Evaluate Blends” interface of of Figure 6. The blend is automatically synthesized using the images for both of the objects and the size and position data from their shape masks. Object A is cropped to it’s shape mask and is scaled, resized, rotated and positioned to fit the size and position of Object B’s shape mask.

6. Evaluation

The blends produced by the algorithm are presented to users in a list that show Object A, Object B and the automated blend. See Figure 6. Users evaluate the blend by judging that it meets both criteria: (1) both objects are blended (2) both objects are individually recognizable. If so, they save the blend. They can make small adjustment to the masks to correct for small errors in annotations from previous steps. If the masks were drawn imprecisely, it can be tweaked in this interface before saving.

Iteration

If no blends are found, or the user wants to improve blends they can iterate. A naive way to iterate would be to simply find and annotate more images, and hope for more blends. However, users can also use their knowledge of the design pattern to see in what specific ways they should refine their

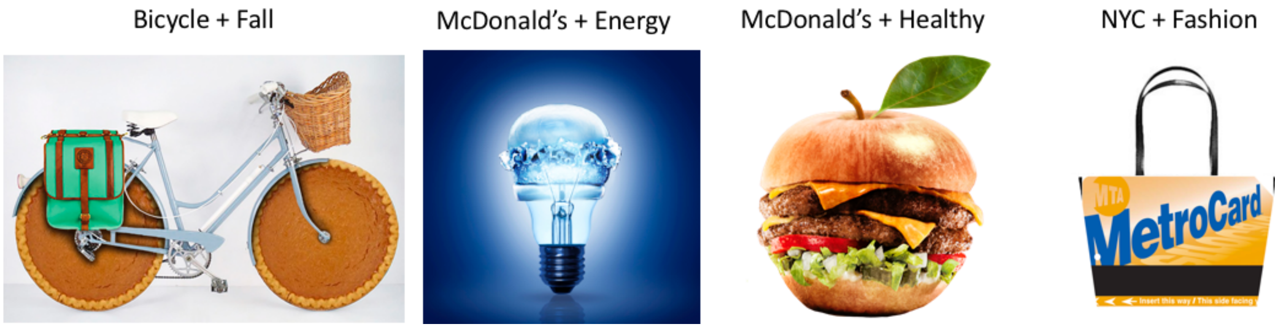

Figure 7: Examples of 4 of the 11 blends produced in the first iteration of the pipeline, with aesthetic editing by an artist.

search in order to maximize the possibility of finding good blends. For example, if all the *Starbucks* images are circles and all the *summer* images are rectangles, they can find more blends by finding more *summer* images that are circles.

Output

The system will often output many visual blends for a given concept pair. If the user wants to improve the aesthetics of the blends, they can either edit it themselves, or hire an artist from an online labor market to produce a professional-quality image based on the output.

EVALUATION

The aim of the VisiBlends is to decompose the iterative design process into microtasks to enable multiple people to collaborate in creating a visual blend. We address the following research questions:

1. **Independent microtasks**: Can users independently work on separate steps of the pipeline to create blends?
2. **Group collaboration on blends for messages**: Can groups of trained users collaboratively use the pipeline to create a blend for a message?
3. **Novice blend creation**: Does VisiBlends help novices create better blends than their innate process?

Study 1: Independent Microtasks

We often think of design has a process that needs centralized control - one person to who either does all the work, or is in control over all work done by others. However, if we can decompose the design process, we can enable contributors to work independently without the oversight of a central planner. To test whether the VisiBlends pipeline enables decentralized collaboration, we ran a study that made blends for 16 concept pairs where each step of the pipeline was done by a different person.

We recruited 7 university students (5 male, 2 female) for a 1-hour study who were paid $20. They spent 30-minutes completing an interactive tutorial on all stages on the pipeline. Afterwards, they spent 30-minutes doing microtasks. To ensure independence, each user participated in only one stage of the pipeline for any given concept pair. Although each person only worked on one step of each blend, it is important to note that everyone was trained on the entire pipeline. In early iterations we found that people were only successful at completely microtasks when they knew abstractly where each microtask fit into the entire pipeline. We expand on this in the discussion.

We tested the pipeline on 16 concept pairs made from 8 concepts: 4 brand or product concepts: *New York City*, *bicycles*, *Lego Toys*, and *McDonald's* and 4 nouns or adjectives to associate with a brand or product: *fall (season)*, *healthy*, *energy*, and *fashion*. Some of these combinations make sensible messages like "Riding a bike is healthy" and some are less sensible like, "Legos are fashionable." The goal at this stage is to test the feasibility of the making blends for many random concepts pairs. The concept pairs we picked had never been run before and there was no guarantee it was possible to find a blend for any of the inputs.

For each concept, the pipeline first collected a brainstorm of 40 ideas from 5 MTurk workers (they were paid $0.25 for 8 ideas). Next, the study participants independently completed the steps of pipeline: one person found 10 images of the first concepts, a different person found 10 images of the second concept, a third and fourth person annotated the two sets of images, the blending algorithm automatically produced visual blends for another user to evaluate. For each of the 16 concept pairs, there were a total of 320 brainstormed items, 80 images, and 80 images annotations. Our results show that the first iteration of the pipeline produced at least one visual blend for 11 of the 16 concept pairs. Figure 7 shows 4 successful blends found on the first iteration.

To find blends for the 5 remaining concept pairs, we asked a new user to look at the data from previous users and iterate to find blends. The naive way to iterate is to find and annotate more images, and hope this results in a new blend. Instead, users did something much faster. They looked at the shapes that were missing from one set of images and then specifically searched for objects with that shape. Their knowledge of the *Single Shape Mapping* design pattern helped them narrow their new search.

In summary, when running the pipeline on independent workers, the pipeline found blends for 11 of 16 concept pairs on its first iteration, and all 16 concept pairs on its second iteration. When iterating, the design pattern makes it easy to refocus the search.

Figure 8: Blends produced for the 5 messages in Study 2, with aesthetic improvements done by the users.

Study 2: Group Collaboration on Blends for Messages

After establishing that novices can use the pipeline in independent microtasks, we ran 5 case studies where groups of 2 or 3 people familiar with the pipeline collaborated to make visual blends for a message. During the study the group members worked in one room, each with their own laptop.

Each message is either a headline for a news article, a public service announcement or an advertisement inspired by events or concerns on campus. The groups were given messages and their concept pairs:

1. **News**: “Football linked to brain damage”
   Concept pair: *football* + *dangerous*
2. **PSA**: “Wash your hands. It's the smart move.’
   Concept pair: *Hand-washing* + *smart*
3. **Ad**: “Joe's Coffee: Open late”
   Concept pair: *Joe's Coffee* + *Night*
4. **Ad**: “Panel Discussion: Women in CS”
   Concept pair: *Women* + *Computer Science*
5. **Ad**: “Join the Philosophy Dept's Holiday Celebration”
   Concept pair: *Philosophy* + *Christmas*

On average, groups brainstormed more than 20 items and found more than 20 images for each concept. They found at least 3 good blends for each concept pair. Figure 8 contains posters made featuring these blends.

When using the pipeline for real messages, we discovered that it is easier to find images for some concepts than others. *Football* and *Dangerous* have many associated objects, while *Philosophy* does not. Thus, users found many visual blends for *football* + *dangerous* on the first iteration. *Philosophy* + *Christmas Party* had fewer blends because users struggled to find symbols of Philosophy. *Joe's Coffee* is a local chain and although users could brainstorm many symbols related to the concept, they could not find many images online. To compensate for the lack of images, they focused their search on finding more images of the other concept to increase the chances of finding a blend. Based on availability of symbols and images, users could choose which tasks were worth spending time on.

Based on the symbols and images found, users sometimes had to relax constraints to meet the goal. The best symbols for *Philosophy* were images of thought-experiments like the trolley problem and Searle's Chinese room. However, these are scenes, not single objects, as required for visual blends. Still, they were able to replace key elements of the scene with a *Christmas* symbol and still satisfy the definition of a visual blend. Although these images are more complicated than desired, they convey the idea of *Philosophy* + *Christmas* very well. *Women* + *CS* both have many symbols, but the group wanted to find symbols for *women* that were not gender stereotypes. Despite having two women and one man in the group, they struggled to find non-stereotypical symbols for *women*. They ended up blending a perfume bottle with a QR code. They liked the visual appearance of the blend, but the symbol was not ideal. In any design problem, it's hard to meet all the constraints, and users have to decide where they want to compromise.

Participants got many benefits from collaboration. They built off each others' brainstormed ideas and images. They annotated the same image in multiple ways based on how they modeled the object. They were surprised and delighted to see images they found being unexpectedly blended with an image another user found. They enjoyed looking at the blends together and seeing the reaction of other users in person rather than evaluating blends in isolation. They also made minor edits like correcting each others' annotations. In all cases, the users were engaged and the process was leaderless.

In all 5 case studies, iteration was used to improve existing blends by finding slightly different versions of the same objects. For example, in the *football* + *dangerous* blend, the first iteration blended a side-view of a red football helmet with the skull and crossbones. Although this qualifies as a visual blend, users iterated on it by finding a white, cartoon helmet that fit the aesthetic of the skull and crossbones better. The first iteration of the visual blending pipeline is great at finding shape matches, but iteration was always beneficial in refining the secondary visual aspects of the blend, such as color and style.

Study 3: Novices with and without VisiBlends

VisiBlends aims to scaffold the design process and make it collaborative. A key question is whether these features actually help novice designers make visual blends compared to their innate process without help. To test this, we ran a controlled study of individual novice users and compared their creation of visual blends with and without VisiBlends.

We recruited 13 undergraduates (11 female, 2 male) with no formal training in graphic design. Each person was introduced to the concept of visual blends with a definition, three annotated examples showing how two objects were blended into one object, and four exercises for them to annotate and check their answer. Each person was then asked to make six visual blends. Half the participants were assigned to the control-first conditions where they made visual blends without VisiBlends for the first set of three concept pairs, then with VisiBlends for the second set of three concept pairs. The other half were assigned to the VisiBlends-first condition, where they made visual blends first with VisiBlends, then without it. After making blends, we interviewed participants about their experience and collected demographics. This setup allowed us to compare visual blends made with and without the system. The study took a maximum of 1 hour and participants were paid $20.

In the control condition, people were asked to make visual blend mock-ups in Google Presentations. Before starting the tasks, they were given a warm-up to learn Google Presentations and ask any questions about the task. They were given 5 minutes for the warm-up to create mock-ups for the concept pair *apple (fruit)* + *energy*. All participants found the time sufficient to become familiar with the task and how to perform operations such as image search, copy and paste, bring to front, crop, and transparency adjustment.

After the warm-up, they were given five minutes for each concept pair to make as many visual blends as they could. The first set of concept pairs: *McDonald's* + *dangerous*, *Bicycle* + *smart*, *Football* + *autumn*. Next, they were given a tutorial on the design pattern and the VisiBlends tool and asked to create visual blends for the second set of concept pairs: *Joe's Coffee* + *morning*, *New York City* + *night*, *Local University* + *computer science.*

In the VisiBlends condition, people were asked to make visual blends using the VisiBlends tool with the benefit of reusing work from previous users. The main way VisiBlends uses collaboration is using many people to brainstorm symbols for concepts. Thus, we gave users the annotated symbols for each concept, from which VisiBlends automatically creates mock-ups. Users must then sift through all the results selecting good blends or improving them.

All the concept pairs in the study had never been made before, although it reused the symbols for concepts made in previous blends. For example, the concept pair *McDonald's* + *dangerous* had never been produced before. However, *McDonald's* symbols were taken from when users made blends for *McDonald's* + *healthy* and symbols for *dangerous* were reused from symbols found while creating blends for *football* + *dangerous*. One of the strengths of VisiBlends is that it allows brainstorms and symbols to be reused in new combinations. However, it is not guaranteed that there will be any matches for the new combination.

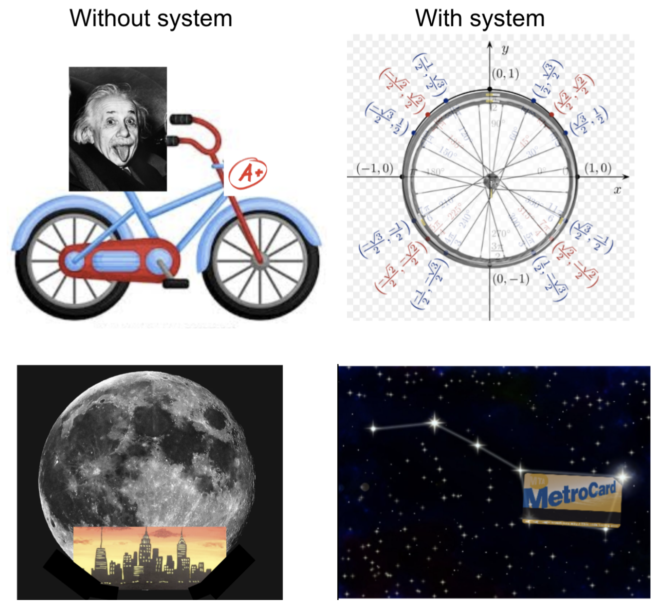

Figure 9: Examples of images made with and without VisiBlends for *bicycle* + *smart* and*New York City* + *night*. The images made without the system are not visual blends because they don't integrate two objects into one. The images made with the system are blends because they replace the main shape of one symbol with a matching shape of another symbol, making the objects look blended.

Results

During the study, participants created 332 attempts at visual blends of which 243 were successful. Blends were evaluated based on whether or not they met all of the objective qualifications of a visual blend – there are two objects blended into one such that each object is individually identifiable. The quality of the symbols chosen was not a part of the evaluation. We assumed that the symbols conveyed their intended meaning. The evaluation was performed by the authors. Both the number of attempted and the number of successful blends were counted for each person and concept.

Results show that using the tool dramatically increased the number of successful visual blends. Using a two-tailed t-test assuming unequal variance, we compared the number of successful blends made by users in the VisiBlends-first condition to the control-first condition. On average, participants made 5.67 successful blends per concept pair in the VisiBlends condition and 0.55 successful blends in the control condition ($t(22) = 5.91$, $p<0.001$). In the VisiBlends condition, users made more attempts at blends and had a higher success rate. On average, people made 6 attempts per concept pair with VisiBlends and 2.72 in the control ($t(22)=3.46$, $p=0.002$). Additionally, in the VisiBlends condition, users had a 96% success rate, as opposed to a 21% success rate without it ($t(18)=9.28$, $p<0.001$). Figure 9 shows examples of blend attempts with and without the tool for *bicycle* + *smart* and *New York City* + *night*. Although all the images contain symbols of both concepts, only the two made with the system are visual blends - they integrate two objects into one rather than placing an object on top of the other.

For the participants in the control condition who started without VisiBlends we measured the performance after introducing VisiBlends and found there is a significance increase in number of successful blends after they were introduced to VisiBlends. Their average number of successful blends jumped from 0.56 to 5.56 ($t(18)=4.88$, $p<0.001$) Conversely, for participants who saw VisiBlends first we measured performance after removing VisiBlends, and found their performance was much worse. With VisiBlends, they averaged 5.67 successful blends but after removing the tool, they produced an average of only 0.67 blends ($t(21)=5.84$, $p<0.001$). Once VisiBlends was removed, those participants had a very similar performance to those in the control group (averaging 0.67 blends, as opposed to 0.56 blends). Thus, using VisiBlends has a large and significant effect on performance even after the tool is removed, demonstrating the utility of the system beyond just having knowledge of the process.

After making blends in both conditions, participants were interviewed about their experience. In general, people expressed that using VisiBlends was a lot easier than making it themselves. P10 from the control condition said *"[The task is] harder than I thought. Using the tool was a lot easier. You already know what shapes to use."* When describing what was hard in either condition, 11 of the 13 participants mentioned that in the control condition, finding images was difficult, especially for abstract concepts. P6 said: *"Smart is such an abstract word. How do you picture smart? What is the symbol for smart?"* People often assume that Google Image Search will provide good symbols, but they are surprised to find that for abstract terms like "smart", "autumn" or "computer science" it returns beautiful images of scenes but very few images of objects.

None of the participants mentioned that making mock-ups in the control condition was hard. However, it did take time. P8 said *"This is taking a lot longer than I thought"*. 8 of 13 people mentioned that the tool produced higher quality mock-ups. Additionally, two people said that better mock-ups made evaluation easier. P5 said: *"It was easier seeing [system mock-ups] to know if they were good or bad. When you're making it yourself it's hard to know if it's good."*. Being able to evaluate your own work is crucial to success. Mock-ups should not be perfect, but they do have to be good enough to evaluate.

The biggest downside of the tool was that although it was easier, 4 of the 13 participants said it was more fun to make blends on their own based on the freedom to use whatever images they wanted. P2 said *"Making my own was a little more interesting. I could search for my own images and use my own sense of humor."* P11 said *"[The control condition is] nice because I can use whatever picture I want."*Although finding symbols is hard and takes time, it can feel like a game. For this study we wanted to test whether reusing symbols made the task easier, which is supported by the data. We did not consider how much fun it was. Outside of a study setting, it would be ideal to balance fun and ease by allowing people to use previously found symbols and add symbols that expressed their taste.

When asked about their strategies for making blends without VisiBlends, participants expressed a few different strategies. Two people focused on finding images: *"I thought of two objects then see if I could find a way to make them come together. I thought of them separately before the actual blend." (P3)* Five people mentioned looking for images with similar shapes that they could replace. *"[I] think of words, [and] just Google them. [Then] try to find similar shapes."* One other focused on the technical considerations first: *"I wanted to find things with transparent backgrounds so I could superimpose them."* Only one person had a strategy involving flare-and-focus. She collected multiple images of each concept on the desktop before choosing a pair to combine.

From interviews and observations, it appears that without VisiBlends, most people innately use a strategy that focuses on one of the constraints of the task, and often sacrifices the others. People that focus on symbols often sacrifice shape fit, people who focus on shape fit often sacrifice on symbols – they start using scenes rather than objects, or use minor aspects of an object rather than its main part. In Figure 9, the bottom left images places the New York City skyline in front of the moon. The skyline is an entire scene, and thus it can not fully integrate with the other object. The top left images places Einstein's face on a bicycle, but does not integrate him into the bicycle. Also, that image places an "A+" symbol where the basket should be, but it is not noticeable because the basket is a minor part of the bicycle. Sometimes successful blends can be found by focusing on one constraint, but participants describe it as *"look[ing] for coincidences" (P11)*. By scaffolding the design process to emphasize all the constraints, VisiBlends does not have to rely as much on coincidence.

DISCUSSION & FUTURE WORK

This paper presents a system that decomposes the design process into independent microtasks for a visual communication task. This leads to many interesting follow up questions about the nature of decomposing design.

Decomposing the design process into independent steps

Brainstorming is a crucial step in the design process that can obviously involve many people. However, to arrive at an effective solution, those idea have to be synthesized into cohesive output. Many people intuit that to have a cohesive output, the synthesis must come from one person's decisions. However, we found a way for users to collaborate through independent tasks by coordinating the work through a design pattern - a shared representation of how items from the the brainstorm can be coherently synthesized.

In early iterations, we found that although users can perform tasks independently, they need to have a high level understanding of how the tasks fit together. Users who weren't familiar with the entire pipeline performed poorly on their microtasks and asked many questions like "how different should the images I find be?" This is a hard question to answer with an explicit rule. Instead, when users were given training on the entire pipeline, they saw the purpose of each microtask and

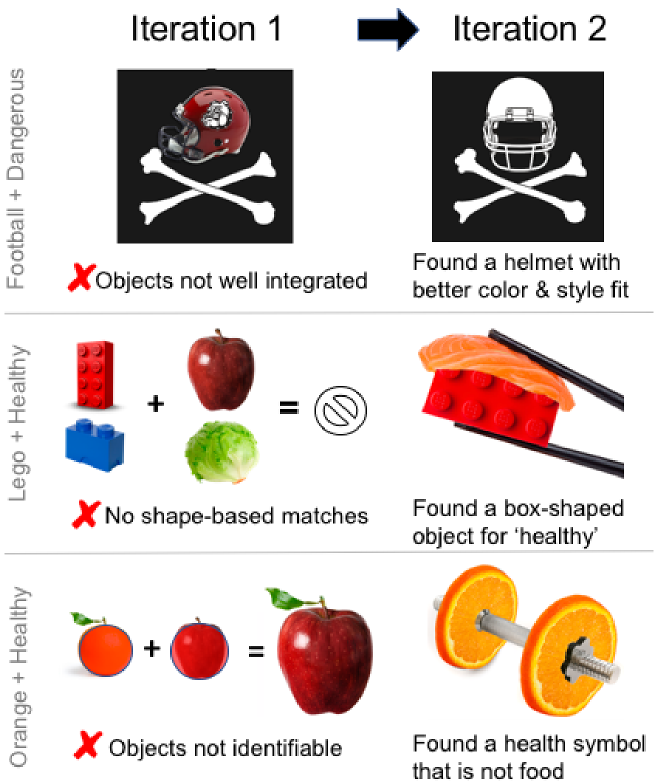

Figure 10: Three cases of iteration.

could implicitly answer those questions by reasoning about later stages in the pipeline: "The images should be different enough so that there is enough variety for the pipeline to find blends. Variety in shape matters, but variety in color does not." Design tasks are too complex to state all the explicit rules, however, a high-level understanding of the process and practice can provide tacit knowledge useful for collaborators to make good judgments.

Three uses of iteration in the design process

When following the iterative design process, it is generally good advice to design, prototype and iterate. However, no explicit guidance is given on how to iterate. In the visual blending task, we observed three common strategies for iteration. See Figure 10.

1. **Improving second-order features**. When the system finds blends, the shape fit is usually good, but its secondary visual features like color, texture, or scaling can be improved. To make the objects appear better integrated, users can find different versions of the object that match the style of the blended image as seen in the *football* + *dangerous* blend.

2. **No matches are found**. In the blend between *Lego* + *healthy*, users only found rectangular objects for *Lego* and round objects like apples and lettuce for *healthy*. The algorithm found no matches. However, instead of starting over, users can iterate by searching more precisely for the shape of items that they need to make more blends. Here, the user went back to the brainstorm for *healthy* and saw "fish" and realized that sushi was a rectangular object that could symbolizes healthy and blend with *Lego*.

3. **Emergent constraints**. It is possible to satisfy the the *Single Shape Mapping* pattern and not get a good blend. For *oranges* + *healthy*, apples are an iconic representation of *healthy* in the abstract, but when blended with an orange, the orange is not identifiable. This is a emergent constraint: something particular to this blend that would be hard to anticipate or create a rule for. One way to iterate is to brainstorm for symbols of *healthy* that aren't food. In Figure 10, orange slices are blended with a piece of exercise equipment.

In general, when users iterate they use the definition of a visual blend and their knowledge of the design pattern to understand what went wrong and to refocus their search for the images that will improve the results.

Decomposing other design problems

VisiBlends takes the general design process and tailors it to one specific problem, based on one design pattern. However, the design process and the idea of design patterns is very general, so there is hope that pipelines can be created for other problems. To do so we would need to know what components go into the solution and what abstract design pattern can describe how those components fit together.

Although we often assume creative design problems can't be reduced to an easy formula, many creative tasks do have patterns: stories have the Hero's Journey, music has chord progressions, proofs have proof techniques, software has design patterns and even academic papers have an abstract structure that advisors pass on to students. We could create pipelines that take this abstract knowledge and help users apply it by defining elements that need to go into the patterns and encourages them to iterate until the design pattern is satisfied.

There was no existing design pattern for visual blends, so we had to find the pattern by looking at examples and testing theories. To find design patterns it is important to ignore the surface level details and focus on the elements that are more fundamental to human cognition. For visual blends, shape was important to a blend. For a domain like persuasive writing, psychological principles of emotional states may be the key elements of a design pattern.

CONCLUSION

Visual blends are an advanced graphic design technique to draw users' attention to a message. Achieving this effect is challenging because there are two opposing goals: blending two objects into one while ensuring both objects are still recognizable. The VisiBlends systems help novices collaboratively create visual blends by decomposing the process for creating them into microtasks.

The process of creating visual blends has no obvious surface-level pattern. However, we discovered a deeper abstract structure: blend two objects that have the same basic shape but other identifying visual features. This task has implications for decomposing a broad range of design problems into independent microtasks.

REFERENCES

1. 2017. AEIOU Framework. https://help.ethnohub.com/guide/aeiou-framework. (19 September 2017).
2. Maneesh Agrawala, Wilmot Li, and Floraine Berthouzoz. 2011. Design Principles for Visual Communication. *Commun. ACM* 54, 4 (April 2011), 60–69. DOI: http://dx.doi.org/10.1145/1924421.1924439
3. Maneesh Agrawala and Chris Stolte. 2001. Rendering Effective Route Maps: Improving Usability Through Generalization. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '01)*. ACM, New York, NY, USA, 241–249. DOI: http://dx.doi.org/10.1145/383259.383286
4. Christopher Alexander, Sara Ishikawa, Murray Silverstein, Max Jacobson, Ingrid Fiksdahl-King, and Shlomo Angel. 1977. *A Pattern Language - Towns, Buildings, Construction*. Oxford University Press.
5. Michael S. Bernstein, Greg Little, Robert C. Miller, Björn Hartmann, Mark S. Ackerman, David R. Karger, David Crowell, and Katrina Panovich. 2010. Soylent: A Word Processor with a Crowd Inside. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 313–322. DOI: http://dx.doi.org/10.1145/1866029.1866078
6. Joel Chan, Steven Dang, and Steven P. Dow. 2016. Improving Crowd Innovation with Expert Facilitation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1223–1235. DOI: http://dx.doi.org/10.1145/2818048.2820023
7. Lydia B. Chilton, Juho Kim, Paul André, Felicia Cordeiro, James A. Landay, Daniel S. Weld, Steven P. Dow, Robert C. Miller, and Haoqi Zhang. 2014. Frenzy: Collaborative Data Organization for Creating Conference Sessions. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1255–1264. DOI: http://dx.doi.org/10.1145/2556288.2557375
8. Lydia B. Chilton, Greg Little, Darren Edge, Daniel S. Weld, and James A. Landay. 2013. Cascade: Crowdsourcing Taxonomy Creation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1999–2008. DOI: http://dx.doi.org/10.1145/2470654.2466265
9. 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 Trans. Comput.-Hum. Interact.* 17, 4, Article 18 (Dec. 2010), 24 pages. DOI: http://dx.doi.org/10.1145/1879831.1879836
10. Douglas K. Van Duyne, James Landay, and Jason I. Hong. 2002. *The Design of Sites: Patterns, Principles, and Processes for Crafting a Customer-Centered Web Experience*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
11. Charles Forceville. 1994. Pictorial Metaphor in Advertisements. *Metaphor and Symbolic Activity* 9, 1 (1994), 1–29. DOI: http://dx.doi.org/10.1207/s15327868ms0901_1
12. Erich Gamma, Richard Helm, Ralph E. Johnson, and John Vlissides. 1995. *Design patterns: elements of reusable object-oriented software*. Vol. 206. 395 pages. DOI:http://dx.doi.org/10.1093/carcin/bgs084
13. Gaoping Huang and Alexander J. Quinn. 2017. BlueSky: Crowd-Powered Uniform Sampling of Idea Spaces. In *Creativity & Cognition*.
14. Peter H. Kahn, Nathan G. Freier, Takayuki Kanda, Hiroshi Ishiguro, Jolina H. Ruckert, Rachel L. Severson, and Shaun K. Kane. 2008. Design Patterns for Sociality in Human-robot Interaction. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction (HRI '08)*. ACM, New York, NY, USA, 97–104. DOI: http://dx.doi.org/10.1145/1349822.1349836
15. Joy Kim, Mira Dontcheva, Wilmot Li, Michael S. Bernstein, and Daniela Steinsapir. 2015. Motif: Supporting Novice Creativity Through Expert Patterns. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1211–1220. DOI: http://dx.doi.org/10.1145/2702123.2702507
16. Mackenzie Leake, Abe Davis, Anh Truong, and Maneesh Agrawala. 2017. Computational Video Editing for Dialogue-driven Scenes. *ACM Trans. Graph.* 36, 4, Article 130 (July 2017), 14 pages. DOI: http://dx.doi.org/10.1145/3072959.3073653
17. Greg Little, Lydia B. Chilton, Max Goldman, and Robert C. Miller. 2010. TurKit: Human Computation Algorithms on Mechanical Turk. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 57–66. DOI: http://dx.doi.org/10.1145/1866029.1866040
18. Kurt Luther, Amy Pavel, Wei Wu, Jari-lee Tolentino, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. 2014. CrowdCrit: Crowdsourcing and Aggregating Visual Design Critique. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW Companion '14)*. ACM, New York, NY, USA, 21–24. DOI:http://dx.doi.org/10.1145/2556420.2556788

19. Paul Merrell, Eric Schkufza, Zeyang Li, Maneesh Agrawala, and Vladlen Koltun. 2011. Interactive Furniture Layout Using Interior Design Guidelines. In *ACM SIGGRAPH 2011 Papers (SIGGRAPH '11)*. ACM, New York, NY, USA, Article 87, 10 pages. DOI: http://dx.doi.org/10.1145/1964921.1964982

20. Daniela Retelny, Sébastien Robaszkiewicz, Alexandra To, Walter S. Lasecki, Jay Patel, Negar Rahmati, Tulsee Doshi, Melissa Valentine, and Michael S. Bernstein. 2014. Expert Crowdsourcing with Flash Teams. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 75–85. DOI: http://dx.doi.org/10.1145/2642918.2647409

21. Allison Sauppé and Bilge Mutlu. 2014. Design Patterns for Exploring and Prototyping Human-robot Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1439–1448. DOI: http://dx.doi.org/10.1145/2556288.2557057

22. 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 (UIST '16)*. ACM, New York, NY, USA, 609–624. DOI: http://dx.doi.org/10.1145/2984511.2984578

23. Robert J Sternberg. 2011. *Cognitive Psychology*.

24. Rajan Vaish, Snehalkumar (Neil) S. Gaikwad, Geza Kovacs, Andreas Veit, Ranjay Krishna, Imanol Arrieta Ibarra, Camelia Simoiu, Michael Wilber, Serge Belongie, Sharad Goel, James Davis, and Michael S. Bernstein. 2017. Crowd Research: Open and Scalable University Laboratories. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 829–843. DOI: http://dx.doi.org/10.1145/3126594.3126648

25. Margot van Mulken, Rob le Pair, and Charles Forceville. 2010. The impact of perceived complexity, deviation and comprehension on the appreciation of visual metaphor in advertising across three European countries. *Journal of Pragmatics* 42, 12 (2010), 3418 – 3430. DOI: http://dx.doi.org/https://doi.org/10.1016/j.pragma.2010.04.030

26. Anbang Xu, Shih-Wen Huang, and Brian Bailey. 2014. Voyant: Generating Structured Feedback on Visual Designs Using a Crowd of Non-experts. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1433–1444. DOI: http://dx.doi.org/10.1145/2531602.2531604

27. Lixiu Yu, Aniket Kittur, and Robert E. Kraut. 2014a. Distributed Analogical Idea Generation: Inventing with Crowds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1245–1254. DOI: http://dx.doi.org/10.1145/2556288.2557371

28. Lixiu Yu, Aniket Kittur, and Robert E. Kraut. 2014b. Searching for Analogical Ideas with Crowds. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1225–1234. DOI: http://dx.doi.org/10.1145/2556288.2557378

29. Lixiu Yu and Jeffrey V. Nickerson. 2011. Cooks or Cobblers?: Crowd Creativity Through Combination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1393–1402. DOI: http://dx.doi.org/10.1145/1978942.1979147

30. Haoqi Zhang, Matthew W. Easterday, Elizabeth M. Gerber, Daniel Rees Lewis, and Leesha Maliakal. 2017. Agile Research Studios: Orchestrating Communities of Practice to Advance Research Training. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 220–232. DOI: http://dx.doi.org/10.1145/2998181.2998199

31. Haoqi Zhang, Edith Law, Rob Miller, Krzysztof Gajos, David Parkes, and Eric Horvitz. 2012. Human Computation Tasks with Global Constraints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 217–226. DOI: http://dx.doi.org/10.1145/2207676.2207708