SymbolFinder: Brainstorming Diverse Symbols Using Local Semantic Networks

Savvas Petridis
sdp2137@columbia.edu
Columbia University
New York, New York, USA

Hijung Valentina Shin
vshin@adobe.com
Adobe Research
Cambridge, Massachusetts, USA

Lydia B. Chilton
chilton@cs.columbia.edu
Columbia University
New York, New York, USA

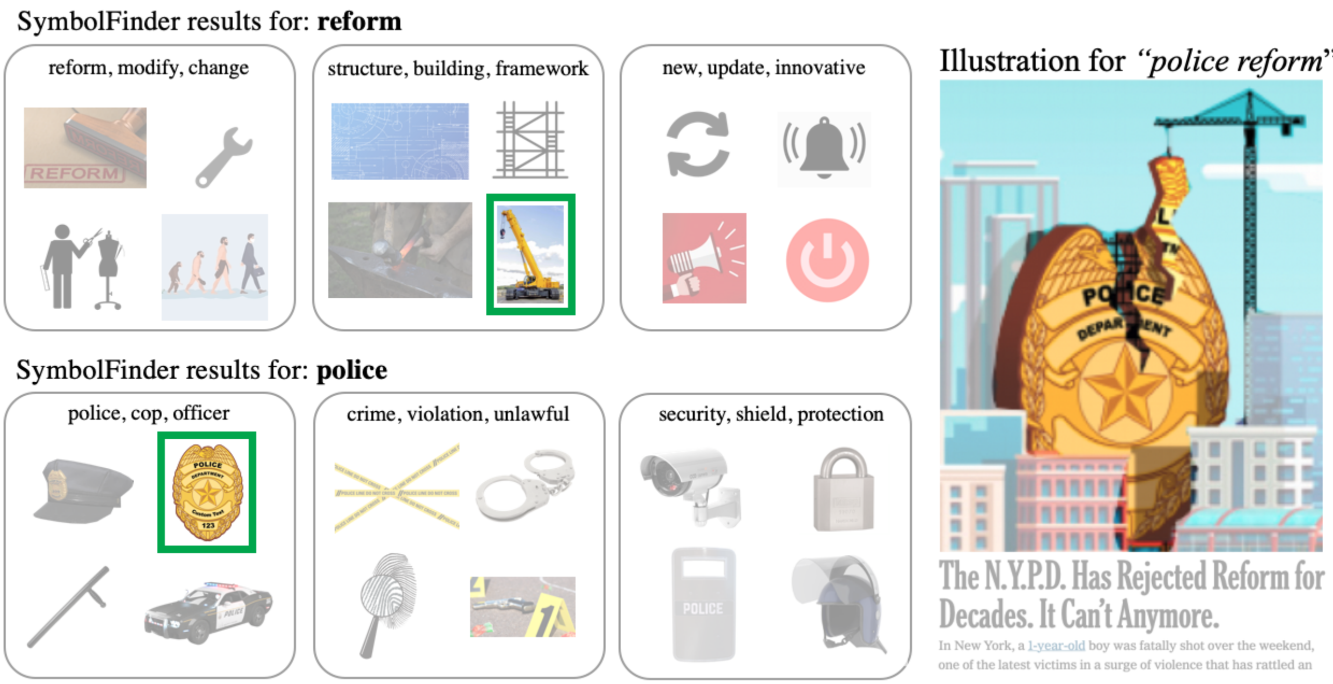


Figure 1: SymbolFinder takes as input abstract concepts like *reform* and *police* and helps users brainstorm many diverse objects that symbolically represent those concepts. With this diverse set of symbols, novice designers find it easier to make compelling symbolic illustrations.

ABSTRACT

Visual symbols are the building blocks for visual communication. They convey abstract concepts like *reform* and *participation* quickly and effectively. When creating graphics with symbols, novice designers often struggle to brainstorm multiple, diverse symbols because they fixate on a few associations instead of broadly exploring different aspects of the concept. We present SymbolFinder, an interactive tool for finding visual symbols for abstract concepts. SymbolFinder molds symbol-finding into a recognition rather than recall task by introducing the user to diverse clusters of words associated with the concept. Users can dive into these clusters to find related, concrete objects that symbolize the concept. We evaluate SymbolFinder with two studies: a comparative user study, demonstrating that SymbolFinder helps novices find more unique symbols for abstract concepts with significantly less effort than a popular image database and a case study demonstrating how SymbolFinder helped design students create visual metaphors for three cover illustrations of news articles.

UIST '21, October 10–14, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8635-7/21/10...$15.00
https://doi.org/10.1145/3472749.3474757

CCS CONCEPTS

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

KEYWORDS

brainstorming, symbols, interactive tool, design

ACM Reference Format:
Savvas Petridis, Hijung Valentina Shin, and Lydia B. Chilton. 2021. SymbolFinder: Brainstorming Diverse Symbols Using Local Semantic Networks. In

The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21), October 10–14, 2021, Virtual Event, USA. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3472749.3474757

1 INTRODUCTION

Visual symbols play a vital role in daily communication. They are used in public signs, user interfaces, logos, and advertisements to convey important information (Figure 2). Symbols quickly and effectively convey abstract ideas using representative concrete objects. For example, "lost and found" is represented by an umbrella and a glove and "search" is represented by a magnifying glass. These concrete objects are also the building blocks for more creative symbolic illustrations, such as the "global warming" PSA in Figure 2, which combines ice cream, a symbol of *melt*, with a symbol of *Earth*. Through these concrete objects, we can enable communication that is quickly understood and universal.

While there has been a great deal of work in the graphic arts and in icon design on how to create icons from an image of an object [31] [20], the problem of how to find these symbolic objects for an abstract concept has been relatively overlooked. Visual language is constantly evolving. New symbols are constantly being created to represent new experiences, organizations, and interactions on interfaces [28]. Novices with little to no experience in graphic design are also creating symbols for logos, websites, slide decks, mobile apps and games. Novices have difficulty not only designing icons from concrete objects, but also finding concrete objects to represent the concepts they want to symbolize in the first place.

Finding symbols is particularly challenging for novice designers when (1) the concepts they would like to represent are very abstract and (2) they want to combine them to create more complicated meanings. One such visual design challenge that inspired SymbolFinder and embodies both of these problems is visual metaphors: illustrations that combine symbols to convey a complex meaning, like the "police reform" illustration in Figure 1. The two symbols in a visual metaphor must be combined in such a way that their shapes blend naturally and the combined design accurately reflects the emotional tone of the message [13] [25]. To accommodate such constraints and create many design alternatives, it is essential to find a diverse set of symbols for the abstract concepts being depicted. However, converting these abstract concepts into a diverse set of visual symbols is hard for novice designers, preventing them from effectively combining them to convey a message.

In order to understand the challenges and workflow of novice designers, we conducted a formative study, where novice participants used Google Images to find symbols for abstract concepts. We observed that novices relied almost exclusively on recalling their own associations about the concept to search for related images. They often had difficulty brainstorming many different related words, and ended up fixating on a narrow set of associations, which represented a limited aspect of the concept being symbolized. Novices needed help to explore diverse ideas, which is crucial to finding an effective and creative solution [45] [58]. Finally, novices struggled to convert abstract associations into concrete objects and actions that could visually represent the concept.

Inspired by these observations, we created SymbolFinder to help novices find compelling visual symbols for abstract concepts. SymbolFinder helps users brainstorm associations by providing related words from an expansive word association data set. By clustering the related words into groups, each of which represents a related but distinct aspect of the concept, SymbolFinder encourages users to explore a broad range of related contexts, rather than fixating on a narrow set of associations. To create these clusters, SymbolFinder constructs a semantic network of word associations and detects highly connected communities of words. Finally, SymbolFinder helps users find imageable objects and actions by organizing words related to each cluster by word-concreteness.

Figure 2: Four visual symbols, from four domains: transportation hubs, human-computer interfaces, logos, public service announcements.

This paper presents the following contributions:

- SymbolFinder: an interactive interface for finding concrete images to represent abstract concepts.
- A technique for applying local semantic networks to word association data to help users perform a broad and deep brainstorm.
- An evaluation showing that users found on average 49% more unique symbols using SymbolFinder than they did using Google Images. Additionally, SymbolFinder was perceived to require significantly less effort and mental demand.
- A case study of novice designers using SymbolFinder to find the assets they need to create more than 10 different visual metaphor prototypes for each of 3 news articles.

2 RELATED WORK

2.1 Visual Symbols

Symbols are fundamental in visual communication and are used in a variety of contexts. They accompany headlines in news articles [28], represent actions in computer interfaces [54] [31], guide people in transportation hubs [48], represent corporations in logos [41], and form associations with products in advertisements [34]. There are many advantages in communicating ideas with symbols. Symbols often require less space to encapsulate an idea than using the word itself, saving space in interfaces, maps, and signs [27]. People can more quickly and easily recognize symbols than words because of our innate visual processes [33] [54]. Symbols are more universally understood than words across cultures, which is why they are used and designed for international transportation hubs [48] [40]. Finally, depicting ideas pictorially aids their memorability and recognition [6] [5]. For these reasons, we built SymbolFinder, to help convey more abstract ideas visually.

2.2 Brainstorming and Exploration Tools

Many tools have been created to help people brainstorm and explore related ideas. These systems are often designed to present a small set of related words or images to inspire new ideas. To present related textual ideas, *InspirationWall* [1] presents a few related topics from a knowledge graph, *V8 Storming* [38] uses word embeddings to find similar words to suggest, and *CrowdBoard* [2] utilizes a real-time crowd to suggest more personalized ideas. Other tools like *Idea expander* [57] and *IdeaWall* [53] present a few related images based on the current spoken ideas of its users. Koch et. al. created a cooperative contextual bandit system which recommends a few images that match the user's current semantic and visual preferences [39]. While displaying closely related words and images is very helpful for finding symbols, it is also necessary that these recommendations encapsulate different aspects of the concept. SymbolFinder organizes a network of words into clusters capturing distinct associations, enabling users to explore diverse contexts and images for an abstract concept.

Clustering is a popular method used to help users understand and explore large datasets. *Scatter/Gather* enables users to interactively choose clusters to find and explore specific documents in a large collection [15]. *Exploratory Labeling Assistant* presents clusters of documents to users as a preliminary step to help them label groups of documents themselves [22]. *Recipescape* clusters recipes for a dish based on the structure of its preparation, enabling users to find recipes with similar or different steps [11]. SymbolFinder clusters word associations to present users with diverse ideas related to the concept being symbolized. Word association data sets are often analyzed as networks, where words are nodes and edges represent associations between them [19] [17]. In this format, they are referred to as semantic networks. We construct a "local" semantic network, consisting of words near the concept being symbolized, and cluster it using a popular network clustering algorithm [4].

SymbolFinder is also closely related to previous work supporting associative browsing. Within this space there are two types of tools: sensemaking and information foraging. Sensemaking tools, such as Apolo [12] and Vigor [51], help users organize and understand information, whereas information foraging tools, like Refinery [37], help users find information. In this framework, SymbolFinder is an information foraging tool. Although SymbolFinder and Refinery are both information foraging systems, they have different goals: Refinery's goal is to find a specific subset of information, whereas SymbolFinder's goal is to help users find multiple, diverse pieces of information. Within associative browsing, there are also two interaction approaches: starting with an example, like Apolo or Refinery, or starting with an overview of the space, like Vigor. Although SymbolFinder and Vigor both provide overviews, they do it in different ways. Vigor, help users understand the overview by providing statistics on different clusters so that users can compare them. In contrast, SymbolFinder splits the workflow into two phases: 1) a breadth phase to get an overview of the clusters and 2) a depth phase to explore the clusters more deeply and find symbols from each cluster.

2.3 Query Expansion and Exploratory Image Search

The queries that users enter when searching for images are often ambiguous and can refer to many different real-world entities. Many researchers have recognized this problem and created tools to help users either clarify their search or find what they're looking for by providing a diverse set of image results. Textual query suggestions are a common technique for helping a user clarify their search. Keywords can be extracted from the most relevant documents associated with the query [59] [9] or taken from commonly occurring pairs of queries from search logs [32] [24]. Zha et. al. improved upon this technique by showing clusters of visually similar images for each keyword to help users preview and compare the images for each keyword [60]. *IGroup* employs a similar technique, by extracting common phrases (n-grams) from the most relevant documents associated with the query and presenting clusters of images for each of these phrases [35]. While keyword suggestions help users disambiguate their queries, they do not let them explore the broader associated meanings and contexts of their search. By using a word associations dataset, SymbolFinder presents broader contexts that expand the user's idea of the query to help them brainstorm.

Instead of using the documents associated with the images, other tools expand queries with knowledge bases to capture diverse intentions. *PARAgrab* takes synonyms, hyponyms, and hypernyms from WordNet [44] and presents these as related searches to users [36]. Hoque et. al. use both the incoming and outgoing of links of the query's Wikipedia page to provide a list of related queries [30]. A separate knowledge base is used to cluster these associations into categories like person, place, and location. *CIDER* adds to this work by spatially arranging the images from these different queries based on their visual attributes [29]. These tools serve to quickly disambiguate a user's search, like separating Denzel Washington the actor from Washington D.C., the place. However, the organization of these related concepts does not capture different meanings and greater contexts associated with the query. For example, for a query like *reform*, instead of returning a list of specific types of *reform*, SymbolFinder presents a set of diverse clusters, each encapsulating a different sense of *reform* like "fix, amend, redo" and "new, update, innovative", to help the user brainstorm. Lastly, ICONATE [61], a system for automatically generating compound icons, expands an abstract query with a manually created concept map. SymbolFinder instead enables users to conduct their own search over the possible associations related to the concept, which helps them understand the space better and find a symbol better suited for their goal.

There exist also a multitude of exploratory image search tools that help users explore diverse results by clustering images. Cai et. al. use text, link, and visual features to cluster a query's image results [8]. Leuken et. al. create a similar system, involving a dynamic weighting function for the visual features, creating clusters that better align with a human's idea of image diversity [56]. Fan et. al. create a visual summary of image results on Flickr by creating a topic network from user-generated tags. This enabled users to view an overview of the various images connected to their query and explore highly connected clusters [21]. By providing clusters of word-associations associated with the query, SymbolFinder also presents an overview of diverse contexts related to the query. However, a crucial component of SymbolFinder is the ability to dive deeper into each cluster and explore concrete words. By incorporating concreteness and in-depth exploration of each cluster, SymbolFinder helps users find objects to symbolize their abstract query.

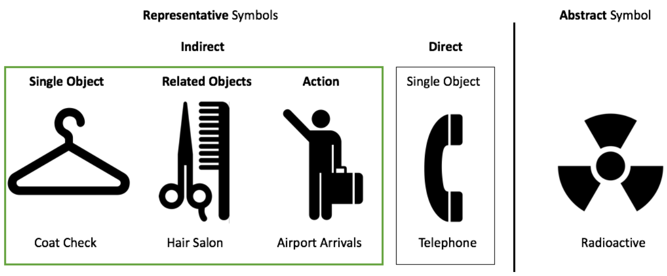


Figure 3: The types of symbols include: representative (indirectly and directly), as well as abstract (radioactive symbol).

3 BACKGROUND: WHAT MAKES A GOOD SYMBOL?

According to the theory of symbols, there are three basic types of symbols: abstract, directly representational, and indirectly representational [42] (Figure 3). A symbol is abstract when an abstract pattern represents the idea, like the radioactive symbol. A symbol is directly representational when its content is an exact representation of its idea, like the telephone symbol in Figure 3. A symbol is indirectly representational when the image content is associated with but not an exact representation of the idea, like the coat hanger, which represents a *coat check* (Figure 3). SymbolFinder was built to help people find indirectly representational symbols for abstract concepts that have a variety of meanings and contexts associated with them. These types of symbols do not require a new design like the *radioactive* symbol and are difficult to find with current image databases, unlike directly representational symbols, as these databases do not enable an exploration of various ideas related to the concept.

A representational symbol can contain three things: a single object, a few related objects, or an action [31]. For example, the coat hanger is the most essential object related to a *coat check*, and thus makes a good single object symbol. Sometimes an extra object makes a symbol more specifically related to the idea it represents. For example, a scissor and a comb together represent a *hair salon* better than either one alone. The two of them together effectively represent the tools a hair stylist uses. Finally, a symbol can also contain an action, like the *airport arrivals* symbol, in which there is a man hailing a taxi. These three categories make up the vast majority of the content displayed in representational symbols.

A good representational symbol is simple and concrete in the content it depicts [20] [31]. The most essential quality of a symbol is that it is recognizable. Its content should contain no more than what is necessary to depict the idea. Visual complexity and extra entities only make them slower to interpret and recognize. From this symbol theory, we establish a set of rules to help users of our system find good symbols (Figure 4). A good, indirectly representational symbol can be:

- **A single concrete object**. The object must be able to represent the concept on its own (Figure 4a).
- **Multiple related objects**. The objects should be related to the concept and to each other, like the combination of the scissor and comb in Figure 3. However, they should not be the same object, like the watermelon slices (Figure 4e), since one is enough to convey the idea. Symbols with more than two representative objects, like the collection of unrelated beach objects in (Figure 4e) are too complicated and can be separated into separate symbols.
- **A concrete action**. The action should be concrete and shown clearly, like the volleyball spike in (Figure 4c), as opposed to the more complex volleyball scene in Figure 4f.
- **No abstract scenes**. Symbols should depict a concrete object or action, instead of abstract landscapes (Figure 4d).

4 FORMATIVE STUDY

Novice graphic designers are designing symbolic illustrations, like logos and visual metaphors, to convey complex meanings. To create these illustrations, each concept requires a diverse set of symbols. This diversity is important to overcome constraints that occur later on in the design process [13]. To better understand the challenges novices face when searching for symbols and how to help them, we conducted a formative study in which we observed participants search for symbols with a popular image database, Google Images. Google Images is the primary tool used by novice and professional icon designers alike to look up visualizations of concepts [61]. Its interface also has many powerful features for exploring related queries including: query suggestions in the search bar, related queries with representative images above the image results, and filters for color and style. Finally, by appending "symbol" or "icon" to a Google search query, users can easily view a wide array of iconography for a particular concept. Therefore, we study how novices use Google Images when searching for symbols because of its ubiquity and its powerful search features.

4.1 Methodology

We recruited 5 participants (3 male, 2 female, average age 24.8) through an email mailing list for recent graduates of a local university. Every participant identified as a novice in graphic design and had used Google Images many times before. Participants were explained that the study was about understanding how novice designers brainstorm different visual representations of abstract concepts. In the task introduction, each participant was shown a slide deck, which introduced them to the concept of good, unique symbols. The slides include a step-by-step introduction to the symbol rules in Figure 4 and explained the importance of finding unique symbols that display different concrete objects and actions, not just the same objects in different colors. To ensure that they understood the rules, participants were asked to complete a quiz where they selected good and unique symbols of *summer* from a set of images. Incorrect answers were discussed with the experimenter.

Once participants felt they understood the task, they moved to the experiment phase of the study. Participants were given 10

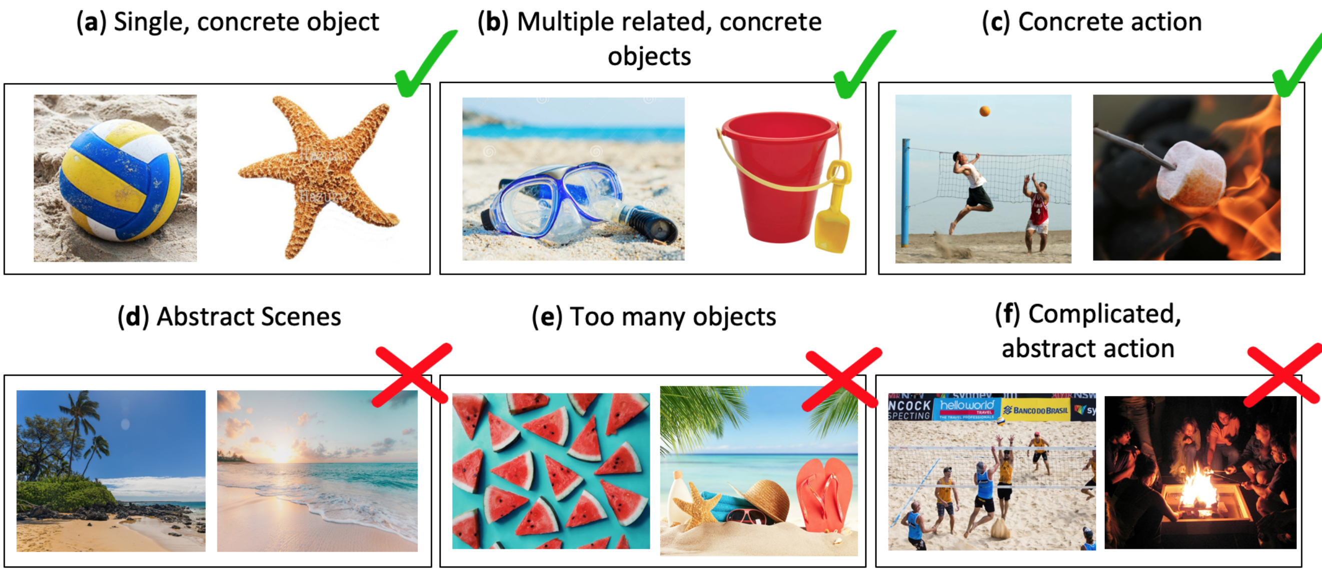


Figure 4: The rules for what makes a good symbol, derived from theory on icons and symbols, and explained with the concept: *summer*. These rules were shown to both raters and participants.

minutes to find as many good, unique symbols for three abstract concepts. In a brainstorm, more ideas are generally better, even if some ideas are not perfect, as they can inspire other, better ideas and can be iterated over later. To encourage a "more is better" brainstorming mindset, we asked them to find at least 20 symbols, which seemed both challenging but doable. The three abstract concepts were *old*, *exciting*, and *innovation*. These concepts were randomly selected from a visual messaging dataset, which contains the most common concepts symbolized in online messages [34]. Participants were asked to think aloud to convey their thought process. After each concept, they were asked to explain the benefits and drawbacks of Google Images, what search terms helped their brainstorming, and their general strategy. The study took at most 1 hour and participants were paid $20 for their time.

4.2 Observations

One author annotated the collected symbols for goodness, according to the symbol rules, and duplicates. Two symbols were considered duplicates if they conveyed the same object or action, regardless of color or image style. On average, participants found 14.6 (standard deviation=2.8) unique symbols for *old*, 10.8 (1.7) for *innovation*, and 11.8 (3.9) for *exciting*. Every participant searched for symbols during the entire allotted 10 minutes.

All five participants were frustrated by the lack of conceptual diversity in the images presented when searching the concept as is on Google Images. P1, P2 and P4 all mentioned that the results for *old* predominantly contained images of old people. Similarly, upon seeing the image results for *excited*, P1 states, "These are all images of the word 'excited'. Or just people looking excited." While there was generally a couple representations of the concept in the first set of images produced by Google, users found that they needed to brainstorm on their own to find different symbols.

The most common strategy to find different images was to search terms related to the concept and scan the image results for new visualizations. For example, P1 searched *ancient*, which he recalled on his own, and met many images of the Parthenon, the Colosseum, and pyramids. This turned out to be a fruitful context, from which he was able to collect an additional three symbols for *old*. Similarly, when seeing only images of excited people in the results for *exciting*, P2 subsequently searched *fun* and *adventure*. In doing so, he found other contexts related to *exciting* like extreme sports. Users had to recall these associations on their own. Therefore, our first design goal for SymbolFinder was to **help users brainstorm related words**, in order to enable recognition over recall.

Users however also struggled to find related words that presented different images and concrete contexts related to the concept. For example, when searching for symbols of *exciting*, P2 searched for images of *adventure* and *explore* and was met with similar images of hiking and camping. While he was able to collect a number of symbols from these searches, it was difficult for him to think of another related word that encapsulated a different flavor of *exciting*. Eventually, he searched the word *suspenseful* and found images of horror movies and theatre which inspired more symbols. From this issue we formed our second design goal: when helping users brainstorm associations, we should ensure that we present diverse ideas in order to help them collect **diverse symbols**.

Once users found a fruitful context, their strategy shifted to searching concrete objects and actions that they would select as their symbols. For example, while searching for symbols of *innovation*, P2 started searching for advanced technology like virtual reality goggles and hovercrafts. Similarly for *old*, P1 and P3 searched for objects old people use like canes and wheelchairs. While more abstract searches like *elder* and *technology* served as inspiration, these highly concrete searches contained the images that would end up being their symbols. When exploring related contexts, users should be able to explore concrete words within these contexts to find representative objects and actions. Thus, our third design goal was to help users **concretize abstract concepts**.

Design Goals. In summary, from the formative study we formed three design goals for the SymbolFinder:

D1: Help brainstorm related words to encourage recognition over recall. Users often recalled related terms to see new visualizations of the concept. By relying on their own memory, they miss obvious symbols and contexts associated with the concept.

D2: Symbol diversity. When helping users brainstorm related terms we should present them a variety of diverse associations so that they can collect diverse symbols from these associations.

D3: Concretize abstract concepts. As well as enabling users to explore diverse associations, they should also be able to explore related concrete terms for each association. This way, users can better find objects and actions to represent the concept.

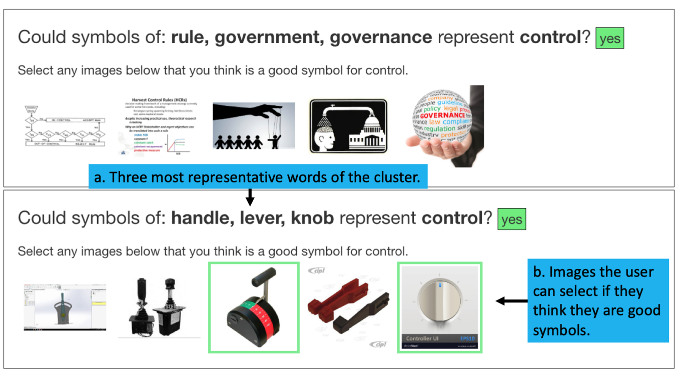


Figure 5: Phase 1 for the concept: *control*. Users select relevant clusters they would like to explore further in phase 2. Each cluster conveys a different association related to *control*, like the government (top) or physical tools we use to control machines. (bottom).

5 SYMBOLFINDER

To address these design goals we present SymbolFinder – an interactive tool that enables novices to find multiple, diverse symbols for abstract concepts. It uses a local semantic network to organize word association data into a hierarchical structure so users can explore diverse contexts associated with a concept. SymbolFinder's interface consists of two phases. Phase 1 is a breadth-first exploration of clusters of associations related to the concept; users select clusters they would like to explore further (Figure 5). Phase 2 is an in depth exploration of associated images and concrete words for each selected cluster (Figure 6).

5.1 Phase 1: Breadth-first concept exploration

Phase 1 addresses D1 (help brainstorm related words) and D2 (symbol diversity). To help users brainstorm a broad set of associations, we enable users to explore clusters of words that represent different aspects of the concept's meaning. Figure 5 shows a snippet of the Phase 1 interface for the abstract concept *control*. In this example, the first cluster is "rule, government, governance" and the second cluster is "handle, lever, knob", which are two distinct aspects of *control*. Users scroll through 10 such clusters and select ones to explore further in Phase 2. For each cluster, the user is posed the following question: "Could symbols of [word 1], [word 2], [word 3] represent [concept]?". The user is instructed to press "yes" if they think it might contain symbols for the concept. There are also 5 images related to these words. Users have the option to select an image if they think it is a good symbol. These images come from three Google Image searches, one for each of the words, where each query is formulated as follows: "[concept] [word]". This is done to keep the results relevant to the concept. The queries for top cluster in figure 5 were: "control rule", "control government", and "control governance". By having users explore a broad set of clusters briefly, we quickly expose them to a diverse set of associations, preventing them from fixation on a single one.

5.2 Phase 2: Image selection within clusters

Phase 2 further supports D1 (help brainstorm related words) as well as D3 (concretize abstract concepts). In phase 2, users further explore the clusters they selected from phase 1 (Figure 6a) and select symbols (Figure 6c). The key part of this interface is the sidebar on the left which is where users explore the clusters (Figure 6a) and recursively explore concrete words related to them (Figure 6b). To support D3, when users select the top level cluster words, they are shown related words sorted by concreteness (Figure 6b). In Figure 6, the user selected the "rule, government, governance" cluster. They then expanded *regulation*, one of the cluster words, and selected *referee*, a related concrete term. They could also view more associations of *regulation* by pressing the "see more" button. As well as exploring the clusters, users can also type associations they think of themselves in the "write your own" text boxes and view images and associations related to their entry. In this way, the sidebar enables users to recognize good symbols as well as use their own thought processes.

The second key part of this interface is the set of Google Image search results that populate the screen when a word is selected (Figure 6c). Four queries are made per word, and they include the word on its own [referee], the word and its parent [regulation referee], the concept and the word [control referee], and finally the word and "icon" appended to the search [referee icon]. We include the parent and concept queries as they help keep images on topic. We include the icon query as they often provide simple images of the action or item we are looking for. When users select an image, its link and metadata are saved. When they are done searching, they can download this data. Together, the sidebar of clusters and the multiple image searches effectively help users to find concrete symbols.

6 IMPLEMENTATION

SymbolFinder is implemented in the Flask web-framework. In the back-end, SymbolFinder uses the Small World of Words (SWOW) word association dataset [16] and a dataset of concreteness ratings

Figure 6: Phase 2 for the concept *control*. Users dive into the clusters they chose from phase 1. (A) On the left sidebar are the clusters, where users can explore related words. (B) Users can also view concrete associations for each related word in the cluster. While *regulation* is quite abstract, *referee* and *military* are more concrete and easier to visualize. (C) On the right are a few Google Image searches for the selected word: *referee*. The user selected two images, indicated by the green boxes.

for English words [7]. Calculating the network clusters and eigenvector centrality of nodes is done with the python library NetworkX [46]. Image search is implemented with Google's Custom Search API [3].

6.1 D1: Helping users brainstorm related words

To help users brainstorm a broad set of concrete associations, SymbolFinder uses word association data to find words that are related to the concept, have diverse connotations, and are concrete. We explored two different options for creating word associations: (1) Glove word embeddings, trained on Common Crawl [50] and (2) Small World of Words (SWOW), a crowd-sourced word association database [16]. Word embeddings are commonly used for comparing the similarity between words [43] and have been used in a number of brainstorming tools to compare the similarity of ideas [53] [10]. SWOW is a large English word association dataset. The dataset was created by having thousands of participants complete a word association task, in which each participant records the first three words they think of when seeing a cue word.

In initial testing, we found that SWOW produced words that were more relevant, diverse, and concrete than those by Glove. For example, for the abstract word *help*, the most related words that Glove produces include words like: *helping* and *need*, which are related, but are not diverse or concrete. Meanwhile, SWOW produces terms like: *donation*, *red cross*, and *tutor*. These terms are all related to *help* and they are diverse in that they represent actions, organizations, and people that *help*. Additionally, at least one term (*red cross*) is concrete. Thus, we chose SWOW to be our dataset for providing related words.

6.2 D2: Diversity using Local Semantic Networks

The SWOW word association dataset contains hundreds of associations for common abstract concepts, which when represented as an unorganized list, is too many for users to go through. Also, many of these associations are similar, which increases time spent and frustration parsing through them. To help users find diverse symbols from this data, these associations must be organized to identify a small number of diverse, yet highly relevant associations with the concept. To do this, we first create a local semantic network, to find all the words relevant to the concept. Then, a network clustering algorithm is run to identify sets of highly connected words that represent distinct associations of the concept.

For example, Figure 7 shows a 2-level local semantic network for the concept *control*. The first level of the network includes words strongly associated with *control* such as *government*, *rule*, *power* and *dominate*. The second level has words associated with the words in the first level. Including two levels introduces a greater variety of words while keeping words relevant. When this network is clustered, some first-level words such as *government* and *rule* are highly interconnected, and thus are merged into one association to present to users.

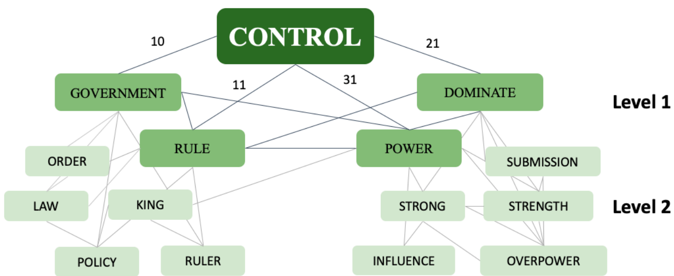


Figure 7: The two-level local semantic network for the word *control*. The first level of the network contains words like *government, rule, power* and *dominate*. The second level contains words like *king, law,* and *policy,* stemming from first layer words.

6.2.1 Constructing a local semantic network. To convert the word association data into a network, each word in the dataset is treated as a node and each association is an edge. The weight of an edge is the number of users who made that association in the word-association task [16]. From the concept being symbolized, a 2-level network is created (Figure 7). To create the first level, the first 60 strongest associations of the root concept, *control*, are added as nodes. We choose the concept's 60 strongest associations to include many associations and to keep the first-level highly related to the root concept. To create the second level, the first 5 strongest associations of each node in the first level is added. We create a second level to include a greater variety of words, but limit it to 5 per node so that few irrelevant words are added. A third level is not constructed, because associations this far from the root tend to introduce a lot of irrelevant words and make the clusters less interpretable.

6.2.2 Clustering the network. Our goal is to create a set of clusters, where each cluster contains highly related words that capture a distinct association of the concept. To cluster the network we considered two algorithms: the Clauset-Newman-Moore [14] and Louvain [4] network clustering algorithms. From initial experimentation we determined that the Louvain algorithm produced more interpretable clusters, as the Clauset-Newman-Moore algorithm tended to produce fewer clusters with a greater number of words, often combining clusters that the Louvain algorithm separated. The Louvain algorithm optimizes modularity, a measure which compares the edge density of the nodes in a cluster to the edge density of the same nodes in a randomly generated network. In our case, the algorithm identifies communities of highly related words by grouping words connected with high edge weights. The algorithm returns a hierarchy of clusters. We use the highest level of clusters (i.e. largest cluster size). For our dataset, the final pass of the algorithm generates about 12 to 20 clusters, which is small enough for a user to explore and large enough to contain a variety of unique ideas related to the concept.

Exciting		Future	
Related	Concrete	Related	Concrete
Fun	Motorcycle	Past	Crystal Ball
Thrill	Race car	Tomorrow	Hovercraft
Happy	Roller coaster	Crystal Ball	Robot
Interesting	Package	Someday	Car
New	Firework	Prediction	Spacecraft

Table 1: Associations sorted by strength of association (related) and by concreteness (concrete) for two abstract concepts: *Exciting* and *Future*. Sorting by concreteness highlights concrete objects that can serve as symbols.

To show users the most relevant clusters first, we sort the clusters by the average importance of their nodes. In this case, importance is defined as the eigenvector centrality of a word. A node has a higher eigenvector centrality if (1) it is connected to many other nodes and (2) its connections are also connected to many other nodes [47]. Sorting clusters in this way prioritizes the most relevant associations for a concept. For example, after sorting clusters for *control*, the most highly associated clusters are at the top of the sidebar in Phase 2, like "rule, government, governance" (Figure 6a). Meanwhile, more niche and perhaps less relevant clusters like "leash, harness, dog" are at the bottom of the list.

6.3 D3: Concretize abstract concepts

Although the clusters contain relevant and diverse associations, the words within them are not guaranteed to be concrete, like the "rule, government, governance" cluster in Figure 6a. Users should be able to explore concrete associations for each related word in the cluster. For example, the word *regulation* is very abstract, but *referee*, *military*, and *tax* are more concrete and thus easier to visualize (Figure 6b). By enabling users to view concrete associations for each cluster word, SymbolFinder helps users find imageable objects, actions, and people to symbolize the concept.

To incorporate concreteness, we use a crowd-sourced dataset of concreteness ratings for 40,000 English words and phrases [7]. Crowd-workers rated words on a scale from 1 (abstract) to 5 (concrete). In phase 2, for each top-level cluster word like *regulation*, we resort its associations by incorporating both concreteness and strength of association (Figure 6b). To include concreteness, we normalize the word's association strength and multiply this value by the word's concreteness score. We sort by this product. Consider the examples shown in Table 1. Instead of abstract related terms like *fun*, the concrete lists provide imageable words like *motorcycle* that can symbolize the abstract concept: *exciting*.

7 EVALUATION

We conduct a within-subjects study to evaluate whether users can find more unique symbols with SymbolFinder than with Google Images for abstract concepts. We also compare the perceived difficulty of finding symbols with SymbolFinder to finding symbols with

Google Images. As stated in the formative study, Google Images is a good baseline because of its ubiquity and powerful search features.

7.1 Methodology

We recruited 10 students via e-mail mailing list at a local university (2 female, 8 male), with an average age of 26.6. The participants had no formal training in graphic design. The study took at most 2 hours and participants were paid 40 dollars for their time. First they were introduced to the problem of finding unique symbols through a 10-minute warm-up task. Next, they were introduced to both systems. They were asked to use both systems to find as many unique symbols as they could within a time limit and rate the difficulty of the task. Lastly, they answered questions in a semi-structured interview about their experience.

In the task introduction, as in the formative study, each participant was shown a slide deck, which introduced them to the concept of good, unique symbols. To ensure that they understood the rules, participants were asked to complete a quiz where they selected good and unique symbols of *summer* from a set of images. Incorrect answers were discussed with the experimenter.

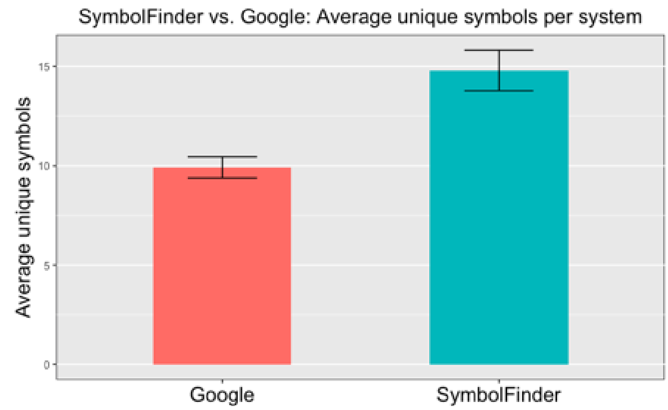
To set up the experiment, six concepts were selected for participants to symbolize. They were randomly selected from the same visual messaging dataset used in the formative study [34], from three levels of concreteness. The most concrete concepts were *fast* (concreteness=0.66) and *art* (0.83). The medium concrete concepts were *dangerous* (0.46) and *rugged* (0.55). The least concrete concepts were *control* (0.38) and *simple* (0.32). Every participant found symbols for the concepts in the following order: *fast*, *dangerous*, *control*, *art*, *rugged*, *simple*. To counter-balance the study, half of the participants were asked to use SymbolFinder for the first three concepts then Google images for the second set of three concepts. The other half did the opposite. This ensured that each condition had one concept from each level of concreteness.

In the experiment phase, participants were randomly assigned to a condition: SymbolFinder-first or Google-first. In both conditions, participants were given a short introduction to the interface, then given 10 minutes to find at least 20 good, unique symbols for each of the three concepts in that condition. From the formative study, we found 20 to be a challenging but appropriate target for the 10-minute time limit.

While they searched for symbols, participants were able to refer back to the good symbol rules, which were printed on a sheet of paper. After each concept, participants were asked to complete a NASA-TLX survey, where they rated their perceived work-load on a 10-point scale. After the experiment phase, we interviewed participants about their experience. They were asked questions which elicited feedback on which system they preferred and how their preferred system helped them complete the task.

7.2 Results

To evaluate the performance of the two systems, we needed to count the good, unique symbols found by participants. Although participants were asked to focus on finding only good, unique symbols, some of them did find duplicate symbols or symbols that did not conform to the rules. This is to be expected during a brainstorm, where participants are not supposed to edit their ideas, but focus on generating more ideas in an attempt to get better ideas.



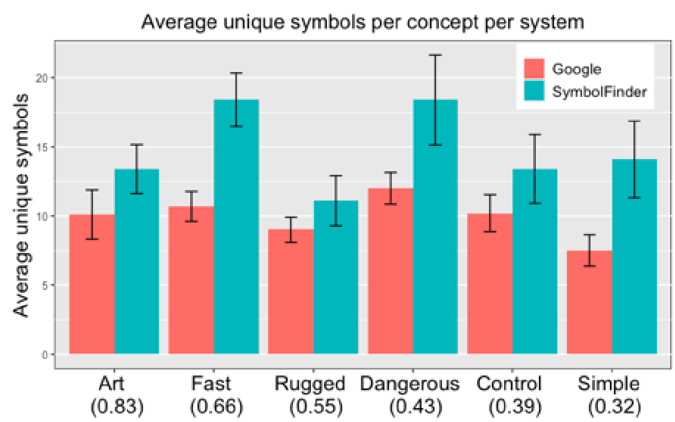


Figure 8: Top) Comparison of SymbolFinder and Google across all concepts. On average, users collected significantly more unique symbols with SymbolFinder than with Google. Bars are standard error. Bottom) Comparison of SymbolFinder and Google for each concept. On average, users collected more unique symbols per concept with SymbolFinder. The bars are standard error. Concreteness is in parenthesis.

	SymbolFinder	Google	p-value
Mental Demand	5.13 (1.41)	6.8 (1.97)	**<0.001**
Physical Demand	1.97 (1.43)	3.97 (2.58)	**<0.001**
Temporal Demand	5.13 (2.55)	6.17 (2.44)	0.11
Performance	6.77 (2.03)	5.9 (1.8)	0.03
Effort	4.87 (1.67)	7.43 (1.52)	**<0.001**
Frustration	3.0 (1.93)	4.33 (1.92)	0.057

Table 2: Comparison of SymbolFinder and Google Images for each category in the NASA-TLX questionnaire. 6 paired-sample Wilcoxon tests, with Bonferroni correction, show that mental demand, physical demand, and effort are significantly lower with SymbolFinder than with Google. Standard deviation is in parenthesis.

Figure 9: Users collected more unique symbols for each concept with SymbolFinder. Above are results for *control*, where the SymbolFinder user collected 15 unique symbols and the Google Image user collected 8.

To eliminate unrelated and duplicate symbols, we recruited two graduate students in design (who did not participate in the study) to annotate the collected images of each participant. For each image collected, they determined if it was a good symbol or not based on the criteria in Section 3 and the examples in Figure 4. Because of the natural subjectivity of this task, we had the annotators label two practice sets of images for good and unique symbols together. Based on the calculated Cohen's Kappa coefficient, the two raters had substantial agreement in their annotations for both goodness and uniqueness. The raters had a 94% agreement on goodness (Cohen's Kappa 0.74) and a 96% agreement on uniqueness (Cohen's Kappa 0.75). To determine the number of good and unique symbols for each participant and concept, we averaged the count annotated by the two raters.

Participants found significantly more unique symbols with SymbolFinder than with Google. To assess whether SymbolFinder helps people find unique symbols compared to using Google, we conducted an analysis of variance on a generalized linear mixed model (GLMM) with Poisson function, where the number of unique symbols is the response variable. This model can account for repeated measures from participants, as well as other factors that could potentially affect the number of unique symbols collected, such as (1) concept concreteness and (2) the order in which the tools were used. Thus, the fixed effects include: *System*, *Concept concreteness*, and *Order*. The random effect is *Participants*. The results indicate a significant effect of *System* ($\chi^2(1) = 57.3, p < 0.001$). There was neither a significant effect of *Concept concreteness* nor *Order*, confirming that the counter-balancing was effective. Following this result, we conducted a paired-sample Wilcoxon test and found that SymbolFinder users collected significantly more symbols than Google users ($V = 59.0, p < 0.001$). With SymbolFinder, participants collected on average 14.8 (stdev=5.5) unique symbols per concept, while Google Image users collected 9.92 (stdev=2.9) (Figure 8). Figure 9 shows unique symbols found for *control* by one participant using Google Images and another using SymbolFinder; the SymbolFinder participant found almost twice as many unique symbols.

SymbolFinder required significantly less mental demand and effort than Google. An analysis of variance test on a GLMM (with Poisson function) was conducted for each NASA-TLX category, with the same fixed and random effects from before. *System* had a significant effect for mental demand ($\chi^2(1) = 9.4, p < 0.005$), effort ($\chi^2(1) = 15.8, p < 0.001$), frustration level ($\chi^2(1) = 7.2, p < 0.01$), and physical demand ($\chi^2(1) = 19.9, p < 0.001$). Neither *Order* nor *Concreteness* had a significant effect on any category. Following these results, we conducted paired-sample Wilcoxon tests, with Bonferroni correction and found: mental demand ($V = 13.0, p < 0.001$), effort ($V = 21.0, p < 0.001$), and physical demand ($V = 4.0, p < 0.001$) were significantly lower with SymbolFinder (Table 2). Users often hit dead-ends of redundant symbols with Google. After exhausting the related searches at the top of the screen, they relied on their own brainstorming to find more symbols, increasing mental demand and effort. Physical demand is less meaningful. The significant result is likely due to users having to copy and paste images from Google into a slide-deck.

SymbolFinder helped participants most when it encouraged multi-faceted interpretation of concepts. While participants found more unique symbols for every concept with SymbolFinder than with Google, the difference was greatest for three concepts in particular: *fast*, *dangerous*, and *simple* (Figure 8). SymbolFinder users found 72% more unique symbols for *fast*, 53% more for *dangerous*, and 96% more for *simple*. For the more concrete concept, *fast*, the SymbolFinder clusters mapped broadly to different categories of fast things, such as animals, vehicles, fast-food, natural events, scientific processes, etc. SymbolFinder users dove into these clusters and found many concrete examples listed for each one. Similarly, for the least concrete concept *simple*, the SymbolFinder clusters mapped to adjacent meanings of *simple*, like *primitive* and *pure*. These associations broadened participants' conception of *simple*, and they were able to collect concrete objects like the caveman wheel and water droplet from them. Finally, SymbolFinder was less useful for *rugged*, where clusters mapped to redundant ideas, such as "hard, rock, stone" and "mountain, craggy, rocky". By broadening participant's interpretation of concepts, SymbolFinder helped users collect more symbols, regardless of the concept's concreteness.

8 CASE STUDY

To understand how SymbolFinder helps novice designers in practice, we deployed the system to a group of three students who make cover illustrations for a university science publication. Many of their illustrations were visual metaphors that combine symbols of two abstract concepts. We observed their process and interviewed them about their experience using SymbolFinder as a brainstorming tool. The team members were not co-located, but did much of their work together synchronously over teleconference software that allows screen sharing. We joined their conference call and observed them in three 90-minute sessions over a period of three months.

We selected three recent articles for them to make illustrations for. One article was selected from their school science publication and the other two were selected from The New York Times (NYT) for content diversity. For each headline, the team had to select the concept pair to represent it:

A1. "Public Health Messaging in Minority Communities and COVID-19's Neurological Effects" (science publication)
Concept pair: *Diversity* (concreteness = 0.45) + *Neurology* (0.5)

A2. "The N.Y.P.D. Has Rejected Reform for Decades. It Can't Anymore." (NYT)
Concept pair: *Police* (0.96) + *Reform* (0.4)

A3. "When the World Shut Down, They Saw It Open - The pandemic has made work and social life more accessible for many. People with disabilities are wondering whether virtual accommodations will last." (NYT)
Concept pair: *Disability* (0.69) + *Participation* (0.52)

From our observations we wanted to answer the following questions:

(1) **Picking concepts**. What kind of concepts do they enter into SymbolFinder? Are they abstract or concrete? How do they choose them?
(2) **Picking symbols**. What is their process of finding symbols and how does SymbolFinder help?
(3) **Combining symbols**. How do they use the symbols to make the illustrations?

8.1 Findings

During the three observation sessions, the team made a total of 32 low-fidelity prototypes that were iterated into 6 high-fidelity illustrations. Figure 10 shows the two high-fidelity illustrations they made for each headline. Their general process involved first discussing what two concepts they would combine from the headlines, then using SymbolFinder together to find multiple symbols for each concept. Next, they viewed all the symbols to consider which of them might be combined in the illustration. They then created a low-fidelity prototype by copying the images into PowerPoint, and if they liked the idea, they would discuss how to improve it to higher fidelity. One person would then create a high-fidelity illustration in Photoshop. Sometimes they would use Google Images to conduct a secondary search for an image that met some stylistic criteria (color, perspective, etc.)

As expected, SymbolFinder was indeed a part of the early brainstorm part of their design process while they were still exploring multiple possibilities. A somewhat surprising observation is that although the tool was built with a single-user in mind, they used it collaboratively with one person "driving" and screen-sharing while the two others looked at the results and commented on terms and images that interested them. In the future, it might be useful for SymbolFinder to be a multi-user system. However, it's also possible that the tool may work well with one main user driving the system and other users contributing ideas. Brainstorming sessions can be run with or without a leader. Having multiple users selecting concepts independently could lead to redundant symbols that would then need to be deduplicated.

8.1.1 Picking concepts. When selecting two concepts to represent a headline there are many possibilities. They could pick very concrete concepts that are easy to represent visually or abstract concepts that are more difficult to visualize. To select concepts, the team read the article title and text to extract multiple potential concepts that best capture the meaning of the article. For example, while working on A2, "The N.Y.P.D. Has Rejected Reform for Decades. It Can't Anymore", the team quickly picked *police*, a very concrete concept (concreteness = 0.96), as the first concept. For the second concept, they identified a few candidates that were all relatively abstract. This included: *reform* (0.4), *law* (0.51), and *scrutiny* (0.45), which are all very abstract words. Ultimately they made illustrations by combining symbols for *police* and *reform* (Figure 10). For the other two articles, they picked two fairly abstract concepts to combine: for A1. they picked *diversity* (0.45) and *neurology* (0.5), for A3 they picked *disability* (0.69) and *participation* (0.52). Given the content of the articles, **at least one of the two concepts they chose to represent it was abstract, thus making SymbolFinder an apt tool for their process.**

8.1.2 Picking symbols. We observed that when they collected symbols, **their focus was to find multiple, diverse representations of the concept**, as opposed to finding the perfect image for one particular representation. SymbolFinder helped them find different representations in two ways: (1) by exposing them to multiple different ideas through the clusters and (2) by presenting them with concrete objects within clusters to find symbols from these different ideas.

In the first phase of SymbolFinder for *reform*, the team found multiple distinct contexts associated with *reform*, many of which led to symbols in phase 2. They selected 7 of 18 clusters shown to them in phase 1, opting for clusters that captured different aspects of *reform*, like "reform, modify, change", "fix, amend, redo" and "new, update, innovative". Meanwhile, they rejected clusters that brought in connotations that did not fit with the overarching message: "police reform", like "political party, progressive, republican", which while related to reforming politics, is not so relevant to police reform. **Phase 1 of SymbolFinder helped the team quickly eliminate these irrelevant clusters**. Of the clusters they chose, the first obvious cluster "reform, modify, change" provided the most symbols at about 40%. However, the team found many useful symbols from the less obvious clusters as well, where 60% of their symbols were spread across 6 other clusters. The second most fruitful cluster was "fix, amend, redo", containing 22.5% of their total

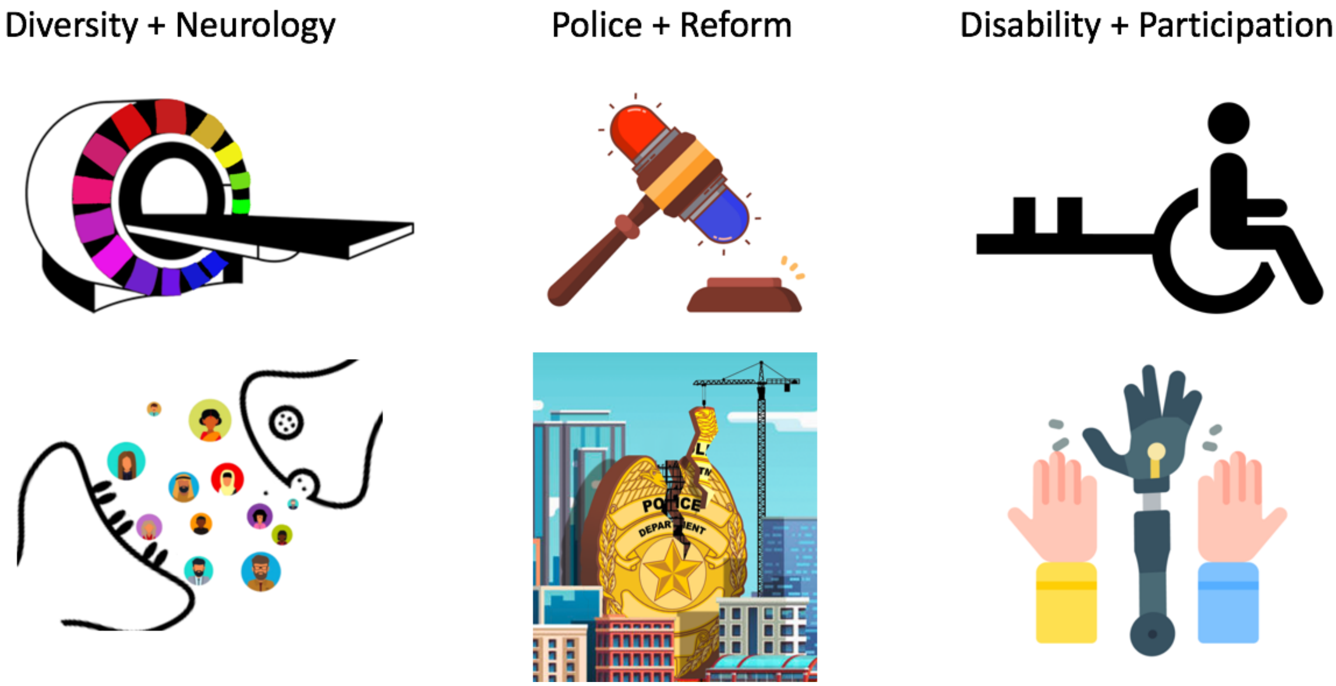


Figure 10: Cover illustrations consisting of two combined symbols for the three articles. The team of student designers found each symbol idea from SymbolFinder and used PowerPoint or Photoshop to create these prototypes. For *diversity + neurology*, the top illustration consists of an MRI machine and a color wheel, and the bottom illustration consists of a synapse and diverse people. For *police + reform*, the top illustration consists of police sirens and a gavel and the bottom illustration consists of a police badge and crane. For *disability + participation*, the top illustration consists of a person on a wheelchair and key, and the bottom illustration consists of a prosthetic arm and raised hands.

number of symbols. On average, the team used 6.83 (standard deviation = 2.8) clusters presented in SymbolFinder, demonstrating that **the clusters were useful for finding multiple distinct visual representations of the concepts**.

In the second phase of SymbolFinder for *reform*, **the team took advantage of the concrete sub-words to collect many different objects associated with that cluster**. For example, while exploring the "structure, building, framework" cluster, the team collected symbols of a *crane*, *scaffolding*, and *blueprint*, which appeared in the concrete sub-words of *building*. Similarly, from the "fix, amend, redo" cluster, they collected images from many concrete sub-words like *toolbox*, *saw*, and *screwdriver*. Concreteness helped the team quickly convert diverse clusters of concepts into symbols.

While picking symbols, a second constraint on the symbol space became apparent: the connotation and tone of the symbol. In phase 2 for *reform*, the team found symbols like the "update bell icon" and the "cycle refresh button" from the "new, update, innovative" cluster. Although both of these symbols did not have the tone or gravity they wanted for a illustration conveying "police reform", they collected them anyway, thinking they could be useful for future illustrations with *reform*. In the end, they were most excited by the *scaffolding* symbol, since its tougher tone and "New Yorkness" fit the article tone well. **Thus, while the team is predominantly seeking a variety of representations for each concept, they do keep in mind the tone of the article when looking for appropriate symbols**.

8.1.3 Combining symbols. To combine two symbols, the team employs a **matching strategy**. After finding symbols for both *police* and *reform*, the team placed the images side-by-side to ideate combinations between them. Commenting on their overall process, P2 explained "we start by choosing a symbol we like from one concept. Then we match that symbol with one from the other concept, usually based on shape or function." They ultimately made 10 initial prototypes, 2 of which were iterated over, shown in Figure 10. Across these 10 prototypes, they used 8 unique symbols of *police*, which came from 4 different clusters. They combined these *police* symbols with 8 unique symbols of *reform*, which came from 5 different clusters. **By having multiple diverse symbols, the team is (1) able to successfully find a match between two concepts with a higher probability and (2) create a diverse set of prototypes to show their client, using 6-8 unique symbols from each concept.**

9 DISCUSSION

In the following section we discuss limitations, future work to improve the system, and generalizable insights for future brainstorming tools.

9.1 Emerging vocabulary

Currently, SymbolFinder is constrained to the associations present in the SWOW dataset. There are two limitations to this: (1) symbols are defined culturally and the associations in SWOW are limited to the backgrounds of the users who built it (2) new or esoteric concepts will not appear in the dataset even though it is quite large. For example, in the past, the team worked on an article where one of the concepts was COVID-19, which did not exist in SWOW. In order to find symbols, the team brainstormed on their own, used Google Image Search, and also tried inputting related terms like *virus* into SymbolFinder. To include a new concept, we could extract related keywords from web search results or from a frequently updated knowledge base like Wikipedia. These keywords could then be linked to current entries in the SWOW dataset. We could also support symbol finding for different cultures by including international word association datasets and by helping users extract associations from international corpora.

9.2 Finding the perfect image

While SymbolFinder is effective for finding many diverse representations of an abstract concept, it is less useful for finding a specific image, once a particular representation is chosen. In the case study, after the team came up with an idea for an illustration using symbols they found with SymbolFinder, they would sometimes perform a secondary search with Google Images to find a particular version of the symbol. For example, while making the police badge and scaffolding illustration (Figure 10), the designer did not use the images of scaffolding they found with SymbolFinder. After imagining the symbol illustration, she had a specific idea for how she wanted the scaffolding to look. She wanted a "consistent background color so it would be easy to remove it". As well as a removable background, she wanted a 2d image that was neither "super busy", containing "overlapping scaffolding", nor too simple and "unnatural" looking. She ended up scanning many images to find the image she used. As well as finding the perfect image that contains the right visual detail, the team mentioned other constraints they consider in the secondary search, including finding images that are free to use and finding symbols of a particular shape (to increase more blend combinations). Fundamentally, SymbolFinder is a brainstorming tool, but in the future, we can incorporate tools to help users find particular versions of an image that fits their purpose.

9.3 Applying SymbolFinder to other visual media and databases

Though built on top of Google Images, we can add many other image databases to SymbolFinder, like the Noun Project [52], Flickr [23], or Shutterstock [55]. These datasets can often provide different types of images, like black and white iconography, to expand the diversity of images users see for a given concept. And beyond images, SymbolFinder can also be expanded to search GIFS and videos. Regardless of the image database or datatype, SymbolFinder's clustered local semantic network will help users explore diverse representations of an abstract concept.

9.4 Generalizable insights for future brainstorming tools

For future brainstorming tools, we believe that word-association networks (like SWOW) can be powerful tools to provide people related words that are relevant, concrete, and diverse. Traditionally, brainstorming tools have used word embeddings like GloVe [50] and word2vec [43] to suggest related ideas. However, current word embedding associations do not have all these desirable properties. This is likely because word embeddings are based on the distributional hypothesis of words: two words are similar if they often appear close together in a corpus [26]. However, the closest words related to an abstract concept can often include other abstract words and antonyms [49]. For example, calculating *control*'s closest words with word2vec yields antonyms like *uncontrollable*, different forms of the same word like *controlling*, and similarly abstract terms like *regulate*. These associations are not concrete and do not capture different associations of *control* and thus are not useful for brainstorming. Meanwhile, the word-associations in SWOW are created by people; they are more closely aligned to people's mental perceptions of words and have been shown to capture word relatedness better than word2vec [18]. For *control*, SWOW includes diverse associations like *government* and *dominate*, as well as concrete associations like *traffic light* and *leash*. For future brainstorming tools, word-association networks can be used to better generate related ideas and organize them, as opposed to using word embeddings trained on large corpora.

A fundamental hurtle of any brainstorming and idea-generation system is preventing fixation. Users can be tempted to dive into a sub-problem and ignore the overall solution space. And while brainstorming, it is critical to first rapidly go through many, different ideas prior to iterating and focusing on a subset. This was an issue in earlier iterations of SymbolFinder; in pilot studies users would spend most of their time parsing through the first few clusters of associations and ignore others that might have been more useful. To prevent fixation it was critical to split the workflow into a breadth and depth phase. Once users were made aware of the solution space in the breadth phase, they spread their attention more evenly across clusters in the depth phase and were able to find more solutions.

Concretizing abstract ideas is a cognitive challenge and is useful for many brainstorming and design tasks. We can use concreteness to convert a vague problem like, "How do we help people eat healthier" into a more actionable and specific question like "How can we make vegetables a more convenient snack food?". The abstract idea of "eating healthy" has many concrete associations, such as consuming more vegetables and cooking more instead of eating take-out. We can apply concreteness to this next set of associations to further refine the question into multiple concrete options like: "How do we make vegetables more snack-like?" and "How can we make cooking as convenient and tasty as take-out?". With a more concrete framing, it is now easier to brainstorm real solutions. Concreteness can be incorporated in future brainstorming tools to help users first refine and better specify the question they are brainstorming.

10 CONCLUSION

This paper presents SymbolFinder, an interactive tool that enables users to find diverse symbols for abstract concepts. In our user study we show that users can find significantly more unique symbols for abstract concepts with significantly less effort with SymbolFinder than with Google Images. We also conduct a case study showing how SymbolFinder is useful for creating cover illustrations of news articles. In the future, SymbolFinder can be applied to other media types, like GIFs, and other image databases. Also SymbolFinder could include tools to help users find a perfect image after a representation is chosen and expand its word association dataset automatically with new concepts.

REFERENCES

[1] Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. 2015. InspirationWall: Supporting Idea Generation Through Automatic Information Exploration. 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, 103–106. https://doi.org/10.1145/2757226.2757252

[2] Salvatore Andolina, Hendrik Schneider, Joel Chan, Khalil Klouche, Giulio Jacucci, and Steven Dow. 2017. Crowdboard: augmenting in-person idea generation with real-time crowds. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. 106–118.

[3] Google Custom Search API. 2020. *Google Custom Search API*. https://developers.google.com/custom-search/

[4] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

[5] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. 2016. Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 519–528.

[6] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. 2013. What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2306–2315.

[7] Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods* 46, 3 (2014), 904–911.

[8] Deng Cai, Xiaofei He, Zhiwei Li, Wei-Ying Ma, and Ji-Rong Wen. 2004. Hierarchical Clustering of WWW Image Search Results Using Visual, Textual and Link Information. In *Proceedings of the 12th Annual ACM International Conference on Multimedia* (New York, NY, USA) *(MULTIMEDIA '04)*. Association for Computing Machinery, New York, NY, USA, 952–959. https://doi.org/10.1145/1027527.1027747

[9] Claudio Carpineto, Renato de Mori, Giovanni Romano, and Brigitte Bigi. 2001. An Information-Theoretic Approach to Automatic Query Expansion. *ACM Trans. Inf. Syst.* 19, 1 (Jan. 2001), 1–27. https://doi.org/10.1145/366836.366860

[10] 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* (Singapore, Singapore) *(C&C '17)*. Association for Computing Machinery, New York, NY, USA, 93–105. https://doi.org/10.1145/3059454.3059455

[11] Minsuk Chang, Leonore V. Guillain, Hyeungshik Jung, Vivian M. Hare, Juho Kim, and Maneesh Agrawala. 2018. RecipeScape: An Interactive Tool for Analyzing Cooking Instructions at Scale. 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, Article 451, 12 pages. https://doi.org/10.1145/3173574.3174025

[12] Duen Horng Chau, Aniket Kittur, Jason I. Hong, and Christos Faloutsos. 2011. *Apolo: Making Sense of Large Network Data by Combining Rich User Interaction and Machine Learning*. Association for Computing Machinery, New York, NY, USA, 167–176. https://doi.org/10.1145/1978942.1978967

[13] Lydia B. Chilton, Savvas Petridis, and Maneesh Agrawala. 2019. VisiBlends: A Flexible Workflow for Visual Blends. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300402

[14] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Physical review E* 70, 6 (2004), 066111.

[15] Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. 1992. Scatter/Gather: A Cluster-Based Approach to Browsing Large Document Collections. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (Copenhagen, Denmark) *(SIGIR '92)*. Association for Computing Machinery, New York, NY, USA, 318–329. https://doi.org/10.1145/133160.133214

[16] Simon De Deyne, Danielle J Navarro, Amy Perfors, Marc Brysbaert, and Gert Storms. 2019. The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior research methods* 51, 3 (2019), 987–1006.

[17] Simon De Deyne, Daniel J Navarro, and Gert Storms. 2013. Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior research methods* 45, 2 (2013), 480–498.

[18] Simon De Deyne, Amy Perfors, and Daniel J Navarro. 2016. Predicting human similarity judgments with distributional models: The value of word associations.. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. The COLING 2016 Organizing Committee, Osaka, Japan, 1861–1870. https://aclanthology.org/C16-1175

[19] Simon De Deyne and Gert Storms. 2008. Word associations: Network and semantic properties. *Behavior research methods* 40, 1 (2008), 213–231.

[20] Ron S Easterby. 1970. The perception of symbols for machine displays. *Ergonomics* 13, 1 (1970), 149–158.

[21] Jianping Fan, Yuli Gao, Hangzai Luo, Daniel A. Keim, and Zongmin Li. 2008. A Novel Approach to Enable Semantic and Visual Image Summarization for Exploratory Image Search. In *Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval* (Vancouver, British Columbia, Canada) *(MIR '08)*. Association for Computing Machinery, New York, NY, USA, 358–365. https://doi.org/10.1145/1460096.1460155

[22] Cristian Felix, Aritra Dasgupta, and Enrico Bertini. 2018. The Exploratory Labeling Assistant: Mixed-Initiative Label Curation with Large Document Collections. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 153–164. https://doi.org/10.1145/3242587.3242596

[23] Flickr. 2020. *Flickr*. https://www.flickr.com/

[24] Bruno M. Fonseca, Paulo Golgher, Bruno Pôssas, Berthier Ribeiro-Neto, and Nivio Ziviani. 2005. Concept-Based Interactive Query Expansion. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management* (Bremen, Germany) *(CIKM '05)*. Association for Computing Machinery, New York, NY, USA, 696–703. https://doi.org/10.1145/1099554.1099726

[25] Charles Forceville. 1994. Pictorial metaphor in advertisements. *Metaphor and symbol* 9, 1 (1994), 1–29.

[26] Zellig S Harris. 1954. Distributional structure. *Word* 10, 2-3 (1954), 146–162.

[27] Kathleen Hemenway. 1982. Psychological Issues in the Use of Icons in Command Menus. In *Proceedings of the 1982 Conference on Human Factors in Computing Systems* (Gaithersburg, Maryland, USA) *(CHI '82)*. Association for Computing Machinery, New York, NY, USA, 20–23. https://doi.org/10.1145/800049.801748

[28] Nigel Holmes and Rose De Neve. 1985. *Designing pictorial symbols*. Rac Books.

[29] Enamul Hoque, Orland Hoeber, and Minglun Gong. 2013. CIDER: Concept-based image diversification, exploration, and retrieval. *Information Processing Management* 49, 5 (2013), 1122 – 1138. https://doi.org/10.1016/j.ipm.2012.12.001

[30] Enamul Hoque, Grant Strong, Orland Hoeber, and Minglun Gong. 2011. Conceptual query expansion and visual search results exploration for Web image retrieval. In *Advances in Intelligent Web Mastering–3*. Springer, 73–82.

[31] William K Horton. 1994. *The icon book: Visual symbols for computer systems and documentation*. John Wiley & Sons, Inc.

[32] Chien-Kang Huang, Lee-Feng Chien, and Yen-Jen Oyang. 2003. Relevant term suggestion in interactive web search based on contextual information in query session logs. *Journal of the American Society for Information Science and Technology* 54, 7 (2003), 638–649.

[33] WH Huggins and Doris R Entwisle. 1974. *Iconic Communication: An Annoted Bibliography*. Johns Hopkins University Press.

[34] Zaeem Hussain, Mingda Zhang, Xiaozhong Zhang, Keren Ye, Christopher Thomas, Zuha Agha, Nathan Ong, and Adriana Kovashka. 2017. Automatic understanding of image and video advertisements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1705–1715.

[35] Feng Jing, Changhu Wang, Yuhuan Yao, Kefeng Deng, Lei Zhang, and Wei-Ying Ma. 2006. IGroup: Web Image Search Results Clustering. In *Proceedings of the 14th ACM International Conference on Multimedia* (Santa Barbara, CA, USA) *(MM '06)*. Association for Computing Machinery, New York, NY, USA, 377–384. https://doi.org/10.1145/1180639.1180720

[36] Dhiraj Joshi, Ritendra Datta, Ziming Zhuang, WP Weiss, Marc Friedenberg, Jia Li, and James Ze Wang. 2006. PARAgrab: a comprehensive architecture for web image management and multimodal querying. In *VLDB*, Vol. 6. 1163–1166.

[37] S. Kairam, N. H. Riche, S. Drucker, R. Fernandez, and J. Heer. 2015. Refinery: Visual Exploration of Large, Heterogeneous Networks through Associative Browsing. *Comput. Graph. Forum* 34, 3 (June 2015), 301–310.

[38] Yui Kita and Jun Rekimoto. 2018. V8 Storming: How Far Should Two Ideas Be?. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) *(AH '18)*. Association for Computing Machinery, New York,

NY, USA, Article 14, 8 pages. https://doi.org/10.1145/3174910.3174937
[39] Janin Koch, Andrés Lucero, Lena Hegemann, and Antti Oulasvirta. 2019. May AI? Design Ideation with Cooperative Contextual Bandits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 633, 12 pages. https://doi.org/10.1145/3290605.3300863
[40] Paul A Kolers. 1969. Some formal characteristics of pictograms. *American scientist* 57, 3 (1969), 348–363.
[41] Aaron Marcus. 1984. Corporate identity for iconic interface design: The graphic design perspective. *Interfaces in Computing* 2, 4 (1984), 365–378.
[42] Siné JP Mcdougall, Martin B Curry, and Oscar de Bruijn. 1999. Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, & Computers* 31, 3 (1999), 487–519.
[43] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. 3111–3119.
[44] George A. Miller. 1995. WordNet: A Lexical Database for English. *Commun. ACM* 38, 11 (Nov. 1995), 39–41. https://doi.org/10.1145/219717.219748
[45] Murli Nagasundaram and Alan R Dennis. 1993. When a group is not a group: The cognitive foundation of group idea generation. *Small Group Research* 24, 4 (1993), 463–489.
[46] NetworkX. 2020. *NetworkX*. https://networkx.org/
[47] Mark EJ Newman. 2008. The mathematics of networks. *The new palgrave encyclopedia of economics* 2, 2008 (2008), 1–12.
[48] American Institute of Graphic Arts. 1974. *Symbol Signs*. National Technical Information Service.
[49] Masataka Ono, Makoto Miwa, and Yutaka Sasaki. 2015. Word embedding-based antonym detection using thesauri and distributional information. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 984–989.
[50] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 1532–1543.
[51] Robert Pienta, Fred Hohman, Alex Endert, Acar Tamersoy, Kevin Roundy, Chris Gates, Shamkant Navathe, and Duen Horng Chau. 2018. VIGOR: Interactive Visual Exploration of Graph Query Results. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 215–225. https://doi.org/10.1109/TVCG.2017.2744898
[52] Noun Project. 2020. *Noun Project*. https://thenounproject.com/
[53] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. 2017. IdeaWall: Improving Creative Collaboration through Combinatorial Visual Stimuli. 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, 594–603. https://doi.org/10.1145/2998181.2998208
[54] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. 2016. *Designing the user interface: strategies for effective human-computer interaction*. Pearson.
[55] Shutterstock. 2020. *Shutterstock*. https://www.shutterstock.com/
[56] Reinier H. van Leuken, Lluis Garcia, Ximena Olivares, and Roelof van Zwol. 2009. Visual Diversification of Image Search Results. In *Proceedings of the 18th International Conference on World Wide Web* (Madrid, Spain) *(WWW '09)*. Association for Computing Machinery, New York, NY, USA, 341–350. https://doi.org/10.1145/1526709.1526756
[57] Hao-Chuan Wang, Dan Cosley, and Susan R. Fussell. 2010. Idea Expander: Supporting Group Brainstorming with Conversationally Triggered Visual Thinking Stimuli. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 103–106. https://doi.org/10.1145/1718918.1718938
[58] Hao-Chuan Wang, Susan R. Fussell, and Dan Cosley. 2011. From Diversity to Creativity: Stimulating Group Brainstorming with Cultural Differences and Conversationally-Retrieved Pictures. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (Hangzhou, China) *(CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 265–274. https://doi.org/10.1145/1958824.1958864
[59] Jinxi Xu and W. Bruce Croft. 2017. Quary Expansion Using Local and Global Document Analysis. *SIGIR Forum* 51, 2 (Aug. 2017), 168–175. https://doi.org/10.1145/3130348.3130364
[60] Zheng-Jun Zha, Linjun Yang, Tao Mei, Meng Wang, Zengfu Wang, Tat-Seng Chua, and Xian-Sheng Hua. 2010. Visual Query Suggestion: Towards Capturing User Intent in Internet Image Search. *ACM Trans. Multimedia Comput. Commun. Appl.* 6, 3, Article 13 (Aug. 2010), 19 pages. https://doi.org/10.1145/1823746.1823747
[61] Nanxuan Zhao, Nam Wook Kim, Laura Mariah Herman, Hanspeter Pfister, Rynson W.H. Lau, Jose Echevarria, and Zoya Bylinskii. 2020. ICONATE: Automatic Compound Icon Generation and Ideation. 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/10.1145/3313831.3376618