COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 41 (2017) 1450–1484

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12418

Familiarity, Priming, and Perception in Similarity Judgments

Laura M. Hiatt, J. Gregory Trafton

U.S. Naval Research Laboratory, Washington, D.C.

Received 30 April 2015; received in revised form 2 June 2016; accepted 13 June 2016

Abstract

We present a novel way of accounting for similarity judgments. Our approach posits that similarity stems from three main sources—familiarity, priming, and inherent perceptual likeness. Here, we explore each of these constructs and demonstrate their individual and combined effectiveness in explaining similarity judgments. Using these three measures, our account of similarity explains ratings of simple, color-based perceptual stimuli that display asymmetry effects, as well as more complicated perceptual stimuli with structural properties; more traditional approaches to similarity solve one or the other and have difficulty accounting for both. Overall, our work demonstrates the importance of each of these components of similarity in explaining similarity judgments, both individually and together, and suggests important implications for other similarity approaches.

Keywords: Similarity; Perceptual Similarity; Priming; Associative Learning; Cognitive Simulation

1. Introduction

Similarity is a critical and pervasive part of human cognition (Medin, Goldstone, & Gentner, 1993). Similarity measures, for example, are integral to object categorization and classification (Nosofsky, 1992). Similarity is also pervasive in problem solving (Novick, 1990), decision-making (Medin, Goldstone, & Markman, 1995), and memory (Roediger, 1990). One interesting aspect of similarity is that asymmetries can arise when making similarity judgments, even of very simple perceptual stimuli (Rosch, 1975; Tversky, 1977). In the past, these asymmetries have been explained in several different ways. Rosch (1975) argued that similarity is based on mapping stimuli onto one another and, intuitively, non-prototypical stimuli map more easily onto prototypical stimuli than vice versa, causing an asymmetry. Tversky (1977) argued that asymmetry is due to weighted

Correspondence should be sent to Laura M. Hiatt, U.S. Naval Research Laboratory, 4555 Overlook Ave SW, Washington, DC 20375. E-mail: laura.hiatt@nrl.navy.mil

feature matching, where the salience of features in the current context determines their weight; others have also found general support for this approach (Glucksberg & Keysar, 1990; Medin et al., 1993).

These two explanations, however, assume that either there is a clear, inherent prototype among the stimuli, or that stimuli have features with a clear, inherent order of saliency (e.g., preferring the regularity of horizontal and vertical lines, or preferring forms that are symmetric). They do not, however, provide any concrete explanations for why these abstract measures may lead to prototypicality or saliency for simple stimuli with few features, such as blocks of different colors.

The issue is further complicated when stimuli become more complex, such as when comparing pairs of objects with shapes or symbols. In cases such as these, a variety of evidence strongly supports structure-based analyses as key to judging similarity (Goldstone, 1994b; Larkey & Markman, 2005). Such approaches involve explicitly identifying features and objects that correspond between stimuli; the quality of the mapping between two stimuli is then the basis for a similarity rating.

As part of our work, we consider three experiments that explore this debate. The first is by Polk, Behensky, Gonzalez, and Smith (2002), who present an experiment that involves perceptual stimuli where the only feature is color, and where the color hues are fairly similar. Despite this simplicity, the experiment showed a significant asymmetry in similarity judgments between different colors when the colors were presented with different frequencies during an irrelevant training task: Colors that had been trained on *less* often were considered more similar to colors that had been trained on *more* often than the other way around. Additionally, stimuli overall became more similar to one another as the experiment progressed. Many prototype and feature matching theories have difficulty explaining this data because of the lack of features to comparatively weight, and the lack of an inherent a priori prototype. Because of the lack of structure in this experiment, structure-mapping based approaches also have trouble accounting for it.

The second experiment involves more complicated visual stimuli, where participants rate the similarity of two pairs of objects (Larkey & Markman, 2005). The pairs differ in the colors, shapes, and spatial orientation of the individual objects. The pattern of the similarity rankings of these stimuli sheds light on the relative importance of different facets of the stimuli's similarity; for example, features shared across pairs (such as each pair having a red object, but with different shapes) as compared to objects shared across pairs (such as each pair having a red square). A third experiment is also considered that manipulated the features of schematic, bird-like stimuli (Goldstone, 1994b). Due to the richer nature of these stimuli, approaches that do not account for structure (such as prototype- and feature-based theories) have difficulty explaining these results.

We will show that the data from all three experiments can be accounted for in a single approach that uses a novel way of looking at similarity. Our approach posits that similarity stems from three main cognitive sources: familiarity, priming, and inherent perceptual likeness. The first, familiarity, represents the concept's strength in memory, based on its frequency and recency of use (Anderson, 2007). The second, priming, represents a concept's relationship with other concepts in memory, based on its past experiences with the concepts both together and apart (Anderson, 1983; Böhm & Mehlhorn, 2009). Perceptual likeness, the third value, represents similarity stemming purely from stimuli's surface appearance. For this measure, we utilize a standard measure for measuring color similitude (Breslow, Ratwani, & Trafton, 2009; Breslow, Trafton, & Ratwani, 2009), as well as a measure for shape resemblance derived from the biologically plausible vision system LVis (O'Reilly, Wyatte, Herd, Mingus, & Jilk, 2013).

Other work has used priming and familiarity for similarity in more abstract terms. The neural network written by Polk et al. (2002) relies on familiarity-based activation patterns to produce asymmetries in similarity ratings. Tversky (1977)'s discussion of salience, and Rosch (1975)'s on prototypicality can also be seen as broadly touching upon familiarity or priming in similarity; other accounts suggesting priming also exist (Kozima & Furugori, 1993; Ulhaque & Bahn, 1992). Some structural alignment approaches (e.g., Goldstone, 1994b; Larkey & Love, 2003) also use a form of priming between features and objects that are used to determine similarity. Their use of these measures, however, is more limited in scope than those we consider (for example, the amount of priming between two features is not learned over time based on experience with different stimuli); further, none of these approaches consider the purest, most basic aspect of similarity, perceptual likeness (Smith & Heise, 1992). Other accounts do consider perceptual likeness and familiarity, but not priming (Petrov & Anderson, 2005). Our work, in contrast, aims to unify many of these different viewpoints by explaining them with the same set of underlying mechanisms.

In sum, the primary contribution in this article is our claim that these three cognitive measures, familiarity, priming, and perceptual likeness, together comprise the underlying components of similarity. To support our argument, we perform an analysis of the relative, and combined, contribution of these measures to similarity ratings in the three experiments introduced above. Viewing similarity in this way has the additional contribution of providing an alternate explanation for structural alignment of simple stimuli: In our approach, structural effects arise naturally out of implicit priming effects. This suggests that the effort required for aligning structure representations when determining similarity can be deferred to more complex cases, leaving simpler cases to be taken care of as part of the process of perceiving and representing the stimuli. After introducing our measures, and describing their fit to experimental data, we discuss this further, as well as discuss how our approach allows us to better relate existing work in similarity to one another.

2. Similarity as priming, familiarity, and perception

Our approach posits that similarity derives from three key measures: familiarity, priming, and perceptual likeness. Familiarity represents how experienced one is with a concept; in other words, how strong the concept is in memory. Familiarity changes over time and with experience, weakening if the concept is not thought about often, and strengthening if it is. Intuitively, those items that are more familiar in memory are often viewed as having higher similarity than less familiar concepts.

Priming represents the strength of the relationship between an item and other items in memory, such as other items being looked at. The strength between items is learned over time, and it is based on how often items are thought about together (strengthening their association) or separately (weakening it). Priming is also directional and potentially asymmetric, in that two items can prime one another to different degrees.

Perceptual likeness, in turn, indicates the similarity of two items' surface appearances, such as how close their colors are or whether they are the same shape. This measure is static and does not change with experience. Perceptual likeness plays an important part of similarity by capturing trends and effects that are dependent purely on perceptual differences, such as color and shape.

Note that our work does not make any strong claims about how familiarity, priming, and perception individually, or combined, are transformed into an overall similarity rating; we currently leave that work to others (e.g., Petrov & Anderson, 2005). Our goal, in contrast, is to investigate how these components account for similarity, and how, together, they can strongly explain many of the effects seen across a broad array of experiments.

We next describe in more detail the specifics of our similarity account, as well as the framework in which we implement it, ACT-R/E (Trafton et al., 2013). ACT-R/E is an embodied version of the ACT-R cognitive architecture (Anderson, 2007). By grounding our approach in this existing, well-studied theory of cognition, we support our overarching desire to explain cognitive phenomena using cognitive process models, as well as connect our work on similarity with a broad spectrum of other work done within the ACT-R framework. We discuss these points further in the general discussion.

2.1. *Cognitive framework for similarity*

ACT-R is an integrated, process-level theory of human cognition in which a "production system operates on a declarative memory" (Anderson, Bothell, Lebiere, & Matessa, 1998). The specifics of the operations depend on the levels of activation of individual items in memory, which determine what memories will next become the focus of attention. This, in turn, affects how the production system operates, potentially altering the future contents of declarative memory and future memory activation levels. Models written in this framework capture the cognitive processes that people go through as they undergo tasks, and can produce behaviors that can be quantitatively and qualitatively compared to experimental data.

2.1.1. *Memories and activation*

Items in declarative memory are stored as slot-based, schema-like structures. They can be used to represent a model's goals, intermediate problem-state representations, or more conceptual fact-based knowledge. Their slot values, or features, can be either primitive (like strings or numbers) or pointers to other items in memory, potentially allowing a hierarchical-like structure to occur. Central to ACT-R's theory of cognition is its account of memory activation, which has been shown to be a very successful predictor of human memory across a variety of domains (Anderson, 1983; Anderson et al., 1998; Schneider & Anderson, 2011).

Intuitively, the activation of items in declarative memory depends both on how much the item has been thought about in the past, as well as how related the item is to other memories that are currently the focus of attention. Activation consists of three primary components: activation strengthening, spreading activation, and noise. Activation strengthening is a function of how frequently and recently the memory has been thought about in the past, and it represents the model's *familiarity* with a concept. It is designed to represent the activation of an item in memory over longer periods of time. Spreading activation, on the other hand, is temporary and context dependent, allowing memories that are currently the focus of attention to activate, or *prime*, other related items for short periods of time. Noise is a random component added in to model the noise of the human brain. They are combined according to the following equation (Anderson, 2007):

$$A_i = AS_i + \sum_j W_j S_{ji} + \varepsilon \tag{1}$$

where A_i is the total activation of item i, AS_i is the activation strengthening of item i, $W_j\ S_{ji}$ is activation spread from item j to item i and sums over all items j that have an outgoing association with i, and ε is noise. Because the default in ACT-R is to exclude noise, and its presence does not affect our results, we ignore noise in this article.

Activation strengthening of a memory item i is calculated according to Anderson (2007):

$$AS_i = \ln\left(\sum_{r=1}^{R} {t_r}^{-d}\right) \tag{2}$$

where R is the number of times item i has been *referenced* (e.g., was the focus of attention, or was explicitly thought about) in the past, t_r is the time that has passed since the rth reference, and d is the strengthening learning parameter, which defaults to 0.5. This equation implies that a memory's activation strengthening grows quickly in the early stages of learning, and more slowly once it is already familiar. More subtly, a side-effect of this equation is that intermediate problem-state representations have an undefined activation strengthening since they have not been thought about in the past. Once (or if) the finalized forms of these intermediate representations are added to declarative memory, they have a calculable activation strengthening. This side-effect is important for the second and third experiments we model.

Spreading activation is spread along *associations* between memories. In addition to considering what items are being referenced at any given time, it also considers what items are in the current *context*. The current context consists of both those items being referenced, as well as the set of items in slot values of the items being referenced that are under consideration. For example, when referencing a goal to compare blocks and considering a color in one of its slots, the goal would be referenced, and both the goal itself, as well as the color, would be in the context.

Associations are directional, and they are created from an item j to an item i when item j is in the current context when item i is being referenced. Note that if both j and i are being referenced, an association is created in each direction. Once established, associations have a corresponding strength value which affects how much activation is passed along the association from item j to item i. Association strengths, intuitively, reflect how strongly item j, when currently being referenced or in context, predicts that item i will be referenced next.

The equations for the associative strength from an item j to an item i in memory are (Harrison, 2014):

$$S_{ji} = mas \cdot e^{\frac{-1}{al \cdot N_{ji}}} \tag{3}$$

$$N_{ji} = \frac{f(C_j R_i)}{f(C_j) - f(C_j R_i) + 1} \tag{4}$$

These equations reflect two parameters: *mas*, the maximum associative strength parameter; and *al*, the associative learning rate. Neither of the two parameters has a default value. The function f is a count function, tallying the number of times that item j has been in the current context, either independently ($f(C_j)$) or at the same time that item i has been referenced ($f(C_j R_i)$).

S_{ji} approaches its maximum value of *mas* when N_{ji} tends to infinity, and it approaches its minimum value of 0 when N_{ji} tends to 0. Intuitively, this means that an association from item j to item i is strengthened when j and i are referenced or (or i referenced while j is in context) at the same time; conversely, it is weakened when j is referenced (or in context) without i. Note that while associations may be present in both directions (i.e., item j activates item i and vice versa), the associations may be of different strengths if the items have not always been referenced or been in context at the same time, or with the same frequency.

These equations are a non-standard adaptation of ACT-R's canonical Bayesian-based priming mechanisms (Anderson & Lebiere, 1998; Anderson & Reder, 1999). We use this adaptation in order to account for the large numbers of associations and objects needed by the experiments we consider here, which ACT-R's original formulation is unable to do. These adapted equations have been successful in modeling associations and priming across a variety of domains (Harrison & Trafton, 2010; Hiatt & Trafton, 2015a,b; Lawson, Hiatt, & Trafton, 2014; Trafton et al., 2013).

Spreading activation sources from a model's goal. The goal has a fixed amount of source activation which it first divides, equally, among all items j which have an *outgoing* association with the goal item i (such as a slot value of item i, or an item that has co-occurred with item i in the past). Note that this first step is in the opposite direction of spreading activation—the source activation is divided among items with *incoming*, not outgoing, connections to the goal.

These j items then use this allocated activation, W_j, as the basis of spreading activation along all of their outgoing associations; the higher W_j an item has, the more activation it spreads to the connected items. Ultimately, each item i receives $\sum_j W_j S_{ji}$ in activation from each item j that has an outgoing association to it, as indicated in Eq. 1.

The two parts to priming, W_j and S_{ji}, allow it to capture different aspects of similarity. For example, it can capture more contextual similarity (due to different allocations of source activation), where the similarity of items' change a depending on the context in which they are situated. It can also capture more prototypical- and structure-based similarity (due to differences in translating source activation to spreading activation). As we will show, this allows models in this framework to explain a variety of effects in similarity ratings.

2.1.2. Perceptual likeness

To quantify the perceptual likeness, or resemblance, between two colors, we rely on a standard measure of color similitude proposed by Breslow, Ratwani, et al. (2009). They introduced a component to ACT-R which supports high-level color processing that can detect both color similitude, as well as brightness differences between colors. It is based on the CIEDE2000 algorithm (CIE, 2001), and it has been shown to match well with human participant data.

To quantify the likeness of shapes, we rely on LVis, the vision model of the Emergent/Leabra system (O'Reilly & Munakata, 2000; O'Reilly et al., 2013). LVis is a neural network-based approach to biologically plausible computer vision. LVis does not consider color as part of its analysis; it pays attention only to the outlines of objects. Its network has three hidden layers, represented as matrices, that are organized similarly to the human visual cortex. In particular, its highest layer, the IT layer, has a *view-specific* encoding, where different parts of the matrix respond selectively to features like corners, angles, and curves with specific orientations. The likeness between two images (such as the outlines of two shapes) is calculated using a simple Euclidean distance measure between their numeric IT matrices. In some ways, this is reminiscent of work by Vinokurov, Lebiere, Herd, and O'Reilly (2011), who represented LVis IT matrices as concepts in ACT-R and classified them using partial-matching and blending algorithms.

If each stimulus only has one color, then using the color likeness measure is straightforward. If each stimulus has more than one, we make the assumption that all pair-wise combinations of color should be evaluated. For shape likeness, we assume that the shapes are evaluated in concert; that is, one IT matrix is constructed for the first stimulus as a whole, and it is compared to the IT matrix for the other stimulus.

2.1.3. Perception and action

ACT-R supports interacting with the world via limited perception and action. Models can view very simple items (such as colored blobs and text) on a simulated computer monitor, and they are provided with the items' appropriate symbolic representations. Models can act on the world by pushing buttons on a simulated keyboard.

3. Rating similarity of color-based stimuli

The first experiment we model in this paper studied asymmetries in similarity ratings of simple perceptual stimuli. There were three phases to the experiment: a pre-test phase, a training phase, and a post-test phase (Polk et al., 2002). In the pre-test phase, participants viewed two patches of different colors and were asked to rate their similarity on a scale of 0 to 9 (0 as highly dissimilar, 9 as highly similar). The colors were five different hues of green and five different hues of blue, designated as *blue1*, ..., *blue5*, and *green1*, ..., *green5*. Greens and blues were never compared to each other; only hues of the same color were shown concurrently. During a trial, the stimuli were presented as part of a text question that emphasized directionality in the judgment: "How similar is (left color patch) to (right color patch)?" Underneath the color blocks were the labels "Blue1" and "Blue2" (or, if appropriate, "Green1" and "Green2"). Each block was 140 × 140 pixels, and the sentence was centered both horizontally and vertically. Once a user entered a rating (by pressing a key from "0" to "9"), the screen was cleared for 500 ms before the next comparison appeared. Each pair of colors was presented twice in each direction for a total of four times each. The order in which the pairs were presented was randomized, except that the same hue was not present in consecutive trials.

In the middle, training, phase, participants saw two patches of the same hue and color but different sizes (125 × 125, 131 × 131, 138 × 138 and 144 × 144 pixels, appearing with equal probability) and were asked to specify which patch was larger. The key part of the experiment is that, during the training phase, two of the five green hues and two of the five blue hues were presented 10 times more frequently than the others, 110 times instead of 11. Half the participants, called "group 1," saw *blue1*, *blue2*, *green1*, and *green2* with a higher frequency; the other half, "group 2," saw *blue4*, *blue5*, *green4*, and *green5* presented more often.

The third phase was a second testing phase that was an exact replication of the first phase. Thirty-five participants' data were analyzed. For more details on the experiment, see Polk et al. (2002).

The experiment produced two main results. First, the similarity ratings were significantly higher in the post-test than in the pre-test. Second, the pre- and post-test results each exhibited different characteristics for "forward" (less frequent color on the left, more frequent color on the right) versus "backward" (more frequent color on the left, less frequent color on the right) comparisons. In the pre-test, the ratings were symmetric, with no significant difference between ratings made in the different directions. In the post-test, however, the similarity ratings showed a significant asymmetry effect: Specifically, forward comparisons were judged as significantly more similar than backward comparisons. There were also non-significant trends in the data reflecting differences stemming from color and group; specifically, group 1's ratings were lower, overall, than group 2's, and the ratings for green were, overall, lower than for blue.

3.1. Model

Our model of this experiment is quite simple. It starts out with no *a priori* declarative memories, but with the procedural knowledge it needs to complete the task. At the beginning of each trial, the model has the goal of completing the trial by rating the similarity of the two objects, or comparing their heights, as appropriate. Colors are represented as memory items with numeric slots for the colors' RGB values, a very simple way to represent them. Blocks are represented simply, as well, including slots for their location on the computer "screen" as well as their color.

For each trial during the two testing phases, the model follows the experiment directions by looking at the color block on the left. Once the model sees the left block, it retrieves the block's color from memory and looks for the block on the right. Then, when the model sees the block on the right, it retrieves that block's color from memory as well. Once the model is thinking about both colors, it proceeds to rate the colors' similarity using familiarity, priming, and perception.

Because of the comparison directionality inherent in the study, when considering familiarity and priming, the model considers only the stimulus being compared to; here, the stimulus on the right. To determine its rating, then, the model considers the perceptual likeness of the two colors, the familiarity of the right color, and the priming of the right color. More specifically, the model calculates the RGB similitude of the two colors and looks at the total activation (both strengthening and spreading) of the right color at the time of the judgment. Once the model has determined these measures, it presses a key to finish the trial and waits for the next stimuli to appear.

Note that color is considered here by familiarity, priming, and perceptual likeness in qualitatively different ways. Familiarity and priming consider color symbolically, in terms of the *concept* of the color. Perceptual likeness, in contrast, considers colors at a lower level, by looking, mathematically, at the similitude of their RGB values.

During a training trial, the model first looks at the color block on the left. While continuing to think about this block, it looks for another block of the same color. Once it sees the right block, the model compares their heights and responds accordingly.

During each trial, many associations are created between the many items involved. Key to our discussion here, during a testing trial, associations are created (or strengthened) in both directions between each of the colors and the current goal, as well as between the two colors. During a training trial, an association is created (or strengthened) from the color to the current goal. Fig. 1 shows this in diagrammatic form.

In terms of parameters, the associative learning rate, which affects the rate at which associations are strengthened, was set to 6.5, which represents a fairly brisk rate of learning. There is no standard value for this parameter. The strengthening learning parameter was set to 0.4 instead of its default of 0.5. All other parameters were set to their default values.

3.1.1. Model explanations

First, familiarity accounts for why later comparisons are, overall, more similar than earlier comparisons. In the beginning of the experiment, colors do not have a very high

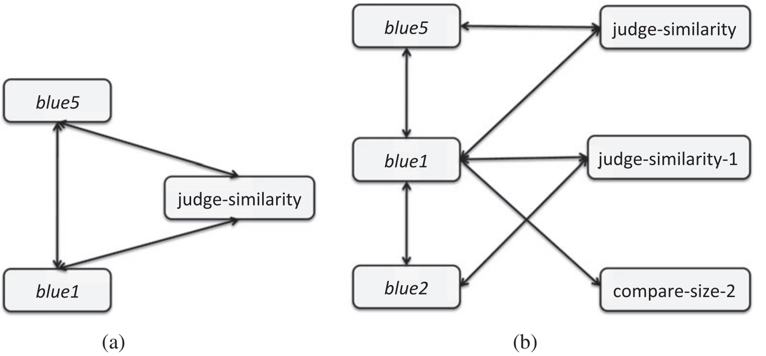


Fig. 1. Key associations at various phases of model execution. Here, the model performed two pre-test trials (with colors *blue5*/*blue1*, and *blue1*/*blue2*, respectively), and one training trial (where *blue1* is the color). Note that in order to maintain clarity, this diagram is simplified from the model's actual associative network (e.g., it does not contain associations involving blocks, which do not affect priming here). (a) Key associations after 1 testing trial. (b) Key associations after 2 testing trials and 1 training trial.

strengthening activation. During the pre-test, however, their activation increases as the colors are seen and considered many times. Throughout the training phase, the colors' strengthening activations decay a little, since the colors are not explicitly referenced. Then, during the post-test, strengthening activation values again increase, surpassing their pre-test values and leading to higher overall similarity ratings in the post-test than in the pre-test. Since all colors are shown equally during the pre-test and post-test phases, however, familiarity does not account for any asymmetry effect or any other color-specific effects.

As we have mentioned before, associations between two items can have different strengths in the two different directions, potentially causing asymmetries to arise in the amount of activation spread along them. Consider, on an intuitive level, Fig. 1b. Here, as its greater number of associations implies, *blue1* has been in context more times than *blue5*; this means that the association *blue1* → *blue5* is weaker than the association *blue5* → *blue1*. And while *blue1* and *blue5* are primed by items other than each other, such as from the goal, the amount of this other priming is the same (subject to ordering affects). Therefore, *blue1* will, on average, receive more priming than *blue5* due to the fact that it has appeared, in the past, in more contexts.

As a result of this asymmetry, priming in the model explains the different effects for the pre- and post-tests. For the pre-test, the model suggests that differences in similarity ratings of forward and backward comparisons are based solely on ordering effects of the stimuli. Given enough participants, these ordering effects average out over time to result in pre-test forward and backward comparisons that are equal. For the post-test, however, less frequent colors will prime more frequent colors more than the frequent colors will prime them back; that is, priming in the model accounts for how forward comparisons are ranked as more similar than backward comparisons.

Perceptual likeness, in turn, accounts for color-specific effects stemming from the differences between the green and blue hues. It does not explain differences depending on the direction of the comparison, or on whether the test is a pre- or post-test.

3.1.2. Model fit

In addition to the experimental results published in the original article (Polk et al., 2002), we also examined more detailed aggregate data provided to us by the authors. The data included the averages, for each participant, of ratings for trials of each condition (e.g., the average rating for each participant of all pre-test forward trials of blue hue, etc.). Since our model is sensitive to the order in which stimuli are presented, we used our model to simulate data from 1,000 experimental runs in order to allow effects to better converge on the model's true predictions. Each experimental run was generated according to the original study's methodology.

One issue with modeling similarity studies is the fact that they often use Likert rating scales as dependent measures. Our goal is to understand the possible existence and strength of the relationships between familiarity, priming, perception, and similarity; therefore, we used a theoretically light method (linear regression) of converting these measures' values to Likert ratings. Specifically, we assumed each participant has their own individual transformation function between the three similarity measures and their corresponding similarity ratings. We therefore created a linear regression model for each participant that best matched the three measures, averaged across the 1,000 model runs, to each participant's aggregate similarity ratings. We also created analogous linear models that translated each of the three measures, on their own, to each of the participants' similarity ratings, in order to investigate their individual contributions to explaining the data.

When looking at the combined contribution of familiarity, priming, and perceptual likeness, we begin by summing the familiarity and priming activation values into a single activation value. This assumption is imposed, and supported, by the cognitive framework we use: Part of ACT-R's theory is that activation is always considered as a whole. Then, color similitude is considered a second independent variable for the linear regression model. All together, familiarity, priming and perceptual likeness capture the data extremely well, with $R^2 = .924$, as shown in Fig. 2.

Both familiarity and perceptual likeness, alone, were significant predictors of the human data (familiarity: $p < .05$; color: $p < .01$). Familiarity captured the large difference between the pre- and post-tests, but it did not capture either the differences in conditions stemming from color or the asymmetry in the post-test (Fig. 3a). In contrast, perceptual likeness captured the differences between the blue/green and group1/group2 conditions, but it did not capture any of the effects between the pre- and post-tests, or between forward and backward comparisons (Fig. 3b).

Priming, in contrast, strongly explains the post-test asymmetry. Despite this, it was not found to be a significant predictor of the data as a whole ($p = .231$), in large part because it does not capture the significant increase in ratings between the pre-test and the post-test (Fig. 4). Importantly, however, because it does capture the post-test asymmetry, it significantly increases the fit of our approach, overall, to the data ($R^2 = .877$ without priming and, as reported earlier, .924 with). Additionally, as we will see, it proves to be an important part of our account for Experiment 2 and 3. Overall, therefore, we conclude that familiarity, priming, and perceptual likeness are all important pieces of similarity in this experiment and, together, do an excellent job of accounting for the data.

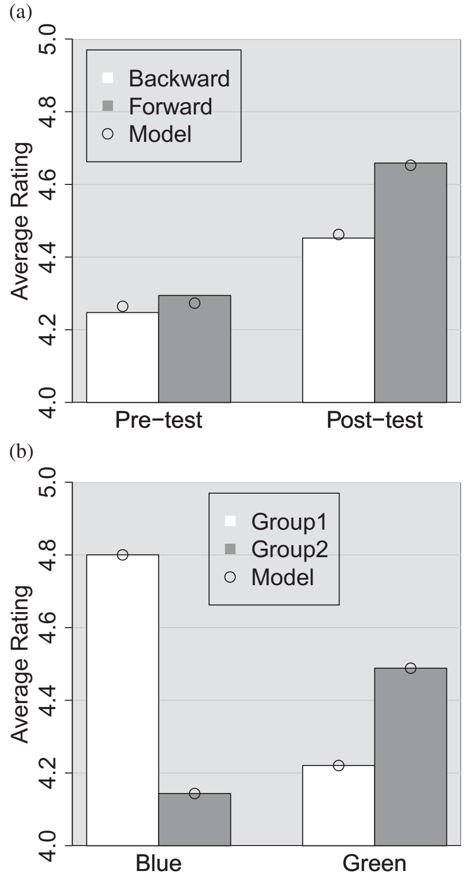


Fig. 2. Average similarity ratings of conditions separated by frequency manipulation and color. Graphs include data from the experiment itself, as well as the ratings generated by the model. (a) Average similarity ratings of various conditions of forward and backward comparisons for both the pre-test (before stimuli frequency was manipulated) and the post-test (after stimuli frequency was manipulated). (b) Average similarity ratings of various conditions of group1 and group2 comparisons for both blue and green hues.

3.2. Discussion

In this section, we described a model that uses our account of similarity, including familiarity, priming, and perceptual likeness, to explain the similarity ratings of different pairs of simple perceptual stimuli, and showed a strong account of the experimental data presented by Polk et al. (2002). First, familiarity explains why the ratings increase over time. As the stimuli appear repeatedly throughout the experiment, the stimuli's familiarities are strengthened, leading to higher similarity ratings. The second significant effect, the post-test asymmetry, is explained via priming. Perceptual likeness was also found to be a strong predictor of the data, accounting for the various trends in the data between colors and groups.

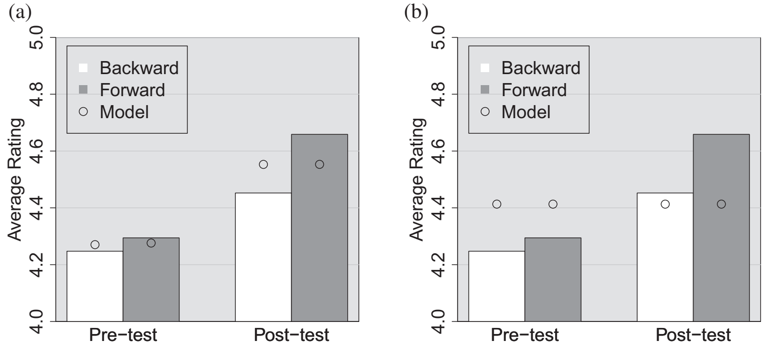


Fig. 3. Average similarity ratings of forward and backward comparisons for both the pre-test (before stimuli frequency was manipulated) and the post-test (after stimuli frequency was manipulated) for activation strengthening and color. Graphs include data from the experiment itself, as well as the ratings generated by the model. Importantly, neither activation strengthening nor color captures the post-test asymmetry. (a) Average similarity rating of the various conditions of forward and backward comparisons for both the pre-test (before stimuli frequency was manipulated) and the post-test (after stimuli frequency was manipulated) when activation strengthening was used as the sole similarity measure. (b) Average similarity ratings of the various conditions of forward and backward comparisons for both the pre-test (before stimuli frequency was manipulated) and the post-test (after stimuli frequency was manipulated) when color similitude was used as the sole similarity measure.

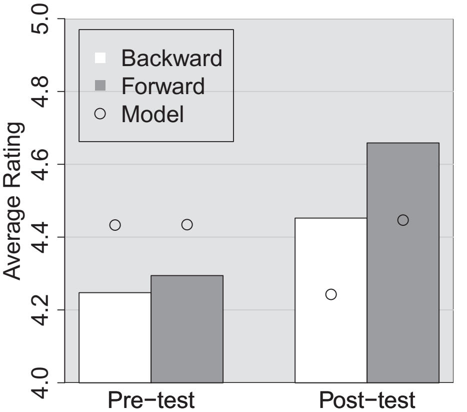


Fig. 4. Average similarity ratings of forward and backward comparisons for both the pre-test (before stimuli frequency was manipulated) and the post-test (after stimuli frequency was manipulated) for priming. Graphs include data from the experiment itself, as well as the ratings generated by the model. Priming explains the asymmetry in the post-test condition, but not the increase in ratings between the pre-test and post-test.

Because of the extremely simple nature of the stimuli, the choice of representation of the color blocks makes little difference to our model's results. Instead, they are affected by the blocks' color values (affecting the perceptual likeness measure) and the frequency with which the colors appear (affecting both the familiarity and priming measures).

To account for these results, the study's authors implemented a neural network that simulated the asymmetry effect by measuring the ease with which the network switches between different activation patterns; those that are more stable (e.g., high-frequency patterns) were easier to assimilate to. Using this single similarity measure, their model qualitatively accounts for the asymmetry effect; however, it is not described as capturing either the increased ratings over time or the color effect, nor is a quantitative comparison performed. Our work, therefore, provides a stronger account of this data, since it qualitatively and quantitatively explains all three effects in a single, coherent model.

Finally, it is important to note that our approach does not unilaterally predict that similarity increases over time; in fact, there is evidence to the contrary (e.g., Goldstone, 1994a; Levin & Beale, 2000; Sjoberg, 1972). Our model allows for this effect. The average strength of priming in many cases decreases with long-term exposure to the items being compared. Although in this experiment it is balanced by the higher activation strengthening, in other cases it could ultimately lead to lower similarity.

4. Rating similarity of pairs of perceptual stimuli

The second experiment we consider involved similarity judgments of pairs of perceptual stimuli (Larkey & Markman, 2005). Each object in a pair had two distinct features: a color (red, yellow, green, or blue) and a shape (circle, square, triangle, or star). In each trial, participants were shown two pairs of objects simultaneously and were asked to rate the pairs' similarity on a scale of 1 (low) to 6 (high). The colors and shapes of the objects in the first pair were selected randomly (and always were different for the two objects), as was whether the two objects were arranged horizontally or vertically. The second pair reflected three systematic modifications of the first pair. First, the spatial relationship between the two objects (i.e., horizontal vs. vertical) was the same on half of the trials and different on half of the trials. Second and third, the colors and shapes of the objects in the second pair were each generated based on a manipulation of the shapes and colors of the objects in the first pair.

To explain the shape and color manipulations, the authors of the study abstractly represent a pair's features using letters. The first object pair always has its colors and shapes each represented as A for the top or left object and B for the bottom or right object. In other words, the first object pair's colors are referred to as "AB" (A color for the top/left object, B color for the bottom/right object). The first object pair's shapes are referred to as "AB" in the same way.

Then, there are nine possible manipulations for each of color and shape to create the second pair: AB (nothing changed), BA (switch values), AA (copy the first value), BB (copy the second value), AC (replace the second value with a new value), CB (replace the first value with a new value), CA (replace the second value and switch values), BC (replace the first value and switch values), and CD (replace both values) (see Fig. 5).

The three manipulations (color manipulation, shape manipulation, and spatial relationship manipulation) happened independently, leading to a total number of 162 unique

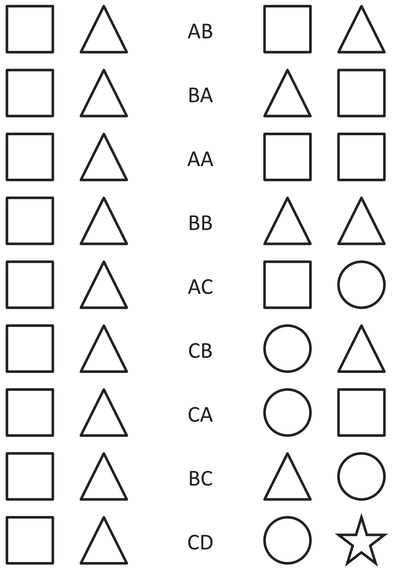


Fig. 5. The nine possible methods for manipulating the shape feature. The shapes of the first pair of objects are always different and are considered shapes "A" and "B"; with this notation, they are then methodically manipulated. The color feature (not shown) is done in an analogous way. (This figure is adapted from Larkey & Markman, 2005.)

conditions. A sample trial is shown in Fig. 6. We denote conditions as "color manipulation/shape manipulation/spatial relationship manipulation" (e.g., CA/CB/different for the manipulation in Fig. 6). In the original article, the data are collapsed across the two spatial relationship manipulations; the condition CA/CB, it follows, would denote the color manipulation CA, the color manipulation CB, and both the same and different spatial relationship conditions.

Each object was approximately 3 cm × 3 cm when displayed on the monitor. Objects within each pair were displayed as 1 cm apart. The pairs' locations were selected randomly subject to the constraint that they be 11 cm apart. One final note is that when the spatial relation was different, the object on top (or left) of the first pair was randomly chosen to "correspond" (for the purposes of the feature manipulations) to either the right or left (or top or bottom) of the second pair.

There was only one phase to the experiment, during which each participant was presented with one pair from each of the 162 conditions in random order and was asked to rate the pairs' similarity. There were 58 participants in the study. In the original article, the experiment was duplicated with texture and shape instead of color and shape, with nearly identical results; in this paper we consider only data from the experiment that involved color.

In their analysis, the authors consider only a subset of the data: They consider one group of trials where at least one feature manipulation was AB (whether for color or shape), and another group where at least one feature manipulation was BA. They also

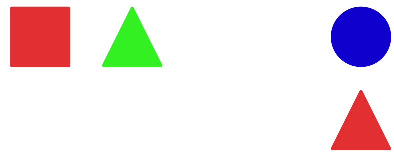


Fig. 6. A possible stimulus pair. Here, the color manipulation is CA and the shape manipulation is CB. In addition, the spatial relationship of the pairs is different. We denote this condition as CA/CB/different.

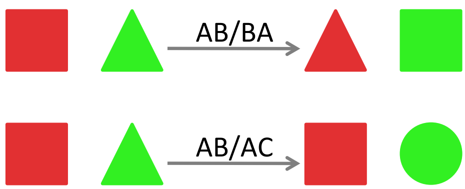


Fig. 7. Two stimuli examples. The first shows "AB/BA," where color's manipulation is "AB" and shape's is "BA"; the second shows "AB/AC," where color's manipulation is "AB" and shape's is "AC."

Table 1
Ordinal relationships between manipulations when at least one manipulation is AB (first column) and at least one is BA (second column)

AB	**BA**
AB	BA
BA	AB
{AA,BB} {AC,CB}	{AA,BB} {CA,BC}
{CA,BC} CD	{AC,CB} CD

Notes. For example, column 1 row 5, "{CA,BC}," indicates data where either: shape's manipulation was AB and color's manipulation was {CA,BC}, or shape's was {CA,BC} and color's was AB. These stimuli earned higher ratings, although not significantly so, than data where either: shape's manipulation was AB and color's manipulation was CD; or shape's was CD and color's was AB. Adapted from (Larkey & Markman, 2005).

combine the manipulations AA and BB, AC and CB, and CA and CB, arguing that they are functionally equivalent; we denote the combined conditions as {AA,BB}, etc. In our analysis, we adopt these conventions, but also provide statistics showing our fit to all of the data without combining these conditions.

The important results of this experiment, as the authors argue, are the relative orderings of AB/BA, AB/{AA,BB}, and AB/{AC,CB}, and, correspondingly, BA/AB, BA/{AA,BB}, and BA/{AC,CB}. The ordering of these conditions helps to shed light on the relative importance of different facets of the stimuli's similarity. For example, one could imagine that a full shared object between pairs (AB/{AC,CB}) would result in a higher similarity than stimuli where all features are the same but no object is shared (AB/BA) (both are represented graphically in Fig. 7); the opposite seems equally plausible.

Accordingly, the authors' primary analysis concerns the ordinal rankings of the similarity ratings for these different conditions, shown in Table 1. The first column shows the rankings for data where at least one manipulation is AB (whether for color or shape); the second, where at least one manipulation is BA. For example, column 1 row 5, "{CA, BC}," indicates data where either shape's manipulation was AB and color's was {CA/BC}, or shape's manipulation was {CA,BC} and color's was AB. The conditions in the column are sorted in decreasing order based on their average similarity ratings by participants. Horizontal lines are present between conditions that were found to be significantly different from one another. Not surprisingly, the results suggest that both shared features and shared objects influence similarity judgments.

4.1. Model

The model has a similar simple structure as the model for Experiment 1, with no *a priori* declarative memories, but with its representations and task knowledge adapted for the more complex task. The model also uses a simple representation for the stimuli that is characteristic of the hierarchical representations used by many structure-based approaches (Goldstone, 1994b; Larkey & Love, 2003). Individual objects are represented with two slots: shape and color. A pair's representation also has two slots, corresponding to the two objects in that pair. The model builds these representations in a straightforward way, creating them as it looks at the objects one by one.

For each trial, the model has the goal of rating the similarity of the two pairs of stimuli. The model begins by finding the object closest to the upper left corner of the screen. It creates an intermediate representation of it, filling in the object's color and shape as appropriate, and adds the final representation to its declarative memory. It repeats this for the second object. The model then creates an intermediate representation of the pair, retrieves each object, and adds each object to the pair's representation. After adding the first pair to its declarative memory, it repeats this process for the second pair. Before adding the second pair to declarative memory, however, it retrieves the first pair from memory so it can compare the two.

Unlike in Experiment 1, the comparisons in this study are not given an explicit direction. Therefore, when considering familiarity and priming, we summed together the strengthening and spreading activations of both pairs.

This model's measure of perceptual likeness is more complicated than the measure of the previous model due to its more complicated stimuli. Prior to the experiment starting, the LVis vision system is trained to recognize, individually, each of the four shapes involved in the stimuli. As part of perceptual likeness, the model calculates the resemblance of each of the six pairwise-combinations of the four objects' colors using the previously discussed color similitude measure. For shape likeness, the model first creates an image for each pair depicting the outline of the two shapes in the appropriate configuration, and then uses LVis to generate, and calculate the likeness between, the two images' numeric IT matrix representations (described previously). Note that while color and shape are considered by both priming and perceptual likeness, then, the measures considers them in a qualitatively different way. Priming considers color and shape symbolically, in terms of the *concept* of the color or shape. The perceptual likeness measures, in contrast, consider them at a lower level, either by looking, mathematically, at their RGB values, or by numerically considering how perceptually similar their shapes are.

So, to summarize, when comparing two stimuli, the model considers: how familiar each pair it sees is, to what degree each pair is primed, and how perceptually like each other the two pairs are. Once the model has found these measures, it virtually presses a key to finish the trial, adds the second pair to memory, and waits for the next one to begin.

During each trial, many associations are created between the many items involved. Pertinent to our discussion here, the goal becomes associated with every item associated with completing the task (e.g., colors, shapes, objects, pairs). Therefore, in trials with more unique colors and more unique shapes (such as if no two objects have the same color or shape), the goal has more incoming connections; in trials with fewer unique colors, fewer shapes and fewer distinct objects (such as if two or more objects have the same color, shape, or both), the goal has fewer incoming connections. This is illustrated in Fig. 8. Additionally, each pair has an incoming association from its component shapes, colors, and objects. This model used the same parameters as the previous model.

4.1.1. Model explanations

While, in theory, familiarity sheds light on similarity for this experiment, in practice it does so in a clumsy way. Familiarity for the second pair, first and foremost, is not meaningful at the time of judgment because it has not yet been added to declarative memory. Familiarity for the first pair, however, has, and in principle it should provide some insight into similarity for this experiment. The familiarity of the first pair at the time of judgment

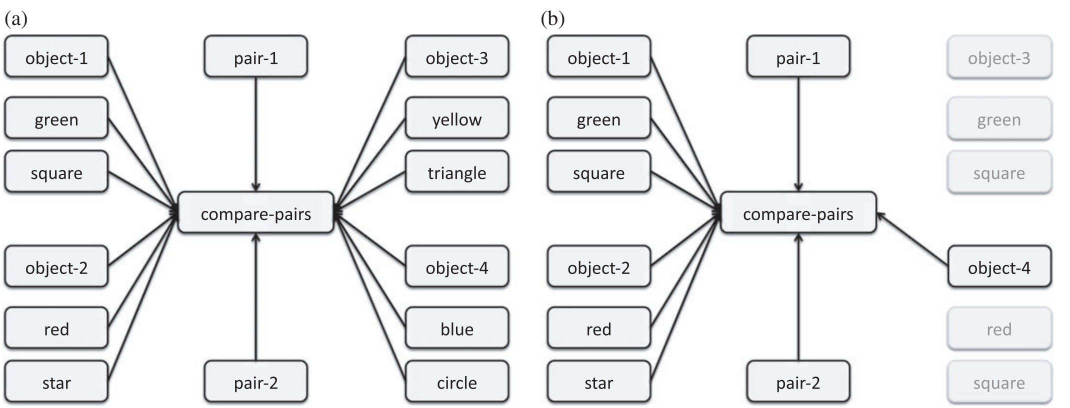


Fig. 8. Two figures showing the contrast between the goal's incoming associations in different conditions. Source activation is divided among incoming associations, so the individual memories in (a) will receive less source activation than those in (b). Note that in order to maintain clarity, this diagram shows only a portion of the model's actual associative network (e.g., it does not contain outgoing associations from the goal, or associations between colors, shapes, and pairs). (a) Incoming associations to the goal when all pairs are made of unique features (CD/CD condition). (b) Incoming associations to the goal when pairs have overlapping features (AB/AA condition). Grayed out boxes indicate repeated items, for clarity.

depends, in part, on how long it takes to build up the representation of the second pair: The faster the second representation is built, the less time will have passed since the first pair was last thought about, and so the more familiar it will be when judging the similarity of the pairs. The time it takes to build up this representation depends, in turn, on how strong the colors, shapes, and objects are in memory. Thus, in conditions with more colors, shapes, and objects in common, those items can be quickly accessed, and the second representation can be built up fairly briskly, leading to a higher familiarity for the first pair. In conditions with few common colors, shapes, and objects, the second representation takes longer to build, and the first pair will have a lower familiarity.

This account, however, is muddled because familiarity depends on an item's entire history in memory. Because of the limited number of possible colors and shapes, a trial's colors, shapes, objects, and pairs are very likely to have been seen before, leading to familiarity being very dependent not only on the shared features of the current pair of objects, but also on whether those features were present in recent trials. To put it another way, familiarity here is very dependent on ordering effects, and we expect that in reality those ordering effects will obfuscate any meaningful contribution of familiarity to understanding this experiment.

Priming also meaningfully explains the data's trends, and it does so regardless of ordering effects. Each pair, in general, has the same number of concepts that spread activation to it: the goal, and the features and objects that are part of it. The amount of spreading activation that each pair receives, then, is dependent largely on the amount of source activation allocated to each of these concepts. The source activation, in turn, depends on the total number of incoming connections to the goal.

To illustrate, in conditions like AB/AA, where the goal has a low number of incoming associations (due to repeated colors, shapes, or objects; Fig. 8b), pairs will receive the most amount of spreading activation; during a CD/CD trial (Fig. 8a), where the goal has the most number of incoming associations from non-repeated colors, shapes, and objects, pairs will receive the least amount. Priming provides much useful information, then, for similarity in this experiment—it can implicitly account for both feature matches (i.e., repeated colors and/or shapes) and higher-level object matches (i.e., repeated objects). This means that, qualitatively, the similarity in our model between two pairs depends almost entirely on the total number of unique colors, shapes, and objects present in the scene.

The exception to this characterization is when there is a repeated color, shape, or object within a pair. Here, the pair itself has fewer incoming associations (because the duplicate color, shape, or object is only connected to the pair once), but that association has a stronger strength (because the repeated item has been thought about with the pair more often). Overall, this results in a gently dampened priming value for that pair. While not critical to understanding this experiment's trends, it becomes important in the next experiment we model (Experiment 3).

Recall that the shape likeness measure compares the perceptual resemblance of the pairs of shapes. In general, then, those pairs that have the same shapes with the same spatial orientation are measured as the most similar, because their numeric IT matrices will be more similar to one another. Following that, again generally, pairs that have shapes in common are measured as more similar than those that do not. In addition, pairs with shapes that have similar visual attributes (such as a triangle and a star both having corners) are also rated as relatively more similar than those with shapes without such attributes in common (such as circle and triangle). Shape, then, accounts for the visual likeness of the pairs overall.

Color likeness, as we describe above, is based on the color resemblance of the pairwise-combinations of the objects in the two pairs of stimuli. Therefore, when considering only color, the pairs in conditions with the most uniform colors (such as the AA and BB color manipulations) are, intuitively, rated as most similar, because the pairwise-combinations of their colors will include more cases where the two colors are the same. It follows that pairs from conditions that introduce a third (or fourth) color would be rated as less similar. As with shape, color accounts for similarity stemming from the diversity of the visual scene overall, but it does not give any intuition about higher level matches.

4.1.2. Model fit

In addition to the experimental results published in the original article (Larkey & Markman, 2005), the authors also provided us with the exact stimuli that each of the 58 study participants were shown. This allowed us to very faithfully replicate the experiment to see how well our model's similarity ratings matched those of the original study participants. We therefore ran the model 58 times, once per participant in the original study, using the same stimuli that the participants saw.

We evaluated the similarity measures using the same methodology as in the previous experiment. To test the model's fit, we assumed each participant has an individual transformation function between the three similarity measures and their corresponding similarity ratings. We therefore created a linear regression model for each participant that best matched the model's three measures to each participant's individual similarity ratings. We also created linear models that translate the three measures, on their own, to each of the participants' similarity ratings, in order to look at their individual contributions. As before, when looking at their combined contribution, we summed familiarity and priming into a single activation value. When considering perceptual likeness, in addition to shape, we considered the sum of the six color similitude values, as well as the three interactions of the pairwise color values.

Using the qualitative ranking evaluation method from the original paper, our model matches the ranking data from the experiment shown in Table 1. Most important, it matches the relative rankings of the AB/BA, AB/{AA,BB}, and AB/{AC,CB} conditions, which we discussed earlier as being particularly useful in capturing the nuances of judgments in this experiment. Quantitatively, Fig. 9 shows the human participant data and the model data on the same graph. A statistical fit of our model to the data produced an R^2 of .960. When the full range of manipulations is considered, not just those with an AB or BA manipulation, our model fit the data with an R^2 of .941. Overall, then, our fit to the data was excellent.

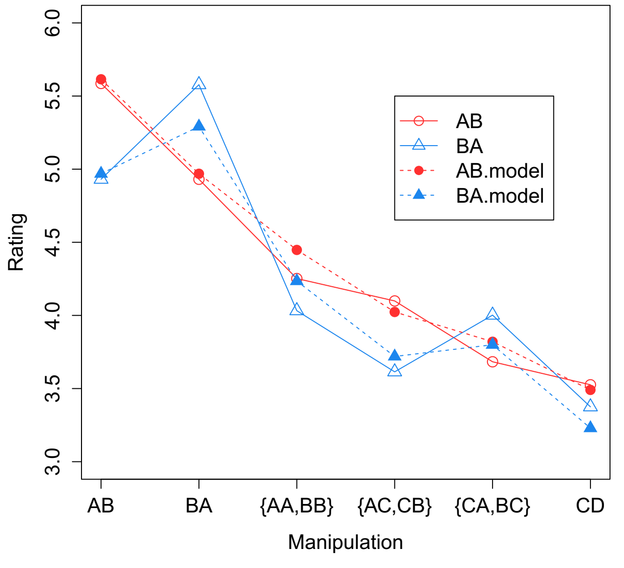


Fig. 9. Average similarity rating by feature manipulation, when the other feature manipulation is AB or BA.

Individually, each of the measures except for familiarity was a significant predictor of the human data (all with $p < .001$). As we suspected, familiarity made no meaningful contributions to the model's explanations of the data because it was too dependent on ordering effects to let its true predictions come through. Color accounted for some of the trends of the data overall, but it was unable to explain the differences between conditions where the same set of colors appears in different configurations (such as the differences between the AB/AB and AB/BA conditions; see Fig. 10). Shape suffers from a similar shortcoming.

Priming, on the other hand, does capture similarity at the object level. As Fig. 11 shows, priming highly differentiates the conditions by the presence of shared objects. In particular, priming places a heavier emphasis on conditions with two identical objects (such as for AB/AB and AB/{AA,BB}) than it does on conditions with simply identical features (such as AB/BA); this distinction, a robust effect seen across a variety of experiments in similarity, is not captured by the other measures.

4.2. Discussion

In this section, we introduced a model in which our account of similarity, using measures of familiarity, priming, and perceptual likeness, explains similarity ratings of different pairs of complex visual stimuli, and we showed that it provides a strong account for the experimental data presented by Larkey and Markman (2005). In the original study, the data were used to differentiate between several transformational and structure mapping approaches to similarity. As Larkey and Markman (2005) explain, one way of differentiating between competing structure-based approaches is in how they handle *matches in*

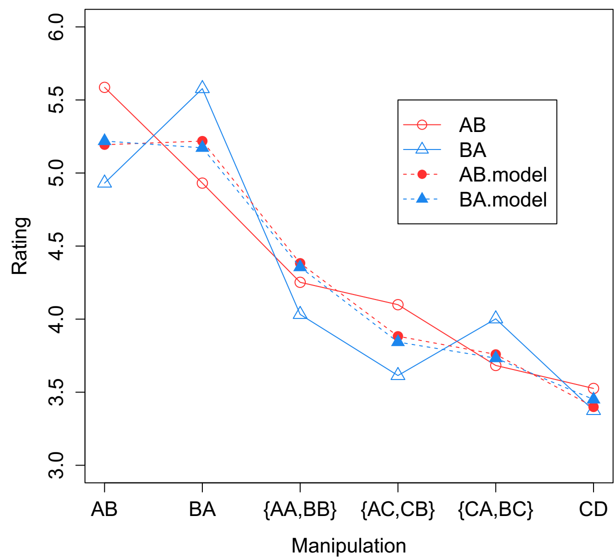


Fig. 10. Average similarity rating by feature manipulation, when the other feature manipulation is AB or BA, and where the only measure considered is color. As this graph shows, while color captures the trends overall, it is unable to account for the differences between the AB and BA results, since they have very similar coloring.

place, or MIPs, where features in an object match in the same role (such as comparing two black cars that have beige interiors), and *matches out of place*, or MOPs, where features in an object match in different roles (such as comparing a black car that has a beige interior with a beige car that has a black interior) (Goldstone, 1994b). While many agree that both MIPs and MOPs are important for determining similarity, accounts differ about their relative contribution to similarity, as well as the complexity of the analysis used to uncover the relational structures among features, roles, and objects (Falkenhainer, Forbus, & Gentner, 1989; Goldstone, 1994b; Larkey & Love, 2003). These data, with their correspondences at both the feature- and object-level, provided a good setting in which to compare these approaches.

Ultimately, in the original paper, a structure mapping approach, SIAM (Similarity as Interactive Activation and Mapping) (Goldstone, 1994b), was identified as the best match since its predicted rankings of similarity match those of the data the best. As we have stated, we believe that these rankings can be explained in a simpler way, without explicit structure mapping.

Our approach does this by implicitly capturing the nuances of MIPs and MOPs in this data without any formal structure analysis. As we described above, the perceptual likeness measures do a satisfactory job of distinguishing between conditions with MOPs and those without. Priming, in turn, implicitly accounts for both MIPs and MOPs because duplicate features and objects result in the second pair receiving higher spreading activation. And although it does not explicitly distinguish between these types of matches,

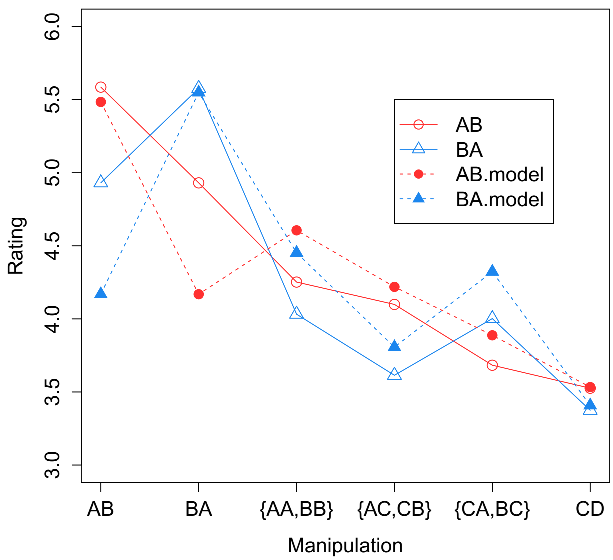


Fig. 11. Average similarity rating by feature manipulation, when the other feature manipulation is AB or BA, and where the only measure considered is priming. This graph illustrates the heavy emphasis that priming places on objects that are identical between the stimuli (indicated by the high ratings for AB/AB and BA/BA); it also illustrates its lesser favoring of MOPs.

MIPs result in an extra boost of similarity since both the underlying features and the parent object itself are duplicate. We argue, then, that our approach is stronger than that of SIAM with respect to these data, because we are able to fit the data quantitatively in addition to qualitatively, and to the full dataset instead of just to the AB and BA conditions subset.

In contrast, the similarity approach of Polk et al. (2002), considered above, is too limited to explain these data. The neural network successfully learned asymmetry based on differential exposure to stimuli, but it is unable to capture the effects here that arise out of perceptual likeness and structural similarity. In fact, given that the stimuli are seen with roughly equal frequency, it would predict, overall, that every stimulus is roughly equally similar to every other.

Similarly, SIAM struggles to explain the main effects of the first study. The priming relationships between the color blocks in SIAM would be symmetric, since they are based purely on features and structure and there are no real features to align; thus, SIAM cannot naturally capture the data's primary asymmetry effect. SIAM is also not sensitive to how often the stimuli are encountered, preventing it from capturing the data's second significant effect where ratings rise over time. SIAM's more specific view of similarity, overall, means that the approach would be unable to make meaningful explanations about the similarity ratings from the first experiment, which we are able to capture using the same approach as we do here.

We also compared our approach to a transformational distance approach that was used to model a dataset that included the same stimuli (but different conditions) as Larkey and

Markman (2005). Hodgetts, Hahn, and Chater (2009) argue that similarity is determined by counting the number of steps (such as swapping objects of features) necessary to transform one stimulus of the pair into the other. We applied their step counts to the experiment here and analyzed its ability to explain the data. Qualitatively, one of the approach's main shortcomings is its inability to adequately capture the differential influence of MIPs and MOPs on similarity. This is because it equally weighs transformations that swap features versus ones that swap entire objects; on the other hand, this allows the model to operate with no free parameters. Quantitatively, on the subset of data analyzed in the original study (where one manipulation was either AB or BA), the transformational approach achieves an R^2 = .86, which is less than our R^2 = .96; on the entire dataset, it achieves an R^2 = .87, compared to our R^2 = .94. Additionally, transformational approaches have difficulty explaining the data in the first study, because of the lack of features to transform.

Still, further study was warranted to support our argument that we can, in fact, account for the role of MIPs and MOPs in similarity. We therefore next modeled one of SIAM's cornerstone experiments. By doing so, we can further confirm the hypothesis that we are able to implicitly account for these low-level structural effects without any formal structure analysis.

5. Rating similarity of schematic stimuli

The third, and final, experiment we consider has stimuli that place a more explicit emphasis on quantifying and qualifying the influence of MIPs and MOPs on similarity. In it, participants were asked to rate the similarity of two schematic birds (Experiment 2 of Goldstone, 1994b). Each bird consisted of four connected parts: a head, upper wing, lower wing, and body. Birds differed from one another based on what symbols appeared in each part; each of the 21 possible symbols consisted of straight lines arranged in geometric patterns. In addition, all birds had identical schematic beaks and tails.

The first bird of the pair always had four unique symbols. The second bird's symbols were then constructed based on one of 15 manipulations of the first bird's symbols. Similar to the manipulations from the preceding experiment, the manipulations are described abstractly, using letters to represent the pairs' symbols, with the first pair always represented as ABCD. Table 2 shows the manipulations.

Fig. 12 shows a sample bird and symbols for the condition BACD. When applying the manipulations, the part of the bird that corresponded to each of the letters was randomly selected, as were the symbols used. The physical left-right order of the birds was also randomized. Each bird was 7.6 cm long, and they were separated by 5 cm.

Participants began with 15 practice trials, followed by 150 trials, with each manipulation being presented 10 times each, in random order.[1] In each trial, participants were shown the two birds simultaneously and were asked to rate the birds' similarity on a scale of 1 (low) to 9 (high). After responding, the screen was cleared and the next pair of birds was shown after a 2 s pause. There were 29 participants in the study.

Table 2
The 15 manipulations used to generate the schematic bird stimuli

Method	Initial Bird	Changed Bird	MIPs	MOPs
1	ABCD	ABCD	4	0
2	ABCD	ABCC	3	1
3	ABCD	ABCZ	3	0
4	ABCD	ABDC	2	2
5	ABCD	ABDZ	2	1
6	ABCD	ABYZ	2	0
7	ABCD	AADC	1	3
8	ABCD	AXDC	1	2
9	ABCD	AAYZ	1	1
10	ABCD	AXYZ	1	0
11	ABCD	BADC	0	4
12	ABCD	BADZ	0	3
13	ABCD	BAYZ	0	2
14	ABCD	BXYZ	0	1
15	ABCD	WXYZ	0	0

Note. Table adapted from (Goldstone, 1994b). MIPs, matches in place; MOPs, matches out of place.

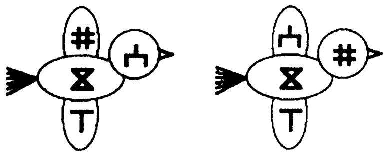
The results analyze the experiment in terms of the relative and combined contributions of MIPs and MOPs. Both MIPs and MOPs were found to be significant factors in the similarity ratings, with MIPs contributing more to similarity than MOPs. Additionally, conditions with repeated symbols presented as a special case. Birds in such conditions have an MOP with the same symbol as an MIP; we call such conditions “duplicate MOP” conditions. As Goldstone (1994b) discusses, such duplicate MOPs seem to contribute little, if any, to similarity ratings.

5.1. Model

We use here the same model as for Experiment 2, but with its stimulus representation and task knowledge marginally adjusted to account for the slightly different stimuli. Our measure of perceptual likeness was hampered because the exact symbols used in the study were not available. Therefore, we defaulted to the simple approach of calculating the likeness of two symbols as “0” if the symbols were different, or as “1” if they were the same; the likeness value of two birds was then the sum of these values for each pairwise combination of symbols. The model uses the same parameters as the previous two models.

5.1.1. Model explanations

Priming, familiarity, and perceptual likeness make the same explanations as in the previous experiment. It is worthwhile, however, to elaborate on several points in order to clarify how they account for the main effects highlighted in this experiment. First, it is appropriate here to frame the discussion in terms of MIPs and MOPs. As we explained



(a) (b)

Fig. 12. Sample birds showing the manipulation condition ABDC. (Figure adapted from Goldstone, 1994b.)

earlier, priming is sensitive to the total number of unique symbols and parts present in the scene. Priming thus suggests that MIPs and MOPs increase similarity because each reduces the numbers of unique symbols and parts in the scene. Priming also provides insight into why MIPs affect similarity more than MOPs: MIPs reduce both the number of unique symbols and parts, whereas MOPs reduce only the number of unique symbols.

We also previously discussed how the results show a special case for "duplicate MOP" conditions, where a stimulus has an internally repeated item. In such cases, the model's representation of the bird will have one fewer incoming association, since it is connected to three unique symbols instead of four. Although the duplicate association strength will be stronger than the others, overall, the additional MOP contributes little priming to the pair, predicting the dampened similarity for these conditions.

A third point of clarification is that, because of the large number of symbols, parts, and birds possible in this dataset, we expect that familiarity will be less dependent on ordering effects and will contribute more meaningfully to our model's fit.

5.1.2. Model fit

The data from this experiment are no longer available; therefore, we used the program "PlotDigitizer" (*Plot Digitizer*, 2015) to extract the values of the summary data points shown in Fig. 7 of Goldstone (1994b). Since our model is sensitive to the order in which stimuli are presented, we used our model to simulate data from 1,000 experimental runs in order to allow effects to better converge on the model's true predictions. Each experimental run was generated randomly according to the specifications of the original study.

We averaged familiarity, priming, and perceptual likeness across these 1,000 runs for each condition. We then created a linear regression model to match total activation and perceptual likeness to the study's data. Overall, the model's fit to the data was excellent, yielding an R^2 of .988. Importantly, in addition to a strong quantitative fit, we also capture the qualitative effects of the original study. As Fig. 13 shows, our account for similarity exhibits both a higher emphasis on MIPs than MOPs, as well as the dampening effect of the duplicate MOP conditions.

Individually, each of the three measures was a significant predictor of the human data (all with $p < 0.001$). Both familiarity and priming contribute to the higher weight for MIPs than MOPs; priming, in addition, accounts for the duplicate MOP conditions. Perceptual likeness also captures some of the data's trends, but, as with Experiment 2,

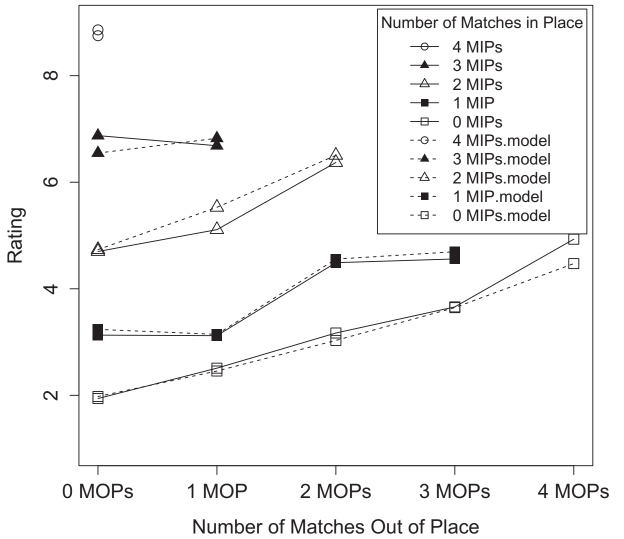


Fig. 13. Average similarity rating by the number of matches in place (MIPs) and matches out of place (MOPs).

weighs MIPs and MOPs equally and so cannot distinguish between conditions with the same symbols, but in different places (such as the ABCD and ABDC conditions).

5.2. Discussion

Here, we described a model using our account for similarity that accounts for a third similarity dataset, Experiment 2 from Goldstone (1994b). Perceptual likeness, while necessarily simplistic for this study, still provides the broad strokes of similarity increasing with shared symbols, whether MIPs or MOPs; familiarity and priming discriminate between MIPs and MOPs by more heavily emphasizing MIPs; and priming provides a strong explanation for why the similarity of duplicate MOP conditions is dampened. As before, priming and familiarity capture these effects implicitly, without any sort of formal structural analysis.

An important point in favor of our approach is that this model, as well as the model for Experiment 2, exhibits these effects and explanations even when different aspects of the model are perturbed. It still strongly fits the data if the task structure is different, such as if it retrieves symbols and parts in a different order. The specific choice of representation also seems to matter little, as long as the basic association structure we have shown here is created as part of the task strategy. Changes to other aspects like the structure of source activation also preserve the model's main effects. This lack of sensitivity to the model's specifics increases our confidence in our overall account.

In the original study, several approaches were considered as candidates for explaining the data, and as with the second experiment we consider here, SIAM was identified as the best candidate. Here, our fit is comparable to SIAM (R^2 = .988 as compared to SIAM's R^2 = .977) while also maintaining the advantages over SIAM that we have previously discussed: namely, that we are able to model the first experiment we consider in this article, which SIAM struggles with; and that our approach does not rely on a separate, formal analysis of the stimuli, but occurs naturally as part of the process of viewing and representing the stimuli to complete the task.

In addition, we argue that our approach has an advantage over SIAM because it extends more naturally to larger and more complicated situations, such as those involved with analogical reasoning. While SIAM has a fixed representation structure, our approach allows for changes in representation: If one's representation of the stimuli were to change over time, familiarity, priming, and perception would adapt and morph along with it. We discuss this point further in the following section.

6. General discussion

Here, we have proposed a novel account of similarity. The account is based on how familiar stimuli are, how much they are primed, and how perceptually similar they are. Models using our account for similarity seem to match well to how humans perform similarity judgments of both simple, perceptual stimuli displaying asymmetry effects, as well as more complicated stimuli that have structural correspondences.

Our approach furthers our understanding of similarity in three main ways. First, it performs an analysis of the relative, and combined, contribution of familiarity, priming, and perceptual likeness to similarity. Across the three studies we consider, we found priming to be the most consistent and reliable discriminators of similarity, capturing both asymmetry as well as structural effects. Perceptual likeness accounts for many of the more general trends of similarity as well. Familiarity, on the other hand, while often important, has a less consistent role in explaining the data from these studies since it is dependent on stimuli's statuses in memory and is heavily influenced by ordering effects.

Second, our approach provides an alternate account for how structural alignment occurs. It posits that, for simple stimuli, structure-based similarity can arise naturally and implicitly out of the combination of familiarity, priming, and perceptual likeness, as representations of stimuli are naturally built up in memory. Third, our approach strongly relates to others in literature and gives strong intuition for why many of these similarity effects occur by providing them with a unified underlying explanation. We further discuss these contributions, as well as other important points, next.

6.1. Similarity, perception, and learning

One aspect of our account that bears further investigation is the depth to which perception influences similarity. Perception has a fundamental role in how humans perceive things as similar in the world, and our results here confirmed its importance in similarity judgments. Some evidence suggests, however, that the strength and quality of its role seem to change over time. In the absence of strong conceptual knowledge, for example, children often rely on perceptual and superficial features to describe why objects are similar; only after more abstract representations are developed do they respond with more relational similarities (Gentner, 1988). Interestingly, our similarity account predicts a similar shift. In the absence of other knowledge, a model would rely purely on perceptual likeness; over time, the effect of perceptual likeness would be modulated somewhat as the model becomes more familiar with concepts and their associations, and familiarity and priming begin to play a larger role.

Similarly, others have argued that perceptual similarity is not a fixed variable, but rather is constantly in flux (Smith & Heise, 1992); and as people gain more experience with items in the world, the perceptual features they attend to shift. While the level of perceptual similarity being referred to in such studies is deeper than what is captured by our perceptual likeness measures, our similarity account is compatible with this view. If the features being attended to while looking at an item were to change, our perceptual likeness measures would reflect that change. Similarly, if one's *representation* of an item's visual appearance changes over time, our account predicts that the item's associations with related concepts will also change, leading to meaningful differences in familiarity and priming.

6.2. *Similarity, prototypicality, and weighted features*

The idea that similarity stems from the notion of prototypicality and/or weighted feature matching is long-standing and well-supported (Rosch, 1975; Tversky, 1977), and it has revealed many interesting effects of how humans perceive similarity. For example, work in these areas has shown asymmetry to occur not only between simple perceptual stimuli, but also between everything from complex shapes and geometric forms, to high-level concepts such as countries, foods, and physical objects (Medin et al., 1993).

Our work on similarity operationalizes and connects the different conceptual and theoretical narrations of similarity that stem from this work. Using learned associations, we are able to explain *how* and *why* prototypically affects similarity: Prototypical items are highly familiar and have asymmetric associations, both leading to higher similarity. We also explain *how* and *why* shared features between objects differentially contribute to similarity: They lead to priming between the objects, and the amount of that priming depends on how correlated they have been found to be with each other in the past. Our work, in a sense, unifies these two separate viewpoints by explaining them with the same set of underlying mechanisms.

Our work not only grounds and supports these different theories on similarity in explicit cognitive mechanisms, but also helps them to explain cases that they did not naturally extend to before. The experiment of pairs of color patches, for example, cannot be explained by these earlier theories because there is no clear prototype (no shade of the blues or greens is inherently more prototypical than another), nor are there multiple features to differentially weigh (color is the only feature). In fact, theories of prototypically and weighted feature matching would have difficulty generating any meaningful similarity values for these stimuli. Using our approach, however, we are able to explain the results of this experiment by effectively learning which colors are prototypical during the course of the experiment.

Our work also accomplishes this in a way that is compatible with some of the compelling arguments against prototypicality and weighted feature matching. Nosofsky (1991), for example, argued against that these asymmetries can rise out of biases associated with specific stimuli, as opposed to the relationship between two stimuli. This bias term strongly correlates with familiarity, which is entirely depending on a single stimulus. It also has ties to priming, since the relationship between two objects depends, in part, in their use outside of that relationship. Familiarity and priming perhaps even more strongly relate to the density portion of the distance-density model (Krumhansl, 1978), which also depends on properties of individual stimuli and their relationship with others outside of the specific judgment being made.

6.3. Asymmetries in similarity

It is also worthwhile to discuss asymmetries in similarity outside of the context of exposure frequency and prototypicality. For example, Hahn, Close, and Graf (2009) found asymmetries in a study involving morphing images into one another. When shown two images from the morph in a forward order, participants rated the images' similarity higher than they did when rating the similarity of two images shown in a backwards order. Our approach is compatible with this finding. Because priming is directional, it is affected by the order in which stimuli are perceived. Specifically, assuming some method of discretization of the video, earlier images in the morph will prime later images in the morph more than the opposite, causing the observed asymmetry.

Although not studied by the experiment we model, the stimuli set from Larkey and Markman (2005) can also give way to asymmetries (Hodgetts & Hahn, 2012). This study found asymmetries in response time in a subset of the manipulations of the experiment, depending on which stimulus is being compared to. Our approach is compatible with both the presence of asymmetry in these results, and with its manifestation as response time. As an intuitive example for how we explain the asymmetry, we refer again to the duplicate MOP condition (such as an AB stimulus being modified to AA) that we have discussed earlier. In this condition, AB will spread more activation to AA than AA to AB, since A's link to AA is stronger than its link to AB. This reasoning generalizes across the different conditions as well.

Our account also predicts that such asymmetries can also manifest as asymmetries in response times. Familiarity and priming together form an activation value that, according to our overarching theory of cognition, determines how easy it is to access a memory. This implies that stimuli with higher familiarity and priming will be more easily accessed in memory, allowing the model to respond more quickly to a query or judgment. Additionally, for more complicated, structural stimuli, the representations of the stimuli can be built up more quickly when there are overlapping features or objects, allowing for a faster response.

6.4. Representation and structure in similarity

One downside of the traditional prototypicality- and feature-based approaches to similarity is that, in addition to being unable to capture similarity between very simple perceptual stimuli, they also cannot easily account for the higher-order relationships, or structure, between different concepts; many argue that these relationships are critical for determining similarity (Medin et al., 1993). As we have discussed, structure mapping approaches attempt to mitigate this shortfall by explicitly considering correspondences and structure as part of their process.

Our approach, in contrast, is able to capture many of the nuances of MIPs and MOPs without any formal structure analysis. In the second and third experiments we model here, familiarity and priming capture implicit structural similarities stemming from shared components between the stimuli. This helps to unify different prototype and weighted feature-based approaches with structure-based approaches by suggesting that the constructs the approaches rely on—such as prototypicality and hierarchical feature relationships—are related.

Priming, however, is only as successful as the items' representations allow, since both rely on overlaps in item representations. While here we join many structure-mapping approaches in assuming simple and appropriate item representations, our work is also compatible with the notion that similar items being compared may be represented differently (e.g., they may be non-aligned). In cases such as this, we believe that some sort of higher level, deliberative processes are necessary to make meaningful similarity judgments (Forbus, Gentner, & Law, 1995; Gentner & Markman, 2005). While we do not model this type of alignment or analogical reasoning here, our approach fits well with these approaches and, once such reasoning has occurred, would naturally take advantage of newly created (and aligned) representations to determine similarity.

An additional, related point, is that we are of the opinion that these differing representations are likely to become more and more common as the items being compared become more and more complex. Therefore, while our account of similarity is able to capture the data shown here, at some point we believe that the stimuli similarity will become too complex for our approach to capture without the higher level, deliberative processes mentioned above. While we are not sure where the line lies between stimuli that require extra reasoning and those that do not, it is something that we look forward to exploring in the future.

6.5. Cognitive framework for similarity

While our more general conclusions about the role of familiarity, priming, and perceptual likeness hold largely independently of the details of their implementations, we also believe that our specific approach to these measures provides distinct benefits to our work. Our use of the ACT-R architecture for familiarity and priming connects our approach with a broader theory of human cognition that has used the mechanisms described here to explain a diverse array of cognitive phenomena, such as memory, diagnostic reasoning, case-based reasoning, and errors during task execution (Altmann & Trafton, 2002; Anderson, 1983; Anderson et al., 1998; Breslow, Trafton, et al., 2009;

Harrison & Trafton, 2010; Hiatt & Trafton, 2015a,b; Lenz & Burkhard, 1996; Mehlhorn, Taatgen, Lebiere, & Krems, 2011). Our work also calculates the familiarity and priming measures in a process-oriented way, in which they are learned and change over time as the model performs the experiments. When viewed in this light, our work suggests that the mechanisms that determine similarity are also used when, for example, priming a memory for retrieval, determining what step to take next in a sequential task, or determining the similarity of sets of medical symptoms to diagnose a disease.

As we have stated before, our goal is to understand the strength of the roles of familiarity, priming, and perceptual likeness in similarity. Therefore, when comparing these measures' values against empirical Likert similarity ratings, we used a theoretically light method of converting these measures' values to similarity ratings. In the cognitive framework that we use in our approach, there are deeper models of such ratings (e.g., Petrov & Anderson, 2005). However, they make auxiliary assumptions that could have clouded our understanding of our results and so we do not utilize them here.

7. Conclusions

In this article, we have presented a novel way of accounting for similarity. Our approach posits that similarity stems from three main sources—familiarity, priming, and inherent perceptual likeness. Using these three measures, our account of similarity explains ratings of both simple, color-based perceptual stimuli that display asymmetry effects, as well as more complicated perceptual stimuli with structural properties; more traditional approaches to similarity solve only one or the other, but have difficulty explaining both. Overall, our work highlights the importance of these components of similarity, both individually and together.

Acknowledgments

This work was supported by the Office of Naval Research and the Office of the Secretary of Defense. The views and conclusions contained in this document do not represent the official policies of the U.S. Navy. Our model for Experiment 1 appeared, in an abbreviated form, in Hiatt and Trafton (2013). Many thanks to Thad Polk and Levi Larkey for making their data available to us, and to Robert Goldstone for trying to. We would also like to acknowledge Wallace E. Lawson, Anthony M. Harrison, and Sangeet S. Khemlani for their helpful discussions during this work.

Note

1. In the description of the original study, the participants are said to have performed 160 trials; as this does not allow each manipulation to be used equally as the experiment is said to do, we consider this value to be, instead, 150 trials.

References

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, *26*, 39–83.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*(4), 341–380.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*, 186–197.

Böhm, U., & Mehlhorn, K. (2009). The influence of spreading activation on memory retrieval in sequential diagnostic reasoning. In *Proceedings of the 9th International Conference on Cognitive Modeling (ICCM)*. Manchester, UK.

Breslow, L. A., Ratwani, R. M., & Trafton, J. G. (2009). Cognitive models of the influence of color scale on data visualization tasks. *Human Factors*, *51*(3), 321–338.

Breslow, L. A., Trafton, J. G., & Ratwani, R. M. (2009). A perceptual process approach to selecting color scales for complex visualizations. *Journal of Experimental Psychology: Applied*, *15*(1), 25–34.

CIE (2001). *Inprovements to industrial colour-difference evaluation* (Tech. Rep. No. CIE 142-2001). Vienna: Commission Internationale de l'Eclairage.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*, 1–63.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*, 141–205.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, *59*, 47–59.

Gentner, D., & Markman, A. B. (2005). Defining structural similarity. *Cognitive Science*, *6*, 1–20.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, *97*, 3–18.

Goldstone, R. L. (1994a). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178–200.

Goldstone, R. L. (1994b). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*(3–28), 10–11.

Hahn, U., Close, J., & Graf, M. (2009). Transformation direction influences shape similarity judgments. *Psychological Science*, *20*, 447–454.

Harrison, A. M. (2014). Listening to the architecture: ACT-R, associative learning, and the fan-effect [abstract]. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX.

Harrison, A. M., & Trafton, J. G. (2010). Cognition for action: An architectural account for "grounded interaction." In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Hiatt, L. M., & Trafton, J. G. (2013). The role of familiarity, priming and perception in similarity judgments. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 573–578). Austin, TX: Cognitive Science Society.

Hiatt, L. M., & Trafton, J. G. (2015a). An activation-based model of routine sequence errors. In *Proceedings of the International Conference on Cognitive Modeling* (pp. 244–249).

Hiatt, L. M., & Trafton, J. G. (2015b). A computational model of mind wandering. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 914–919). Austin, TX: Cognitive Science Society.

Hodgetts, C. J., & Hahn, U. (2012). Similarity-based asymmetries in perceptual matching. *Acta Psychologica*, *139*(2), 291–299.

Hodgetts, C. J., Hahn, U., & Chater, N. (2009). Transformation and alignment in similarity. *Cognition*, *113* (1), 62–79.

Kozima, H., & Furugori, T. (1993). Similarity between words computed by spreading activation on an English dictionary. In *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics* (pp. 232–239). Stroudsburg, PA: Association for Computational Linguistics.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, *85*(5), 445–463.

Larkey, L. B., & Love, B. C. (2003). CAB: Connectionist analogy builder. *Cognitive Science*, *27*, 781–794.

Larkey, L. B., & Markman, A. B. (2005). Processes of similarity judgment. *Cognitive Science*, *29*, 1061–1076.

Lawson, W., Hiatt, L., & Trafton, J. G. (2014). Leveraging cognitive context for object recognition. In *Proceedings of the Vision Meets Cognition Workshop (Held in Conjunction with CVPR)* (pp. 381–386). Columbus, OH: IEEE.

Lenz, M., & Burkhard, H.-D. (1996). Case retrieval nets: Basic ideas and extensions. In *Ki-96: Advances in artificial intelligence* (pp. 227–239).

Levin, D. T., & Beale, J. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception and Psychophysics*, *23*, 1153–1169.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*(2), 254–278.

Medin, D. L., Goldstone, R. L., & Markman, A. B. (1995). Comparison and choice: Relations between similarity processes and decision processes. *Psychonomic Bulletin and Review*, *2*, 1–19.

Mehlhorn, K., Taatgen, N. A., Lebiere, C., & Krems, J. F. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *37*, 1391–1411.

Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity, and classification. *Cognitive Psychology*, *23*, 94–140.

Nosofsky, R. M. (1992). Exemplar-based approach to relating categorization, idenfitication, and recognition. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 363–393). Hillsdale, NJ: Lawrence Erlbaum.

Novick, L. R. (1990). Representational transfer in problem solving. *Psychological Science*, *2*(1), 128–132.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by stimulating the brain*. Cambridge, MA: MIT Press.

O'Reilly, R., Wyatte, D., Herd, S., Mingus, B., & Jilk, D. (2013). Recurrent processing during object recognition. *Frontiers in Psychology*, *4*(124).

Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review*, *112*, 383–416.

Plot digitizer (2015). Available at http://plotdigitizer.sourceforge.net. Accessed December 22, 2015.

Polk, T. A., Behensky, C., Gonzalez, R., & Smith, E. E. (2002). Rating the similarity of simple perceptual stimuli: Asymmetries induced by manipulating exposure frequency. *Cognition*, *82*(3), B75–B88.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*, 1043–1056.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, *7*, 532–547.

Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of hick's law. *Cognitive Psychology*, *62*(3), 193–222.

Sjoberg, L. (1972). A cognitive theory of similarity. *Goteborg Psychological Reports*, *2*(10).

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. *Advances in Psychology—Percepts, Concepts and Categories: The Representation and Processing of Information*, *93*, 233–272.

Trafton, J. G., Hiatt, L. M., Harrison, A. M., Tamborello II, F. P., Khemlani, S. S., & Schultz, A. C. (2013). ACT-R/E: An embodied cognitive architecture for human-robot interaction. *Journal of Human-Robot Interaction*, *2*(1), 30–55.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Ulhaque, E., & Bahn, K. D. (1992). A spreading activation model of consumers' asymmetric similarity judgment. *Advances in Consumer Research*, *19*, 782–786.

Vinokurov, Y., Lebiere, C., Herd, S., & O'Reilly, R. (2011). A metacognitive classifier using a hybrid ACT-R/Leabra architecture. In *Lifelong learning: Papers from the 2011 AAAI workshop* (pp. 50–55). AAAI Press.