Creating Something Different: Similarity and Contrast in Concept Generation.

Nolan Conaway[1], Kenneth J. Kurtz[2], & Joseph L. Austerweil[1]

[1]University of Wisconsin-Madison, Department of Psychology, Madison, WI, USA
[2]Department of Psychology, Binghamton University, Binghamton, NY, USA

1 Abstract

JLA: Joe! Did you know i made a command for you to enter in comments like this??

`\jlanote{Some note}`

Keywords: categorization, concepts, creativity, generation

2 Introduction

The creation of novel concepts and ideas is a highly intriguing – yet infrequently studied – topic of research in human cognition. The creative use of conceptual knowledge is, for example, a core element of scientific investigation, wherein the generation of new ideas is critical to designing experiments and explaining observations. However, due to its complexity, little is presently known about the cognitive processes underlying the generation of new concepts.

Much of what we know about concept generation comes from the foundational literature on creative cognition. In a classic series of reports, Ward & colleagues (Marsh, Ward, & Landau, 1999; Smith, Ward, & Schumacher, 1993; Ward, 1994, 1995; Ward, Patterson, Sifonis, Dodds, & Saunders, 2002) established that category generation is highly constrained by prior knowledge: generated categories tend to consist of features observed in known categories, and they tend to exhibit the distributional properties as found in known categories. In a classic study, Ward (1994) asked participants to generate new species of alien animals by drawing and describing members of the species. People tended to generate species with the same features as on Earth (e.g., eyes, legs, wings), and possessing the same feature correlations as on Earth (e.g., feathers co-occur with wings). Likewise, aliens drawn from the same species tended to share more features with one another compared to members of opposite species.

Much of the work from this area (e.g., Marsh et al., 1999; Smith et al., 1993) focuses on how sample cues (such as an example of a species generated by other participants) can drastically diminish creativity. However, the broader set of observations made by Ward & colleagues provide a great deal of insight into the nature of creative generation. They indicate that people rely strongly on prior knowledge in the creative process, and people generate concepts in accord with what they already know.

Theoretical accounts of these effects were primarily grounded within the categorization literature. For example, the predominant "Path of Least Resistance"

account (see Ward, 1994, 1995; Ward et al., 2002) proposes that, when generating a new species of animal, people retrieve from memory a known subcategory of animals (e.g., *bird*, *dog*, *horse*), and simply change some of the features to make something new. People were thought to change only features that are not characteristic of the retrieved category (e.g., if *bird* was retrieved, the presence of *wings* would not change, but *color* might). This theory incorporates elements of the highly influential basic-level categories framework (Rosch, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), as well as the exemplar view (Medin & Schaffer, 1978; Nosofsky, 1984, 1986). Problematically, however, the experimental paradigms employed in these studies were relatively uncontrolled, precluding the development of formal approaches. Specifically, participants in these studies are typically asked to generate objects that they have a rich degree of prior knowledge about (e.g., plants, animals, tools, toys), and so it is difficult to apply formal models to the data.

Jern and Kemp (2013) recently showed that creative generation could be studied in a more controlled manner through the well-developed methods of an artificial categorization paradigm (see Kurtz, 2015). In their experiments 3 and 4, participants were exposed to members of experimenter-defined categories of "crystals" varying in size, hue, and saturation. Following a training phase during which the experimenter-defined categories were learned, participants were asked to generate novel categories of crystals. In a finding mirroring that of the Ward (1994) studies, Jern and Kemp found that participants generated categories with the same distributional properties as the experimenter-defined categories: for example, after training on categories with a positive correlation between the size and saturation features (larger sized crystals were more saturated), participants generated novel categories with the same positive correlation. This finding is notable, as it demonstrates that category generation can be studied in a well-known and highly controlled experimental paradigm.

The authors evaluated the predictions of several formal models on their data. Most notably, they showed that a Bayesian hierarchical sampling model provided the strongest

account. Their model represents categories as multivariate normal distributions within the space, and from exposure to members of observed categories the model induces a domain-wide representation of the distributional commonalities among the observed categories. When generating a new category, the model samples a new distribution from the domain, and thus the distributional properties of observed categories will be preserved within the generates category.

NBC: Maybe we need a more technical view of the model here?

Jern and Kemp (2013) additionally tested a "copy-and-tweak" model that broadly resembles the earlier "Path of Least Resistance" account: the core proposal is that participants generate new items by copying stored examples from memory and tweaking them to generate something new. The copy-and-tweak model differs from the path of least resistance account in that it notably omits the hierarchical organization of categories, as well as selectivity in which features are changed. Instead, their copy-and-tweak model corresponds to a direct exemplar-similarity approach (e.g., Nosofsky, 1984): participants are thought to generate new items according to their similarity to known members of the target category. This model provided a poor account of the observed data: indeed, the experiments devised by Jern and Kemp were specifically designed to challenge this model. However, it's application is notable as a first step toward explaining category generation using well-known formal approaches from the categorization literature.

2.1   Something Different: A Role For Contrast

It is worth noting that the literature reviewed above provides a somewhat limited view of category generation: the main concern of the published experiments has been on the distributional correspondences between learned and generated categories, and as a result most of the modeling efforts have been towards explaining those effects. In this paper, we seek to identify other important constraints on the processes involved in creative generation. Specifically, we focus on the constraints imposed by category contrast. That is,

in order to generate a novel concept, individuals must produce something that is in some capacity *different* from what they already know. Thus, in a trivial sense, contrast can be viewed as a fundamental constraint on creative generation: new concepts must firstly be different from existing ones.

Although it is evident that people are *capable* of creating new concepts and categories, it is not altogether clear how new concepts are systematically made different from what is already known. The hierarchical sampling model developed by Jern and Kemp (2013), for example, does not provide any strong claims about how generated categories should contrast with learned ones: it is simply assumed that categories are generated in a random location in the domain.

The Copy+Tweak account evaluated by Ward & colleagues (Ward, 1995; Ward et al., 2002) makes a slightly stronger claim: new categories should be .....

3   The PACKER Model

The PACKER model is an extension of the influential Generalized Context Model of category learning (GCM; Nosofsky, 1984). It assumes that each category is encoded by a set of exemplars within a k-dimensional psychological space, and that generation is constrained by both similarity to members of the target category (the category in which a stimulus is being generated) as well as similarity to members of other categories.

As in the GCM, the similarity between two examples, $s\left(x_i, x_j\right)$, is an inverse-exponential function of distance:

$$s\left(x_i, x_j\right) = \exp\left\{-c \sum_k |x_{ik} - x_{jk}| w_k\right\} \tag{1}$$

where w_k is the attention weighting of dimension k ($w_k \geq 0$ and $\sum_k w_k = 1$), accounting for the relative importance of each dimension in similarity calculations, and c ($c > 0$) is a specificity parameter controlling the spread of exemplar generalization. For simplicity, our simulations will use uniform attention weights, except in our discussion of individual differences.

To generate a new example, the model considers both the similarity to examples from contrast categories as well as the similarity to examples (if any exist) in the target category. The aggregated similarity a between generation candidate y and stored exemplars x is given by:

$$a(y, x) = \sum_j f(x_j) s(y, x_j) \tag{2}$$

where $f(x_j)$ is a function specifying the extent to the exemplar contributes to generation. PACKER sets $f(x_j)$ depending on exemplar x_j's category membership: $f(x_j) = \phi$ if x_j is a member of a contrast category, and $f(x_j) = \gamma$ if x_j is a member of the target category. ϕ and γ are free parameters ($-\infty \leq \phi, \gamma \leq \infty$) controlling the contribution of contrast- and target-category similarity, respectively. Larger absolute values result in greater

consideration of those exemplars, with values of 0 eliminating their effect. A negative value for $f(x_j)$ produces a 'repelling' effect (exemplars are less likely to be generated nearby x_j). Conversely, a positive value for $f(x_j)$ produces an 'attracting' effect (exemplars are more likely to be generated nearby x_j).

PACKER's core proposal is that new categories should be different from existing categories, and same-category exemplars should be similar to one another. This is realized when $\phi \leq 0$, and $\gamma \geq 0$. Negative ϕ values encourage y to be distant from the contrast category (as similarity to contrast category exemplars are subtracted during aggregation). Positive γ values encourage y to be close to other exemplars of the target category. When $|\phi| = \gamma$, the repulsion effect from contrast categories is equal to the attraction effect to the target category.

The probability that a given candidate y will be generated is evaluated using an Exponentiated Luce (1977) choice rule. Candidates with greater values of a are more likely to be generated than candidates with smaller values:

$$p(y) = \frac{\exp\{\theta \cdot a(y, x)\}}{\sum_i \exp\{\theta \cdot a(y_i, x)\}} \tag{3}$$

where θ ($\theta \geq 0$) controls response determinism.

4   Snippets

NBC: Just some random snippets that i don't want to delete yet...

How do people generate concepts such that the product is (A) different from what is already known, and (B) internally coherent?

We also introduce a novel exemplar-based model of category generation, PACKER (*Producing Alike and Contrasting Knowledge using Exemplar Representations*), which explains much of what we observe in the creative use of conceptual knowledge. In the sections below, we introduce the PACKER model and its core theoretical claims.

One of the core principles of PACKER (as well as its name) are inspired by earlier work in category learning (see Hidaka & Smith, 2011).
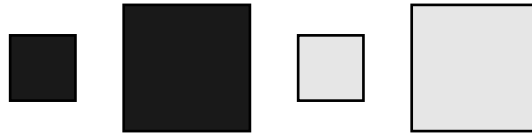
Figure 1: Sample stimuli used in Experiments 1 and 2.

5   Experiment 1

5.1   Participants & Materials

183 participants were recruited from Amazon Mechanical Turk. Participants were randomly assigned to one condition: 64 participants were assigned to the Cluster condition, 61 were assigned to the Row condition, and 58 were assigned to the XOR condition. Stimuli were squares varying in color (RGB 25–230) and side length (3.0–5.8cm). These stimuli are slight variants of those used by Conaway and Kurtz (2016), see Figure 1 for samples. The assignment of perceptual features (color, size) to axes of the domain space (x, y) was counterbalanced across participants.

5.2   Procedure

Participants began the experiment with a short training phase (3 blocks of 4 trials), where they observed exemplars belonging to the 'Alpha' category. Participants were instructed to learn as much as they can about the Alpha category, and that they would answer a series of test questions afterwards. On each trial, a single Alpha category exemplar was presented, and participants were given as much time as they desired before moving on. Each block consisted of a single presentation of each of the members of the Alpha category, in a random order. Participants were shown the range of possible colors and sizes prior to training.

Following the training phase, participants were asked to generate four examples belonging to another category called 'Beta'. As in Jern and Kemp (2013), generation was

completed using a sliding-scale interface. Two scales controlled the features (color, size) of the generated example. An on-screen preview of the example updated whenever one of the features was changed. Participants could generate any example along an evenly-spaced 9x9 grid, except for any previously generated Beta exemplars. Neither the members of the Alpha category nor the previously generated Beta examples were visible during generation. Prior to beginning the generation phase, participants read the following instructions:

> As it turns out, there is another category of geometric figures called "Beta". Instead of showing you examples of the Beta category, we would like to know what you think is likely to be in the Beta category.

> You will now be given the chance to create examples of any size or color in order to show what you expect about the Beta category. You will be asked to produce 4 Beta examples - they can be quite similar or quite different to each other, depending on what you think makes the most sense for the category.

> Each example needs to be unique, but the computer will let you know if you accidentally create a repeat.

Following generation, participants completed a generalization phase wherein they classified novel examples into the Alpha and Beta categories without feedback. On each trial, a single example was presented, and participants were asked to classify it by clicking buttons labeled "Alpha" or "Beta". Participants classified a total of 81 items sampled along a 9x9 grid, including the members of the Alpha and Beta categories (randomly intermixed). These data were, however, collected to address a separate set of questions, and we do not discuss them in this paper.

6 Experiment 2

6.1 Participants & Materials

122 participants were recruited from Amazon Mechanical Turk. 61 participants were randomly assigned to the Middle and Bottom conditions each. The stimuli were exactly as in Experiment 1. Again, the assignment of perceptual features (color, size) to axes of the domain space (x, y) was counterbalanced across participants.

6.2 Procedure

The procedure was exactly as in Experiment 1: participants first completed a short training phase, followed by the generation phase, followed by the generalization phase (data from this phase is not discussed in this report).

7 Acknowledgments

References

Conaway, N., & Kurtz, K. J. (2016). Similar to the category, but not the exemplars: A study of generalization. *Psychonomic Bulletin & Review*, 1–12. doi: 10.3758/s13423-016-1208-1

Hidaka, S., & Smith, L. B. (2011). Packing: a geometric analysis of feature selection and category formation. *Cognitive Systems Research*, *12*(1), 1–18.

Jern, A., & Kemp, C. (2013). A probabilistic account of exemplar and category generation. *Cognitive Psychology*, *66*(1), 85–125.

Kurtz, K. J. (2015). Human category learning: Toward a broader explanatory account. *Psychology of Learning and Motivation*, *63*, 77–114.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of mathematical psychology*, *15*(3), 215–233.

Marsh, R. L., Ward, T. B., & Landau, J. D. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory & Cognition*, *27*(1), 94–105.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*(1), 104.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, *21*(6), 837–845.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar

generation. *Cognitive Psychology*, *27*(1), 1–40.

Ward, T. B. (1995). What's old about new ideas. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 157–178). Cambridge, MA: MIT Press.

Ward, T. B., Patterson, M. J., Sifonis, C. M., Dodds, R. A., & Saunders, K. N. (2002). The role of graded category structure in imaginative thought. *Memory & Cognition*, *30*(2), 199–216.