The Spacing Effect in Inductive Learning

Avril Kenney

9.61

May 11, 2009

Abstract

Learning general principles from specific examples is a key part of acquiring knowledge about the world. Previous studies of such inductive learning have focused primarily on the generalization process and not the context in which examples are actually encountered. It is well established that for memory of individual items, spacing instances apart (spaced presentation) is more beneficial than grouping them together (massed presentation). The experiments described here show that this effect is also present in learning general concepts. In Experiment 1, participants learned the styles of different artists by viewing example paintings, with either spaced or massed presentation. Accuracy in identifying the artists of new paintings was significantly higher for spaced artists than massed. Furthermore, there was no difference in the effect of spacing for new paintings compared with paintings seen during the training phase. In Experiment 2, participants learned names for different categories of abstract shapes that differed in their level of discriminability from one another. The spacing effect was present to the same extent for high-discriminable and low-discriminable categories.

Introduction

Most of the knowledge that people acquire comes not through explicit instruction but through generalizing from observed examples to reasonable models of the world. This type of learning, called induction, is present in every aspect of life. It is used to acquire the structure and vocabulary of language, to determine what relationships hold between classes of objects, and to build representations of causality. People are remarkably adept at learning concepts from even a single instance (Tenenbaum, 1999), and are able to base their generalizations on different information depending on what is relevant in the context (Heit & Rubinstein, 1994).

Previous studies on induction have focused on the generalizations themselves: given a set of examples, how do people categorize them. In everyday learning situations, however, it is likely that instances of a concept will be encountered at different points in time rather than all together. To learn what a concept consists of, people must not only determine the similarities between examples but also remember relevant properties across time.

In the domain of memory, numerous studies have found that concepts are remembered better when repeated presentations are separated from each other in time (spaced) than when they are grouped together (massed) (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006, for one review). This phenomenon is known as the spacing effect, and there is evidence for its influence in memory for things as varied as lists of words (e.g., Glenberg, 1979), mathematical rules (e.g., Gay, 1973), and motor skills (e.g., Shea, Lai, Black, & Park, 2000).

There is no complete explanation for how the spacing effect occurs, but several theories exist. Encoding variability theories claim that when presentations are further apart, the contexts in which each item occurs are more varied, and therefore more associations are formed, making it more likely that some cue will enable recall of the items (Glenberg, 1979). Deficient processing theories claim that after the initial presentation of an item, people are primed to process it or other things similar to it, so immediate presentations of the same stimulus are processed less thoroughly and therefore not encoded into memory as well (Cuddy & Jacoby, 1982).

For inductive learning, it seems that spacing could be either beneficial or detrimental. On the one hand, spacing helps memory of individual items, and learning general concepts is similar except that the items are not exactly alike. On the other hand, to generalize from examples one has to compare them with one another to find similarities, and intuitively this should be easier when they are all presented close together. Previous work on how spacing affects memory for anything other than the exact items presented is somewhat sparse and provides no clear conclusion, but suggests that in some situations spacing does not facilitate learning.

In two experiments by Dellarosa and Bourne (1985), participants had to learn the gist of a sentence by reading or hearing it multiple times. When every presentation of a given sentence was the same, spacing helped memory more than did massing. However, when the presentations of a given sentence were paraphrases of each other rather than verbatim repetitions, the spacing effect was significantly reduced, and when the repetitions were spoken by a different voice, the spacing effect disappeared completely. The proposed explanation is that even in massed trials, the difference in surface form is enough to cause complete processing of the sentence in each presentation. If this is the case, then spacing would not be expected to help in inductive concept learning, because the examples presented are all different from one another.

Another study demonstrated that inductive learning of rule-based categories is actually hindered by spacing (Kurtz & Hovland, 1956). Participants viewed images consisting of shapes that varied on several dimensions and had to learn from these examples which combinations of features corresponded to which of four different nonsense names. The properties were learned better in the massed condition, where all instances of any one name were presented in sequence, than in the spaced condition, where two instances of any name were always separated by at least one instance of a different one. This suggests that at least in some cases, it is easier for people to generalize when multiple examples of a given concept are presented together.

However, a recent set of experiments by Kornell and Bjork (2008) found strong evidence that spaced presentation can in fact help inductive learning. Participants were instructed to learn the styles of several different artists by viewing example paintings, with each artist's work presented either massed or spaced, and then were shown new paintings and had to identify which of the artists had painted each one. People correctly identified paintings much better for artists who had been presented spaced than those who had been presented massed, demonstrating that the spacing effect can be present even when the learning instances are not identical.

One possible explanation for the discrepancy between these findings is a difference in discriminability of the concepts being learned: it could be that the styles of artists used by Kornell and Bjork (2008) are difficult to distinguish whereas the shape arrangements from Kurtz and Hovland (1956) are relatively easy to tell apart from one another. It has previously been found that similarity of items can interact with order of presentation, for example when learning the names of multiple abstract figures (Gagné, 1950). If the figures that looked very similar to one another were grouped together, people tended to make more overgeneralization errors initially, but they ultimately learned the names faster than from other presentation orders.

A similar result shows that the type of material intervening between presentations of an item can affect the strength of the spacing effect (Cuddy & Jacoby, 1982). In this experiment participants were given two presentations of word pairs, separated by zero, four, or eight other items, and then were tested on recall of the second word from each pair, cued by the first word. Recall overall was higher when the items intervening between the two presentations were other word pairs than when they were arithmetic problems. This effect of intervening item type was greater for the shorter spacing interval, suggesting that highly similar intervening material causes faster forgetting of previous presentations and therefore more complete processing of subsequent ones.

The present study aimed to test the effect of spaced versus massed presentation on learning concepts from examples. Experiment 1a replicated Kornell and Bjork (2008), and Experiment 1b extended this result to account for direct effects on memory as opposed to induction. Experiment 2 compared the two presentation styles for a group of some high-discriminable and some low-discriminable concepts.

Experiment 1a

This experiment tested the spacing effect on a type of concept that could only be learned from examples – the painting styles of different artists. Participants were shown paintings by twelve artists, half massed and half spaced, and then were tested on their ability to identify new paintings by the same artists.

Methods

Participants

Participants were 32 members of the MIT community (23 male, 9 female), primarily undergraduates (median age approximately 21). Each was given two candy bars for participating. All participants in this experiment (and the others reported in this paper) signed a consent form approved by the MIT Committee on the Use of Humans as Experimental Subjects.

Materials

Stimuli were JPEG images of paintings by twelve different artists: Bruno Pessani, Ciprian Stratulat, Georges Braque, Georges Seurat, George Wexler, Henri-Edmond Cross, Judy Hawkins, Marilyn Mylrea, Philip Juras, Ron Schlorff, Ryan Lewis, and Yie Mei (identical to the stimuli used by Kornell & Bjork, 2008). The content of the paintings was similar across artists, consisting of natural scenes such as trees, houses, and boats. The experiment was conducted on laptop computers, using Psychtoolbox for MATLAB. Each image was resized such that its diagonal was half as long as that of the computer screen, since not all paintings had the same dimensions.

Design

Conditions (massed and spaced) were varied within subjects. For each participant, the paintings of half the artists were presented massed, and the paintings of the remaining artists were presented spaced. Assignment of artists to conditions was not random but was counterbalanced across subjects, with artists divided into two groups; 17 subjects viewed Group 1 massed and 15 viewed Group 2 massed. The training segment was divided in to six blocks each of massed (M) and spaced (S), in the order MSSMMSSMMSSM for all subjects. A massed block consisted of six paintings by the same massed artist, while a spaced block consisted of one painting by each of the six spaced artists in a random order. The test segment was divided into four blocks, each consisting of one painting by each of the twelve artists. The order of paintings and artists was otherwise completely random in both the training and the test phases.

Figure 1: Examples of three paintings by the same artist and three by different artists.

Procedure

Participants were told that they would be viewing paintings by different artists and should try to learn each artist's style so that they would be able to identify new paintings by the same artists. They were then seated in front of the computer screen at whatever distance they chose. Each painting in the training segment was presented for 3 seconds, with the image centered in the middle

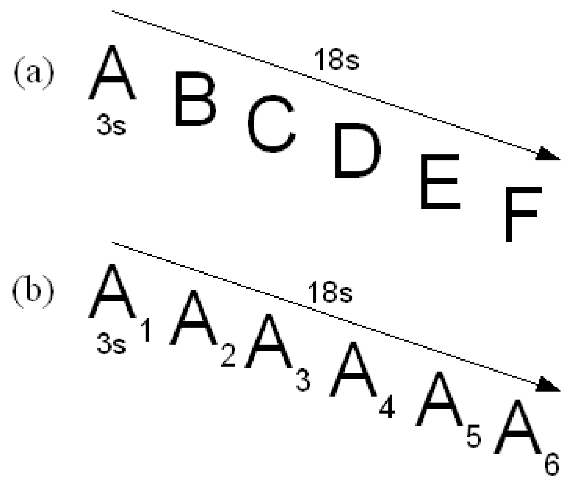

Figure 2: Design of blocks in the training phase for Experiment 1a. Each spaced block (a) was six paintings by different artists, and each massed block (b) was six paintings by the same artist.

of the screen and the artist's full name displayed above it at a fixed distance from the top of the screen.

After the 72 training paintings had been shown, a screen appeared instructing participants to count backward by 3's from 547 out loud for 15 seconds (this distractor task was taken directly from Kornell & Bjork, 2008), after which the test segment began. Each painting appeared in the middle of the screen, with 13 buttons below it: 12 buttons of equal size, one for each artist, were displayed in two rows of six, and to the right was a smaller button labeled "Don't Know." The order of artists' names on the buttons was constant across trials and participants, alphabetical by first name. After the subject clicked a button, the screen was cleared and a feedback message was shown for 1 second, saying "Correct!" if the subject had chosen correctly, and otherwise saying "The correct answer was:" followed by the correct artist's name.

Results

Performance was significantly better on artists whose paintings were presented spaced than those whose paintings were massed. Of the 32 subjects, 25 correctly identified more paintings by spaced artists, 6 correctly identified more paintings by massed artists, and 1 performed equally well on both (sign test, $p < .001$). The effect of condition (massed or spaced) within subjects was highly significant (paired-samples t-test, $t(31) = 4.81, p < .0001$); individuals' performance is plotted in Figure 3. A 2×4 ANOVA, with condition and block as within-subjects factors, shows a significant main effect of block ($F(1,31) = 8.75, p < .01$), indicating that performance improved over the course of the test phase (averages across subjects are plotted in Figure 4). A marginally significant interaction between condition and block ($F(1,31) = 3.49, p = .07$) suggests that the effect of block can be explained by improvement in the massed condition alone.

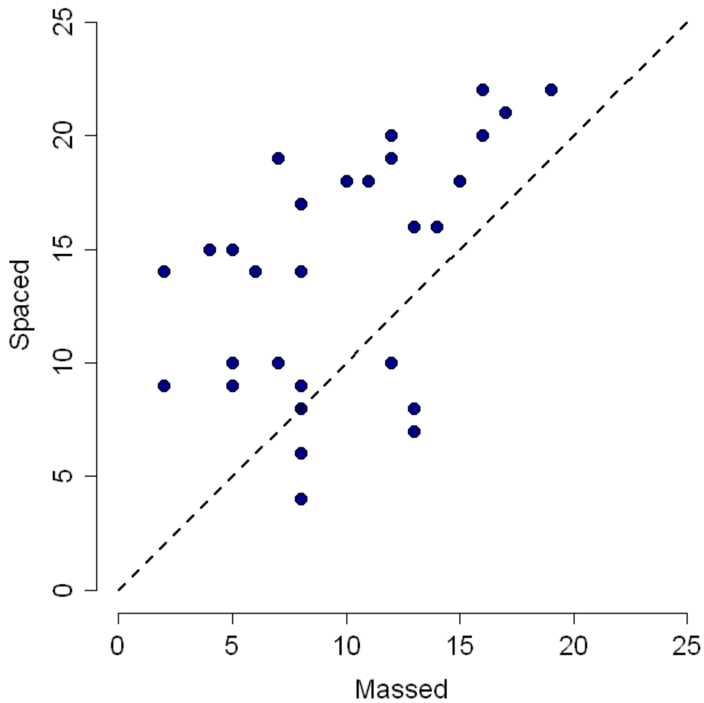

Figure 3: Performance of individual subjects. Each point corresponds to one subject, with the number of correct responses for spaced artists plotted against the number of correct responses for massed artists (both out of a possible 24). The dotted line represents equal performance on the two conditions.

Eight of the participants indicated that they had been familiar with at least one of the artists prior to the experiment. A 2×2 ANOVA, with condition (massed or spaced) and familiarity (familiar or not familiar) as within-subjects factors, shows a significant main effect of familiarity ($F(1,29) = 5.36, p < .03$): subjects performed better on artists with whom they were familiar. However, there was no interaction between condition and familiarity ($F(1,29) = 2.66, p = .11$).

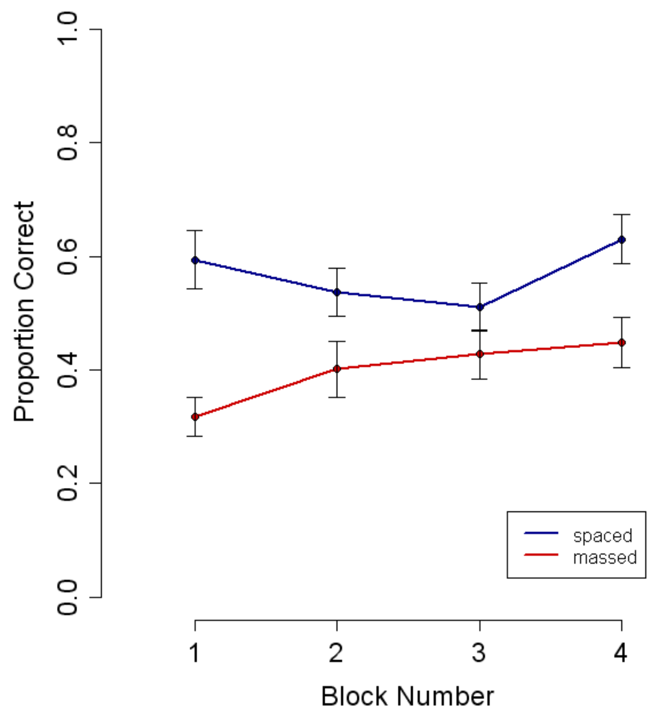

Figure 4: Average performance across the four test blocks. Error bars represent standard errors.

Experiment 1b

To test whether spacing in Experiment 1a had been specifically helping induction, as opposed to only indirectly improving performance on the inductive task by helping memory, a second version was conducted in which subjects were tested on both new and previously-seen paintings.

Methods

Participants

Subjects were 28 adults (15 male, 13 female), median age 24. None of them had participated in Experiment 1a. They participated as volunteers or received candy bars, money, or extra credit in an introductory psychology course.

Materials

Stimuli were identical to those used in Experiment 1a.

Design

The design was the same as in Experiment 1a, except that in the test phase, two of the paintings presented for each artist were paintings that had been shown during the training phase, and the other two were new. The order of new paintings relative to learned ones was random, as was which paintings were selected for each artist.

Procedure

The procedure was the same as Experiment 1a, except that during the test phase, subjects did not receive feedback as to whether their responses were correct, in order to prevent additional learning. (Before running Experiment 1b, an additional four subjects were tested using the design in Experiment 1a but without feedback. They did not show any effect of test block on performance, and their results were otherwise qualitatively the same as those in Experiment 1a.)

Results

A 2×2 ANOVA, with condition (massed or spaced) and painting type (learned or new) as within-subjects factors, showed a significant main effect of condition: subjects correctly identified more paintings by spaced artists than massed artists ($F(1,27) = 7.44, p < .05$). However, there was no interaction between painting type and condition ($F(1,27) = 1.77, p = .19$), indicating that the spacing effect was present to the same extent regardless of whether the exact stimulus had been seen before. There was also a main effect of painting type ($F(1,27) = 5.35, p < .05$): subjects were more likely to correctly identify paintings that they had already seen than other paintings with similar styles.[1] Means are shown in Figure 5.

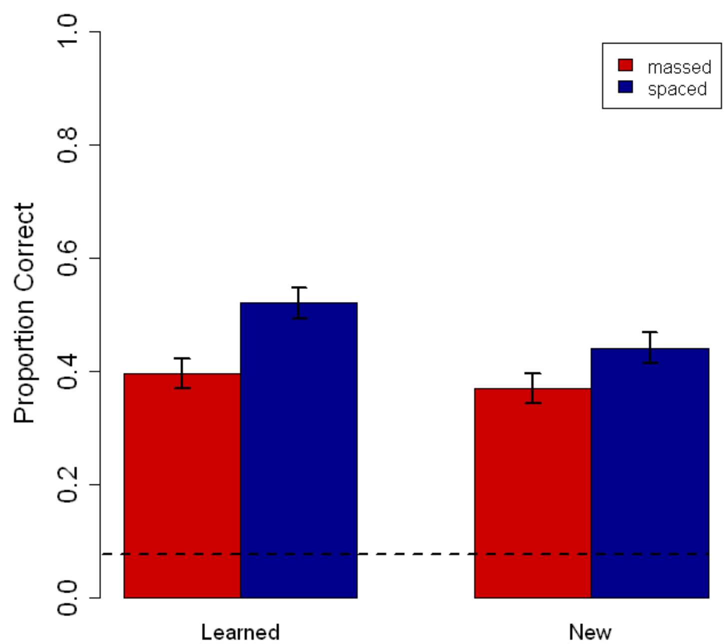

Figure 5: Average performance in the massed and spaced conditions for learned paintings (which had been seen during the training phase) and new paintings (which had not been seen before). Error bars represent standard errors, and the dotted line represents chance performance.

Experiment 2

The purpose of this experiment was to test the idea proposed by Kurtz and Hovland (1956) that spaced presentation is more beneficial to learning when the concepts being interleaved are less discriminable from one another.

[1] Interestingly, post-hoc paired-samples t-tests show that this advantage for learned paintings over new paintings is present when considering spaced artists alone ($t(27) = 2.39, p < .05$) but not when considering massed artists alone ($t(27) = 0.98, p = .33$).

Methods

Participants

Participants were 33 adults (15 male, 18 female), median age 26, recruited through the MIT Brain and Cognitive Sciences subjects website. None had participated in Experiments 1a or 1b. Each received either money or extra credit in an introductory psychology course for participating.

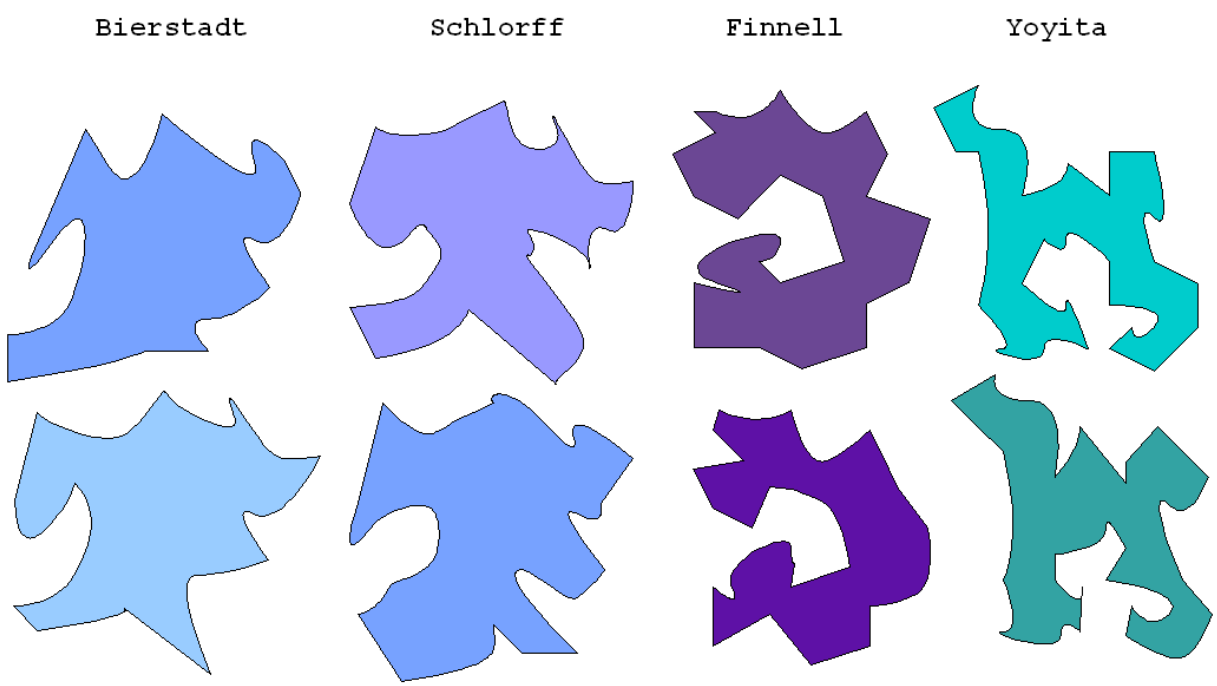

Figure 6: Examples of two low-discriminable ("Bierstadt" and "Schlorff") and two high-discriminable ("Finnell" and "Yoyita") concepts from one of the two sets used in Experiment 2.

Materials

The stimuli were 96 pictures of abstract shapes, drawn using OpenOffice Impress. The pictures were divided into two sets, each of which consisted of four high-discriminable and four low-discriminable concepts (see Figure 6 for examples), with six instances of each. For each concept, all six instances had very similar colors and shapes. Within each set, the four low-discriminable concepts were similarly shaped, and there was some overlap of colors across concepts, whereas the high-discriminable concepts were uniquely shaped, and their colors were somewhat similar to those of the other concepts but with no overlap. Each concept was given the last name of an artist (either from Experiment 1 or similar) as its name; assignment of names to concepts was fixed. Names were chosen to be fairly different from one another, and within each set no two concepts' names had the same first or last letter.

Design

The experiment was divided into two halves: in each half, the participant learned one set of concepts, with either massed or spaced presentation, and then was tested on those concepts. For the training phase in the massed half, four instances of each of the eight concepts were presented in succession, and in the spaced half, training was divided into blocks, with each of four blocks containing one instance of each of the eight concepts. Both halves used the same format for the test phase as in Experiment 1: there were four blocks, each containing one instance of each concept. Two of the test-phase instances of each concept had been seen in the training phase, and two had not. Order of all concepts and instances was randomized. The order of the two halves was counterbalanced across subjects, and pairing of concept sets with conditions was random.

Procedure

Participants were seated at the computer, and the experimenter provided a verbal description of the experiment along with instructions shown on the screen. It was explained that they would be shown multiple examples of several different abstract shapes, along with the associated names, and that the examples for each name would not be exactly the same, so they should try to learn the general concept for each one. They were shown a few examples of simpler name-shape pairings than used in the experiment, as well as a corresponding example of clicking the appropriate name button.

As in Experiment 1, in the training phases each image was displayed in the center of the screen for 3 seconds with its name shown above it, and in the test phases there were buttons below the images with the names of the eight concepts. There was no “Don’t Know” button. No feedback was provided as to the correctness of the answer chosen. Between training and test phases, and between the first and second halves of the experiment, screens informed participants of what they had just finished and what they would do next.

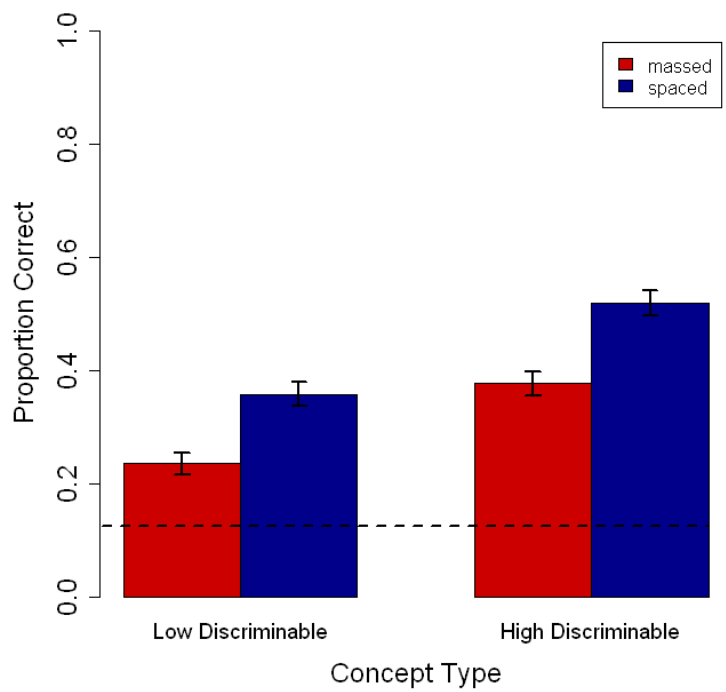

Figure 7: Average performance in the massed and spaced conditions for low-discriminable and high-discriminable concepts. Error bars represent standard errors, and the dotted line represents chance performance.

Results

A $2 \times 2 \times 2$ ANOVA (condition (massed or spaced) by image type (learned or new) by discriminability (high or low)) shows significant main effects of both condition ($F(1,32) = 24.365, p < .001$) and discriminability ($F(1,32) = 28.657, p < .001$): participants performed better on concepts that were presented spaced and concepts that were highly discriminable. Critically, there was no interaction between condition and discriminability ($F(1,32) = 0.08, p = .78$). Means are shown in Figure 7. As in Experiment 1b, there was no interaction between image type and condition ($F(1,32) = 2.55, p = .12$). In this experiment, though, there was also no main effect of image type ($F(1,32) = 2.84, p = .10$).

Discussion

Experiment 1a replicated the findings of Kornell and Bjork (2008), demonstrating that concepts can be learned better from spaced presentation than from massed even when each presentation of a concept is a different instance. Experiment 1b showed that spaced presentation helps in identifying new instances of a concept as much as in identifying previously-seen instances. Experiment 2 found no difference in the spacing effect for learning low-discriminable and high-discriminable concepts, suggesting that interleaving contrastive negative examples is not a primary cause of how spacing helps inductive learning. These results are not completely consistent with either of the two presented accounts for the spacing effect.

Deficient processing theories would predict little or no benefit from spacing, because each presentation of a concept is different from the previous ones even when they are massed. They would additionally predict a more prominent spacing effect for high-discriminable than low-discriminable items, because low-discriminable items differ less from the surrounding ones and so spaced presentation should be relatively similar to massed. Yet all of the experiments found a strong effect of spacing, and there was no difference based on discriminability. One complicating factor, though, is that these designs involved learning paired associates: in addition to learning the style of each artist or the gist of each abstract shape, people also had to learn which name it went with. Since the names were identical for every example of a given artist or concept, it could still be the case that processing of the *names* was deficient in the massed condition.

Encoding variability theories would predict more accurate identification of learned images than new ones, because the recall cues are exactly the same as what was originally learned. In Experiment 1b this prediction held, but in Experiment 2 there was no difference in performance on the two types of images. This difference could be because paintings often have distinct, identifiable features that might be memorable, whereas the abstract shapes from Experiment 2 had no such obvious features. It could also be related to the fact that the abstract shapes of each type were more similar to one another than the paintings of each artist, so perhaps the details differentiating instances of the shapes simply could not be remembered.

These issues bring up the important consideration of what role memory plays in induction. There are two opposing ways in which memory is required for an inductive task such as the ones used in this paper. First, participants must remember the learned concepts in order to give correct responses during the test phase, so in this sense the spacing effect should apply just as it does when all presented instances are identical, or perhaps with a smaller magnitude according to deficient processing theories. Second, each time a new instance of a previously-seen concept is presented, participants must remember the examples they have already seen in order to determine the similarities among them, so in this sense massed presentation might benefit induction by allowing people to notice patterns among examples from the same concept.

A noteworthy fact about the stimuli used here is that a person could generalize about an artist's style reasonably well from any one of the paintings, so all subsequent presentations served to reinforce and revise an already-present representation. Therefore, the first of the memory factors mentioned above, remembering concepts from the study phase to the test phase, should play a stronger role. This contrasts with designs like that of Kurtz and Hovland (1956), in which there were specific rules and participants did not know which were relevant in categorization until after seeing multiple examples of each concept. In this case no meaningful generalizations could be drawn from a single example, so the second memory factor, remembering items between study presentations, should be more influential. Results from these two studies suggest that this distinction is indeed an important one, for Kurtz and Hovland (1956) found higher accuracy from massed training than spaced, whereas the present experiments found the reverse.

The key observation from this difference is that in order to learn concepts from examples, one must be aware of which features are relevant in characterizing which concepts. Unless it is clear what about an instance is important to remember, people have to try to remember everything, so that they can later check for similarities when other instances are presented. The results of Experiment 2, however, suggest that people are not encoding each entire image and then finding patterns among them, since correct identification was equally likely for the images they had not seen as for the ones they had.

It is possible that in the experiments reported here, the inductive task is relatively unrelated to the memory task. That is, the concepts could be remembered in terms of general characteristics rather than specific features, so that each presentation of an example amounts roughly to a repetition of an already known concept. This explanation would account for the equal performance on learned and new images in Experiment 2, because only an overall idea is being encoded and not individual examples. It would also account for the absence of any difference in the spacing effect for low-discriminable and high-discriminable concepts: low-discriminable ones should be more difficult to generalize from, but this factor should not be influenced by the placement of contrastive examples. The experiment by Kurtz and Hovland (1956), in contrast, *required* fairly detailed memory for the individual instances in order to generalize at all, so it would be expected that spacing would be detrimental to the process of induction.

Conclusion

This paper has presented evidence that the spacing effect can occur even in learning concepts of which the presented examples differ from one another. The results do not support several proposed theories of the spacing effect, and suggest that different mechanisms are involved depending on what sort of concepts are being learned. It will be important for future work to address the relationship between spacing of presentations and specific types of inductive learning.

References

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354-380.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior*, *21*(4), 451-467.

Dellarosa, D., & Bourne, L. E., Jr. (1985). Surface form and the spacing effect. *Memory & Cognition*, *13*(6), 529-537.

Gagné, R. M. (1950). The effect of sequence of presentation of similar items on the learning of paired associates. *Journal of Experimental Psychology*, *40*(1), 61-73.

Gay, L. R. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. *Journal of Educational Psychology*, *64*(2), 171-182.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*(2), 95-112.

Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(2), 411-422.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the enemy of induction? *Psychological Science*, *19*(6), 585-592.

Kurtz, K. H., & Hovland, C. I. (1956). Concept learning with differing sequences of instances. *Journal of Experimental Psychology*, *51*(4), 239-243.

Medina, T. N., Trueswell, J., & Gleitman, L. (2008). *When the shoe fits: Cross-situational learning in realistic learning environments.* Paper presented at the 33rd Annual Boston University Conference on Language Development.

Shea, C. H., Lai, Q., Black, C., & Park, J.-H. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human Movement Science*, *19*(5), 737-760.

Tenenbaum, J. B. (1999). *A Bayesian framework for concept learning*. Unpublished doctoral dissertation, MIT.