Base-Rate Neglect in Pigeons: Implications for Memory Mechanisms

Thomas R. Zentall (zentall@pop.uky.edu)
Tricia S. Clement (tcharb1@pop.uky.edu)
Department of Psychology, University of Kentucky
Lexington, KY 40506 USA

Abstract

In delayed matching-to-sample, there is an initial (or sample) stimulus and two test (or comparison) stimuli. When pigeons are trained to match, they presumably choose between the two comparison stimuli according to their ability to remember the sample. When the sample cannot be remembered, comparison choice should depend on the history of reinforcement associated with each of the comparison stimuli (i.e., the comparison base rates). In the present research, pigeons acquired two matching tasks in which samples S1 and S2 were each associated with one of two comparisons C1 and C2 (equal experience with both trial types), and samples S1 and S3 were each associated with one of two other comparisons C3 and C4 (equal experience with both trial types). As the retention interval increased, the pigeons showed a bias to choose the comparison (C1 or C3) associated with the more frequently occurring sample (S1). Thus, pigeons are sensitive, not just to the probability of reinforcement associated with the each of the comparison stimuli (i.e., the base rates) which were equal, but also to the (irrelevant) likelihood that each of the samples was presented (i.e., base-rate neglect).

Introduction

Humans are known to underestimate the effect of base rates associated with probability of being correct (Kahneman & Tversky 1972). In a classic problem proposed by Tversky and Kaneman (1980, p. 62), participants are told that 85% of the taxis in a city are green while only 15% are blue (the base rates). They are also told that a witness to a hit-and-run accident involving a taxi identified the taxi as blue. Furthermore, they are told that under similar conditions witnesses correctly identify the color of a taxi 80% of the time. When participants are then asked, "What is the probability that the taxi involved in the accident was actually blue?" most of them say that it is very likely that the taxi is blue. In making this judgement the participants fail to consider sufficiently the base-rate probabilities. When base rate is considered, the conditional probability of correctly identifying a blue taxi is p(blue|judgement correct) = p(blue) x p(correct) = .15 x .80 = .12, whereas the probability of saying it was blue when it actually was green is p(green|judgement incorrect) = p(green) x p(incorrect) = .85 x .20 = .17. This means that the probability of being correct under these conditions is only .12/(.12 + .17) = .41, or less than 50%. Thus, humans often fail to consider sufficiently the probability of being correct in the absence of the eye-witness information. Although there are certain conditions under which humans can be induced to perform more accurately (e.g., Gigerenzer & Hoffrage, 1995), base-rate neglect is likely responsible form any exaggerated fears such as air travel, walking the streets of New York City, and having one's children killed at school by a fellow student.

An analogous situation can be designed for an animal using a matching-to-sample task. Matching-to-sample is a conditional discrimination in which the identity of the initial or sample stimulus indicates which of two (or more) test or comparison stimuli is correct (Skinner, 1950). According to Hartl and Fantino (1996), comparison choice for pigeons should depend on two factors, the relative probabilities of reinforcement associated with the comparisons (i.e., the base rates) in the absence of the sample, and the conditional probability of each comparison being correct given presentation of one of the samples (i.e., the actual sample event or the evidence, given the base rates). In the case of matching-to-sample, the probability of reinforcement given the sample is typically 1.0. This ensures that the task has been adequately acquired and that the contingencies have been adequately experienced. Biases can be introduced by manipulating the ratio of samples and the probability of reinforcement for choices of the matching comparison (see Goodie & Fantino, 1995, for similar findings with humans, but see also Goodie & Fantino, 1996, for exceptions).

Control by the comparisons alone can be increased by degrading the samples at the time of comparison choice (i.e., by increasing the probability of poor memory, or in the taxi example, of an identification error). One way to degrade the samples is by introducing a delay between the offset of the sample and the onset of the comparisons. Assuming that the comparison stimuli are correct equally often over trials, and that the probability of reinforcement is the same for a correct response to each comparison, one would expect that with increasing delay, the slopes of the pigeons' retention functions would be quite similar (see Grant, 1991; White &

Wixted, 1999).

The analog to base rate in a matching task is the probability of being correct in the absence of information about the sample (i.e., the relative probability of reinforcement associated with each of the comparison stimuli). According to White and Wixted (1999), pigeons should be sensitive to base-rate probabilities, but generally the base rates and the probability of sample presentation are the same (both generally 0.5). In the present experiment we asked if pigeons are able to estimate the probability of a correct comparison response when the sample probabilities are different from the base rates. There are a number procedures that might be used to manipulate the relative frequency of sample (S) presentation while maintaining equal probability of reinforcement for comparison (C) choice (i.e., equal base rates). In the present experiment, we chose to introduce a second 2-sample-2-comparison matching task. Each of the two matching tasks involved a different pair of comparison stimuli but the two tasks shared a common sample. Thus, the two tasks can be represented S1-C1, S2-C2 and S1-C3, S3-C4 (with C1 and C2 always appearing together and C3 and C4 always appearing together). If each of the four trial types appears equally often, each of the comparisons would be associated with reinforcement on 25% of the reinforced trials. However, the same would not be true of the samples. S2 and S3 would each be presented on 25% of the trials, whereas S1 would be presented on 50% of the trials. Under conditions with no delay, one would expect a high level of matching accuracy and no bias. But if a delay is inserted between the offset of the sample and the onset of the comparisons, errors should increase. If comparison choice depends on the reinforcement contingencies associated with comparison choice, errors should not result in a comparison bias. In the absence of memory for the sample, the probability of reinforcement of comparison choice should be 50% for either comparison in either task. Furthermore, if there is memory for the sample, the conditional probability of reinforcement associated with comparison choice should be the same for either comparison in either matching task. However, if pigeons show a bias by using their reference memory of sample presentations, they should access more instances of S1 than of either S2 or S3 and a bias to choose C1 and C3 may result.

Method

Subjects

The subjects were eight White Carneaux pigeons, purchased as retired breeders (5-8 years old) from the Palmetto Pigeon Plant (Sumter, SC). The pigeons were maintained at 80% of their free-feeding body weights throughout the experiment and were caged individually with grit and water continually available in the home cage. The pigeons were maintained on a 12:12-h, light:dark cycle. All pigeons had previously served in an unrelated study involving simple simultaneous discriminations.

Apparatus

The experiment was conducted in a standard BRS/LVE (Laurel, MD) sound attenuating pigeon test chamber. Three rectangular response keys (2.5 cm high x 3 cm wide and 1 cm apart) were aligned horizontally and centered on the response panel. Mounted behind each response key was a 12-stimulus inline projector (Industrial Electronics Engineering, Series 10, Van Nuys, CA,) that could project a red hue or a green hue onto the any of the three response keys or a plain white field onto the center response key. In addition, the left and right projectors could project a white circle and a white dot. A houselight located at the center of the chamber ceiling provided general illumination. A rear-mounted grain feeder was centered horizontally on the response panel midway between the pecking keys and the floor of the chamber. When operated, the feeder was accessible through a 5.0 x 5.5 cm lit aperture in the response panel. Reinforcement consisted of 2.0-s access to Purina Pro Grains. White noise and an exhaust fan mounted on the outside of the chamber masked extraneous noise. The experiment was controlled by a microcomputer located in an adjacent room.

Procedure

Training All pigeons were placed directly on 0-s-delay matching-to-sample training. At the beginning of each trial, the center key (sample) was illuminated. Following 10 responses to the sample, the sample was turned off and the side (comparison) keys were illuminated. Comparison stimuli were presented randomly with respect to location, with the restriction that a particular hue could not occur on the same side key for more than three consecutive trials. One response to either comparison constituted a choice and terminated the trial. Correct comparison responses resulted in a 2-sec presentation of food and a 10-sec intertrial interval. Incorrect choices resulted in the 10-sec intertrial interval alone.

For each pigeon, training consisted of a hybrid matching task involving three sample stimuli (one per trial) and two pairs of comparison stimuli (one pair on each trial). On one fourth of the trials, one of the hues served as the sample (S1) with red and green comparison stimuli (C1 and C2) on the side keys and, for example, red was correct. On another fourth of the trials, a different hue sample (S2) was presented with the red and green comparison stimuli and, for example, green was correct.

On half of the remaining trials, S1 was again presented as the sample and circle and dot were presented as the comparisons (with, for example, dot correct). On the remaining fourth of the trials a third hue was presented as the sample (S3) and circle and dot were presented as the comparisons (with circle correct).

The three sample hues were counterbalanced such that each hue served as the one-to-many sample for 2-3 pigeons and each of the remaining samples was associated with the hue comparisons for at least one pigeon. Sessions consisted of 96 trials and were conducted 6 days a week. For each pigeon, criterion was met when the correct comparison for each trial type was chosen on at least 90% of those trials for two consecutive sessions. Following criterion performance, each pigeon received five sessions of overtraining.

Retention test On the following session, each pigeon was transferred to a mixed-delay matching procedure in which the offset of the sample was separated from the onset of the side keys by a dark retention interval of 0, 2, 4, or 8 s. For each of the trial types, there was an equal number of trials involving each retention interval. The retention test consisted of 2 sessions and the reinforcement contingencies were the same as they were during training. In all analyses of results, the .05 level of statistical significance was adopted.

Results

Training

Sessions to criterion (two successive sessions at 90% correct) for the one-to-one portion of the task was 10.2 when the comparisons were hues and 11.1 when the comparisons were shapes. Sessions to criterion for the one-to-many portion of the task was 13.6 when the comparisons were hues and 13.8 when the comparisons were shapes. A mixed-effect analysis of variance performed on the acquisition scores, with task (one-to-one vs. one-to-many) and comparison dimension (hues vs. shapes) as factors, indicated that neither effect nor the interaction was statistically reliable, $F(1,7) = 2.08$, >1, and >1, respectively.

Retention Test

Data from the retention test were pooled over the 2 test sessions and were subjected to a repeated-measures ANOVA, with task component (one-to-one vs. one-to-many) and Delay (0, 2, 4, and 8 sec) as factors. Most critically, the ANOVA indicted that there was a significant Task Component x Delay interaction, $F(3,21) = 4.37$. There was also a significant effect of Delay, $F(3,21)$ 44.01. The effect of Task Component was not quite significant, $F(1,7) = 4.79$. The retention data are presented in Figure 1.

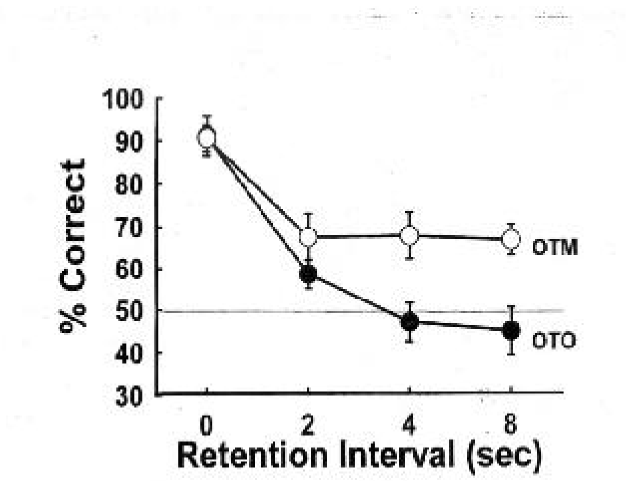

Figure 1. Retention functions following training in which two and samples, S1 and S2, were associated with comparison stimuli, C1 C2, respectively and S1 and S3 were associated with comparisons C3 and C4, respectively. Thus, S2 and S3 were involved in one-to-one matching (OTO) with C2 and C4, while the third sample, S1, was associated with two comparison stimuli, C1 and C3 (one-to-many matching, OTM). In training and test, each comparison was associated with reinforcement on 50% of the trials and C1 and C2 always appeared together as did C3 and C4.

Discussion

According to traditional instrumental views of conditional discrimination learning (i.e., Hartl & Fantino, 1996), the probability of a comparison choice should be determined by the conditional probability associated with each comparison stimulus, given the sample, and, if the sample is unavailable or forgotten, with the probability of reinforcement associated with each comparison (independently of the sample). Thus, the choice a particular comparison (e.g., C1) should depend on both the number of sample-comparison pairings (e.g., S1-C1) that are followed by reinforcement, as well as the number of reinforcements associated with that comparison, independent of the sample (Wixted, 1993). In the present experiment, the conditional probability of reinforcement associated with each of the comparisons, given one of the samples, was equal. Furthermore, the probability of reinforcement associated with choice of either comparison was also equal. Thus, in the present experiment, given presentation of C1 and C2, the only relevant sample-comparison associations determining comparison choice should be S1-C1 and S2-C2. If so, delay-induced sample degradation should have had a symmetrical effect on comparison choice and the retention functions should have been parallel and overlapping.

In the present experiment, clearly divergent retention functions were found. These results require the modification of current theories of delayed conditional discrimination performance (e.g., White & Wixted, 1999) because pigeons choice behavior is influenced not only by the probability of reinforcement associated with responding to each of the comparison stimuli and to the conditional probabilities associated with choice of the comparison stimuli as a function of memory for the sample but also by the relative frequencies of the samples. When delays are introduced, as the delay increases, pigeons have an increasing tendency to select the comparison associated with the more frequently presented sample, even though that sample was not presented more often than the alternative sample in the context of either comparison pair. It is as if, on trials when memory for the sample is poor, presentation of the comparisons causes the pigeons to consult their reference memory for the overall probability of sample presentation (independent of the comparison pair).

Of broader interest, such use of reference memory in delayed matching may be a general phenomenon. However, the use of sample frequency independently of other more relevant measures may be apparent only with a design such as that used in the present research because in the more typical design, either hypothesis makes the same prediction.

Alternatively, in the present experiment, although the pigeons had equal opportunity to acquire each of the four sample-comparison associations, the more frequent presentations of the S1 sample could have allowed it to be more efficiently coded, better maintained in memory, or more easily retrieved from memory. That is, at the time of comparison choice, when the S1 stimulus had been the sample, it may have been more accessible than the S2 or S3 stimuli were when they had been the sample. But if the difference in slope of the retention functions was attributable to differences in sample accessibility at the time the comparisons were presented, both the S1 and the S2/S3 functions should have approached 50% correct with increasing retention interval. Instead, the S1 retention function appears to have leveled off, while the S2/S3 retention function declines below chance at delays of 4 and 8 sec. Such retention functions suggest that rather than better retrieval of the S1 sample, the pigeons developed a comparison bias to choose the comparison associated with the more frequently presented sample.

This comparison bias in pigeons is analogous to the base-rate neglect shown by humans when they fail to consider sufficiently the base-rate probability of occurrence of an event.

References

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review, 102, 684-704.

Goodie, A. S., & Fantino, E. (1995). An experimentally derived base-rate error in humans. Psychological Science, 6, 101-106.

Goodie, A. S., & Fantino, E. (1996). Learning to commit or avoid the base-rate error. Nature, 380, 247-249.

Grant, D. S. (1991). Symmetrical and asymmetrical coding of food and no-food samples in delayed matching in pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 17, 186-193.

Hartl, J. A., & Fantino, E. (1996). Choice as a function of reinforcement ratios in delayed matching to sample. Journal of the Experimental Analysis of Behavior, 66, 11-27.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgement of representativeness. Cognitive Psychology, 3, 430-453.

Skinner B. F. (1950). Are theories of learning necessary? Psychological Review, 57, 193-216.

Tversky, A, & Kahneman, D. (1980). Causal schemas in judgements under uncertainty. In M. Fishbein (Ed.), Progress in social psychology (pp. 49-72). Hillsdale, NJ: Erlbaum.

White, K. G., & Wixted, J. T. (1999). Psychophysics of remembering. Journal of the Experimental Analysis of Behavior, 71, 91-113.

Wixted, J. T. (1993). A signal detection analysis of memory for nonoccurrence in pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 19, 400-411.

Author Notes

This research was supported by National Institute of Mental Health Grants 55118 and 59194.

Correspondence should be addressed to Thomas R. Zentall, Department of Psychology, University of Kentucky, Lexington, KY 40506-0044. Email, zentall@pop.uky.edu