

First trial rewards promote 1-trial learning and prolonged memory in pigeon and baboon

Robert Cook^a and Joël Fagot^{b,1}

^aDepartment of Psychology, Tufts University, Medford, MA 02155; and ^bLaboratory of Cognitive Psychology, Centre National de la Recherche Scientifique, Université de Provence, UMR6146, Pôle 3C, 13331 Marseille Cedex 1, France

Edited by Charles R. Gallistel, Rutgers The State University of New Jersey, Piscataway, NJ, and approved April 10, 2009 (received for review March 26, 2009)

There is a long-standing debate in educational settings on the influence of positive and negative consequences on learning. Although positive rewards seem desirable from an ethical perspective, 1-trial learning has been best demonstrated in the animal literature with tasks using highly salient negative consequences, such as shock or illness, and so far only in tasks requiring the acquisition of a singular stimulus-response association. Here we show that pigeons and baboons can concurrently learn, in a cognitively challenging memorization task, hundreds of picture-response associations after a single exposure and that this rapid learning is better promoted by a positive outcome after the first picture presentation. Further, the early positive outcomes had beneficial effects on the memory of learned acquisitions that was detectable up to 6–8 months after initial training. Beyond their significance for educational policies, these findings suggest that the psychological and brain mechanisms controlling rapid, often 1-trial, learning have a long evolutionary history. They may represent the phylogenetic precursor for the disproportionate impact of first impressions in humans and the phenomenon of fast word learning in children.

monkey | positive reinforcement | fast mapping | picture processing | bird

A long standing debate in educational settings concerns the relative efficacy of positive and negative outcomes in promoting learning and in dealing with the modification of problem behaviors (1, 2). Despite the desirability of using positive rewards, the best animal evidence of rapid learning (i.e., 1-trial learning) come from experiments involving highly salient negative consequences, such as shock or illness (3–7). However, positive rewards may also have important effects on the speed of learning, because rewards have the advantage of directly communicating and confirming what an animal should do as opposed to what behavior should not be performed. Extensive brain structures are known to be associated with positive rewards (8–10).

To further explore the role of positive reinforcement in producing rapid learning in a cognitively demanding situation, we examined learning in a picture memorization task testing 2 highly visual, but distantly related species, pigeons and baboons (11, 12). Here we show that these 2 species can learn very large numbers of concurrent picture-response associations after only a single presentation and that this 1-trial learning and its long-term retention over months was promoted by an early positive reinforcement at the first presentation of each image.

Results

Two pigeons (BF and Linus) and 2 baboons (B03 and B09) had to continually learn and recall right- or left-choice responses to increasing number of pictures presented in a 2-alternative choice task. Correct answers were rewarded by positive reinforcements (food), whereas incorrect answers were mildly punished by a short timeout before the next trial. Because the original mapping of the pictures to the responses was arbitrary, the task required rote learning of each picture-response combination. Sessions consisted of learning trials presenting new picture-response associations mixed with trials retesting memory for already

learned pictures. Over 3 to 5 years of daily testing, new pictures were continually introduced as groups of initially 20 and then 30 randomly chosen pictures, with each animal eventually acquiring thousands of picture-response associations. Earlier reports using this procedure presented information on the long-term memory capacity (11). This paper presents information on the learning of these numerous picture-response associations.

We analyzed the learning curves for each picture-response association by measuring accuracy over the first 10 acquisition trials (presentations 2–11). Any trials that had involved altered stimuli designed to test the nature of stimulus control (scrambled, gray, inverted, etc.) were excluded from the data set, leaving 6,090 and 6,358 separate instances of picture-response learning for B03 and B09, and 2,700 and 1,727 for Linus and BF, respectively. The first trial with each picture was excluded, as it required a guess because of the randomized picture-response assignments (mean first trial accuracy = 50.4%, range 49.1%–50.8%). The pigeons exhibited an initial choice bias to the left side (Linus = 79.1%; BF = 82%), which is explained by the differential outcome procedure involving delivery of the more preferred food on the left side. There was by contrast no strong response bias in the baboons' initial guesses, with B03 making 46% and B09 making 58.5% of their initial choices to the left choice.

Both the pigeons and baboons learned a sizeable proportion of these individual picture-response associations within a single trial. One-trial learning was defined as exhibiting no errors over the first 10 trials after the initial guess. We used this strict criterion because runs of this length were extremely unlikely by chance (probability of occurrence, $P = 0.00097$). In addition, individual performance in these first 10 trials showed an excellent correlation with later performance over subsequent trials with the same pictures (individual range of R values, 0.88–0.92, all $P < 0.001$). With this criterion, the baboons showed 1,533 errorless acquisitions (B09 = 750; B03 = 783) representing 12.3% of their acquisitions. The pigeons had 915 errorless acquisitions (BF = 373; Linus = 541) representing 20.8% of their acquisitions (Fig. 1 *A* and *B*). Because of the pigeons' initial choice bias to 1 side, we separately examined the proportion of errorless acquisitions for both right and left initial choices. The proportion of errorless acquisitions remained similar regardless of initial choice (mean errorless acquisitions following initial left choice, 20.2%; following right choice, 24.1%). This equivalence indicates that this species' initial bias had no influence on the relative frequency of errorless acquisitions. Although they had no bias, the same equivalence was found for the baboons, too. The frequency of these errorless acquisitions for both species was significantly greater than expected by guessing (<7 occurrences for the monkeys and <3 for the pigeons). Because a correction procedure was used to prevent sustained response biases, Fig. 1 *B* and *C* show the percentage of acquisitions as a function of

Author contributions: R.C. designed research; R.C. and J.F. performed research; R.C. and J.F. analyzed data; and R.C. and J.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: joel.fagot@univ-provence.fr.

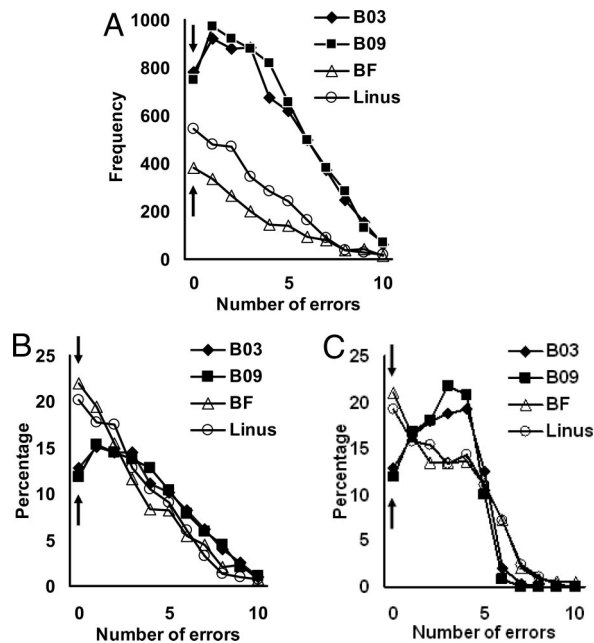


Fig. 1. Distribution of errors over the first 10 acquisition trials. Shown is the frequency (A) of all picture-response acquisitions for the 2 pigeons (open symbols) and baboons (filled symbols) as a function of the number of choice errors made during the first 10 presentations (excluding the mandatory first trial guess) of each image with correction trials excluded. The percentage of acquisitions computed without (B) and with (C) correction trials included is shown. The arrows in A–C point to the errorless acquisitions indicative of 1-trial learning.

number of errors without and with the correction trials included. Both point to the same conclusion that a large proportion of the picture-response associations were learned extremely rapidly, often in a single trial.

One possible source of this rapid learning might be generalized responding from already learned images that looked similar and had the same assigned response as the new pictures. We tested this possibility directly by determining response assignments of the most similar earlier picture for all pictures learned without error. Similarity was determined 2 ways. The first measured similarity used the percentage of matching pixels based on having similar hue, saturation, and luminance values at each location. By using the most similar earlier learned picture determined this way, the percentage of errorless acquisitions with a common “matching” response for an earlier item was very small relative to the chance expectation (B03 = 53%; B09 = 55%; BF = 51%; Linus = 52%). The second method minimized the sum of squared differences for measurements of images’ mean spatial frequency in its RGB color bands, mean luminance, and mean pixel-based entropy. The percentage of errorless acquisitions with a common “matching” response calculated this way was similarly small (B03 = 52%; B09 = 52%; BF = 56%; Linus 50%). Thus, generalized responding from already learned pictures seems to have been minimal. A generalization hypothesis further predicts an increasing frequency of rapidly and poorly learned pictures as the stimulus-response “library” grows. Errorless acquisitions were observed within the first set of 20 new images for the pigeons and by the second new set for baboons, when any chance similarity among the pictures was low. Moreover, the frequencies of errorless acquisitions remained stable for 3 of the animals over testing (mean percent occurrence in 4 vincentized blocks of total acquisitions: 16%, 19%, 18%, 16%), with only B03 showing a monotonic increase in their frequencies (9%, 11%, 15%, 16%). For all animals there were very few acquisitions with no correct responding in the first 10 trials that would have been

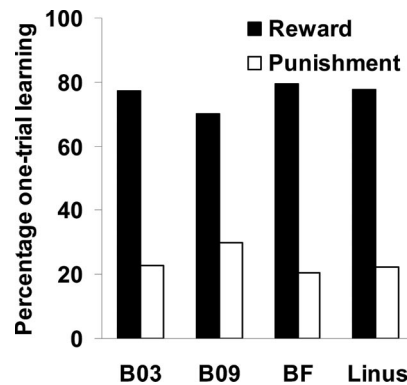


Fig. 2. Effects of a reward or punishment after the guess in the initial trials on the percentage of picture-response associations subsequently learned in 1 trial by each animal.

indicative of any strong interference from generalization (mean = 0.8%) of “mismatching” responses. There was also no increase of such very poorly learned pictures across blocks (0.9%, 0.8%, 0.9%, 0.7%) as also predicted by this hypothesis. Thus, although some generalization from earlier instances might have occurred on a limited basis, this alternative is unable to explain the pervasiveness of errorless acquisitions displayed by each species.

Additionally, we found that both species were far more likely to learn in 1 trial when their initial first-trial guess was correct and positively reinforced. Fig. 2 shows the proportion of errorless acquisitions as a function of initial first-trial outcome. For each animal, the frequency of 1-trial learning was significantly greater when the first trial was positively reinforced than not (binomial tests, all $P < 0.001$). We also examined how long this initial positive advantage lasted by computing accuracy on all subsequent memory trials that were randomly inserted by the protocol after learning was completed. For this analysis, accuracy on all memory trials was tracked over time based on initial outcome. Initially rewarded pictures supported significantly better accuracy for ≈ 30 –70 presentations of each picture after learning, depending on the animal (Fig. 3). After this point, there appeared to be no systematic difference between these 2 conditions. This reward advantage conservatively lasted at least 6 months in the pigeons and 8 months in the baboons, as determined by looking at the sequential probabilities of positive difference scores (z test for proportions, $P < 0.05$). This duration was estimated from the average time interval

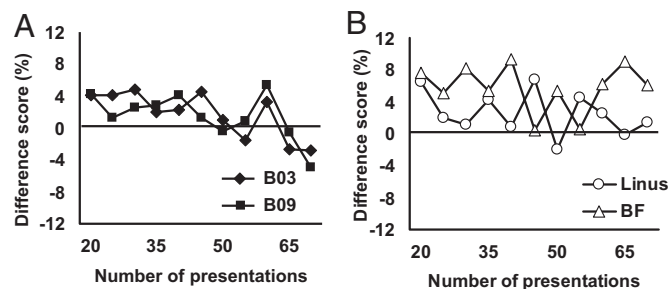


Fig. 3. Long-lasting memory benefits of the early reward. Shown are the mean difference scores for baboons (A) and pigeons (B) during postlearning memory presentations (trial presentations 20–65) as categorized by whether the initial first presentation for all trials was positively reinforced or punished. Positive values indicate that the rewarded pictures supported higher accuracy than those initially punished. This reward advantage was estimated to last at least 6 months in the pigeons and 8 months in the baboons, as determined by adjusting the number of presentations by the average time interval between these memory trials over the course of the experiment. Mean old-item accuracy over this time period (B03 = 76%, B09 = 75%, BF = 78%, Linus = 75%).

between memory trials in each species based on number of trials per session and number of pictures in the memory set.

Discussion

Overall, our results indicate that pigeons and baboons have a far greater capacity for 1-trial learning than previously believed and that an early positive experience can significantly benefit the learning and retention of picture-response associations. Thus far, 1-trial learning has been found with positive rewards involving single associations (4, 6), but here our animals revealed a capacity to often learn in 1 trial when engaged in a much more demanding task requiring that thousands of randomly mapped and competing picture-response associations be learned and remembered. Earlier experiments examining learning set in primates bear some similarity with the present results. They show their ability to learn a win-stay/lose-shift rule to solve problems after 1 presentation. The increments of 1-trial learning over time in B03 may suggest a learning set type strategy, but such an account cannot explain why our other 3 subjects showed no greater propensities to learn in 1 trial over years, and why 1-trial learning occurred very early in the research. Noticeable differences also exist between our findings and those published on learning set. First, learning set is primarily based on short-term memories for approaching or avoiding single objects. Long-term memories for dozens of such discriminations have been found (13), but the number of learned items was very much smaller than in the current research. Second, learning set is better mediated by first trial nonreward (14–16) in sharp contrast to what was observed here. Finally, learning set behavior has been very difficult to create in pigeons, although quickly formed memories of recent image familiarity can be found in this species (17). One-trial learning of stimulus-response association therefore appears a better account of our finding than either a generalization (see *Results*) or learning set account.

From a theoretical perspective, our findings present new challenges to the overriding theories of the last century that have favored gradual views of learning involving the incremental accumulation of repeated experience, as represented by the dominance of linear operator and neural network approaches (18, 19). Such incremental models can, with a significantly large learning rate parameter, mimic very rapid learning in simple associative cases. Such large learning rate parameters, however, are established to create undesirable catastrophic interference during learning and memory (20, 21). With the large number of pictures involved here and the different rates of learning observed from errorless to having many errors with each picture, any learning parameter would need to be stimulus-specific or unpredictably variable to account for what was observed.

We have previously documented that pigeons and baboons can store thousands of picture-response associations (11). Here we add that often this information was acquired and retained very rapidly and often without error. These 2 findings offer a strong support for exemplar approaches to representation in learning and memory (22, 23). The long lasting benefits in memory of the first reward observed here are unlikely based on specific memories of that first event. We rather propose that this initial outcome initiated a cascade of subsequent positive outcomes during learning that promoted the formation of strong associations, which in turn resulted in temporally extended reward-related gain in memory. In primates, such processes may involve

the amygdala (24) and prefrontal cortex (8). In pigeons, such processes likely involve the visual wulst (25). Demonstration of the long-term effect of early positive consequences calls for a special consideration during the first exposure to the to-be-learned situation in human educational settings and animal testing procedures.

Our discovery of extensive 1-trial learning in so distantly related classes of bird and primates suggests that single experience association formation is likely a phylogenetically widespread capacity millions of years old. If so, this ancient mechanism for rapid association formation and memorization may be an evolutionary precursor for the disproportionate impact of first impressions in many human situations (26) and our ability to make judgments based on thin slices of initial behavior (27). It may also have provided the initial foundation for fast word learning and lexicon growth during human ontogeny and evolution (28–30).

Materials and Methods

The subjects were 2 adult 18-year-old male baboons (*Papio papio*) and 2 adult male Silver King pigeons (*Columba livia*). All animals were tested in the same task with minimal modifications made to accommodate each species' size and natural motor and feeding behavior. The baboons were tested in a Plexiglas chamber that permitted free access to a joystick and full view of a computer screen. A trial started after monkeys placed a cursor on the fixation point. A picture was then presented central to the screen during 700 ms immediately after which the 2 response keys were illuminated. Monkeys had to select, by way of joystick manipulation, the response keys associated to the sample picture. The same procedure was followed for pigeons, except that pecks were required to the picture before making their right/left choice to 2 side-choice hoppers, 1 containing mixed grain and the other safflower, respectively. The test program was written in Visual Basic.

For each species, the 120-trials test sessions were composed of 60 old-item and 60 new-item trials. New trials involved 2 presentations of 30 recently introduced pictures that were repeated over sessions until subjects reached criterion (20 pictures were used for the first 18 sets). Because picture-response key assignments were determined on a random basis, our procedure made picture-response learning the mandatory strategy. Acquisition of these new items is the focus of this report. Once learned, each set of new-item pictures was moved to the old-item memory pool, and a new set of 30 pictures to memorize was introduced. The 60 old-item pictures of each session were randomly selected from this ever-increasing pool of previously learned picture-response associations. New-item pictures were introduced in the same order for both species. All trials were differentially reinforced. Incorrect trials gave rise to a timeout of either 3 s (baboons) or 5 s (pigeons) and were immediately presented again until a correct response was given. The number of successive correction trials was limited to 5 for baboons but was unlimited for pigeons. The number of correction trials rarely exceeded 1 in both pigeons (>5% of trials—overall mean of 0.26 correction trials per trial) and baboons (>2% of trials—overall mean of 0.24 correction trials per trial). A 3-s intertrial interval followed each trial. Accuracy and response times were recorded. Baboons had the capacity to be tested with an average of 4 sessions a day, whereas the pigeons were limited to 1 session a day. All pictures were color. The pictures were harvested from various sources and resized to 480 × 300 pixels by using Paint Shop Pro and JASC software. Use of baboons in this research was approved by The Provence Côte d'Azur Regional Ethic committee for Animal Experimental Research and the use pigeons was approved by the Tufts University IACUC.

ACKNOWLEDGMENTS. The authors thank Marc Martin (CNRS) and students (Tufts) for assistance during data collection. R.C. was supported by a grant from the National Science Foundation (0316016). J.F. was supported by European Commission Stages in the Evolution and Development of Sign Use (SEDSU) Grant 012-984 and the Eurocores "The Origin of Man, Language, and Languages" (OMLL) Program.

1. Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev Gen Psychol* 5:323–370.
2. Penney RK, Lupton AA (1961) Children's discrimination learning as a function of reward and punishment. *J Comp Physiol Psychol* 54:449–451.
3. Garcia J, Koelling RA (1966) Relation of cue to consequence in avoidance learning. *Psychon Sci* 4:123–124.

4. Armstrong CM, Devito LM, Cleland TA (2006) One-trial associative odor learning in neonatal mice. *Chem Senses* 31:343–349.
5. Davies DC, Taylor DA, Johnson MH (1988) The effects of hyperstriatal lesions on one-trial passive-avoidance learning in the chick. *J Neurosci* 8:4662–4666.
6. Bardo MT, Bevins RA (2000) Conditioned place preference: What does it add to our understanding of preclinical understanding of drug reward. *Psychopharmacology* 153:31–43.

7. Pearce JM (1978) Relationship between shock magnitude and passive-avoidance learning. *Anim Learn Behav* 6:341–345.
8. Amemori K, Sawaguchi T (2006) Contrasting effects of reward expectation on sensory and motor memories in primate prefrontal neurons. *Cereb Cortex* 16:1002–1015.
9. Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
10. Amiez C, Joseph JP, Procyk E (2006) Reward encoding in the monkey anterior cingulate cortex. *Cereb Cortex* 16:1040–1055.
11. Fagot J, Cook RG (2006) Evidence for large long-term memory capacities in baboons and pigeons and its implication for learning and the evolution of cognition. *Proc Natl Acad Sci* 103:17564–17567.
12. Cook RG, Levison DG, Gillett SR, Blaisdell AP (2005) Capacity and limits of associative memory in pigeons. *Psychon Bull Rev* 12:350–358.
13. Treichler RF (1984) Long-term retention of concurrent discriminations by monkeys. *Physiol Psychol* 12:92–96.
14. Blomquist AJ, Deets AC, Harlow HF (1973) Effects of list length and first-trial reward upon concurrent discrimination performance. *Learn Motiv* 1973:28–39.
15. Brown WL, McDowell AA, Gaylord HA (1965) Two-trial learning-set formations by baboons and stump-tailed macaques. *J Comp Physiol Psychol* 60:288–289.
16. Deets AC, Harlow HF, Blomquist AJ (1971) Effects of intertrial interval and trial 1 reward during acquisition of an object-discrimination learning in monkeys. *J Comp Physiol Psychol* 73:501–505.
17. Todd IA, Mackintosh NJ (1990) Evidence for perceptual learning in pigeons' recognition memory for pictures. *Q J Exp Psychol-A* 42:385–400.
18. Rescorla R, Wagner AR (1972) A theory of Pavlovian conditioning. Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton-Century-Crofts, New York).
19. Rumelhart DE, McClelland JL (1986) *Parallel Distributed Processing* (MIT Press, Cambridge, MA).
20. Ratcliff R (1990) Models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychol Rev* 97:285–308.
21. McCloskey M, Cohen JN (1989) Catastrophic interference in connectionist networks: The sequential learning problem. In *The Psychology of Learning and Motivation*, ed Bower GH (Academic, New York), Vol 24.
22. Chase S, Heinemann EG (2001) Exemplar memory and discrimination. *Avian Visual Cognition*, ed Cook RG. Available at www.pigeon.psy.tufts.edu/avc/chase/. Accessed March 24, 2008.
23. Kruschke JK (1992) Alcové—an exemplar-based connectionist model of category learning. *Psychol Rev* 99:22–44.
24. Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439:865–870.
25. Zeigler HP, Bischof WF (1993) *Vision, Brain, and Behavior in Birds* (MIT Press, Cambridge, MA).
26. Hamilton DL, Katz LB, Leirer VO (1980) Cognitive representations of personality impressions: Organizational processes in first impression formation. *J Pers Soc Psychol* 6:1050–1063.
27. Ambady N, Rosenthal R (1993) Half a minute: Predicting teach evaluations from thin slices of nonverbal behavior and physical attractiveness. *J Pers Soc Psychol* 64:431–441.
28. Heibeck TH, Markman EM (1987) Word learning in children: An examination of fast mapping. *Child Dev* 58:1021–1034.
29. Kaminski J, Call J, Fischer J (2004) Word learning in a domestic dog: Evidence of “fast mapping”. *Science* 304:1682–1683.
30. Clark EV (2003) *First Language Acquisition* (Cambridge Univ Press, London).