for the effects of the retrieval-extinction manipulation on fear memories (*38*), whereas infralimbic plasticity is critical for maintenance of aversive and appetitive extinction memories (*44*, *45*). In the experiments reported here, we found that repeated cocaine-cue retrieval 10 min before daily extinction sessions potentiated the opposite effects of extinction training alone on PKMζ in the infralimbic cortex (increased expression) versus basolateral amygdala (decreased expression) (fig. S6). These findings are consistent with a "dual" effect of the memory retrieval-extinction manipulation on both consolidation of extinction memory and reconsolidation of cue memories.

Investigators have identified several ways to disrupt cue-memory reconsolidation or strengthen extinction learning (*7*, *46*). However, their potential as preventive treatments for addiction is limited because they often rely on pharmacological agents that are either not approved for human use or that can cause problematic side effects. We used established animal models of drug relapse and a standard human laboratory procedure for drug-induced craving to assess a purely behavioral procedure to decrease the motivational effects of drug cues during abstinence. The memory retrieval-extinction procedure decreased cue-induced drug craving and (extrapolating from our rat data) perhaps could reduce the likelihood of cue-induced relapse during prolonged abstinence periods. If our procedure weakens the original drug-cue memories rather than solely facilitating extinction, it may overcome the contextual renewal problems that have limited the clinical effectiveness of traditional extinction procedures (*4*), although this possibility needs empirical evaluation in human addicts. Last, although the cellular mechanisms and brain circuits underlying the long-lasting effects of the retrieval-extinction procedure on drug relapse and craving

remain to be elucidated, our data point to a role for PKMζ activity in the infralimbic cortex and basolateral amygdala.

References and Notes

1. C. P. O'Brien, R. N. Ehrman, J. W. Ternes, in *Behavioral Analysis of Drug Dependence*, S. Goldberg, I. Stolerman, Eds. (Academic Press, Orlando, FL, 1986), pp. 329–356.
2. J. Stewart, H. de Wit, R. Eikelboom, *Psychol. Rev.* **91**, 251 (1984).
3. G. A. Marlatt, *Addict. Behav.* **15**, 395 (1990).
4. C. A. Conklin, S. T. Tiffany, *Addiction* **97**, 155 (2002).
5. M. E. Bouton, *Biol. Psychiatry* **52**, 976 (2002).
6. N. C. Tronson, J. R. Taylor, *Nat. Rev. Neurosci.* **8**, 262 (2007).
7. A. L. Milton, B. J. Everitt, *Eur. J. Neurosci.* **31**, 2308 (2010).
8. J. L. Lee, P. Di Ciano, K. L. Thomas, B. J. Everitt, *Neuron* **47**, 795 (2005).
9. C. A. Miller, J. F. Marshall, *Eur. J. Neurosci.* **21**, 1385 (2005).
10. K. Nader, G. E. Schafe, J. E. Le Doux, *Nature* **406**, 722 (2000).
11. Y. Dudai, *Curr. Opin. Neurobiol.* **16**, 174 (2006).
12. C. M. Alberini, *Trends Neurosci.* **28**, 51 (2005).
13. A. L. Milton, J. L. Lee, B. J. Everitt, *Learn. Mem.* **15**, 88 (2008).
14. J. A. Wouda *et al.*, *Front. Behav. Neurosci.* **4**, 179 (2010).
15. F. Q. Li *et al.*, *J. Neurosci.* **30**, 10351 (2010).
16. H. Sanchez, J. J. Quinn, M. M. Torregrossa, J. R. Taylor, *J. Neurosci.* **30**, 4401 (2010).
17. M. H. Milekic, S. D. Brown, C. Castellini, C. M. Alberini, *J. Neurosci.* **26**, 3010 (2006).
18. J. L. Lee, A. L. Milton, B. J. Everitt, *J. Neurosci.* **26**, 5881 (2006).
19. M. H. Monfils, K. K. Cowansage, E. Klann, J. E. LeDoux, *Science* **324**, 951 (2009).
20. D. Schiller *et al.*, *Nature* **463**, 49 (2010).
21. Y. Shaham, U. Shalev, L. Lu, H. de Wit, J. Stewart, *Psychopharmacology (Berl.)* **168**, 3 (2003).
22. M. E. Bouton, D. Swartzentruber, *Clin. Psychol. Rev.* **11**, 123 (1991).
23. H. S. Crombag, J. M. Bossert, E. Koya, Y. Shaham, *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **363**, 3233 (2008).
24. R. E. See, *Eur. J. Pharmacol.* **526**, 140 (2005).
25. H. de Wit, J. Stewart, *Psychopharmacology (Berl.)* **75**, 134 (1981).
26. U. Shalev, J. W. Grimm, Y. Shaham, *Pharmacol. Rev.* **54**, 1 (2002).
27. Y. Shaham, L. K. Adamson, S. Grocki, W. A. Corrigall, *Psychopharmacology (Berl.)* **130**, 396 (1997).
28. H. S. Crombag, Y. Shaham, *Behav. Neurosci.* **116**, 169 (2002).
29. T. C. Sacktor, *Nat. Rev. Neurosci.* **12**, 9 (2011).
30. R. Shema, T. C. Sacktor, Y. Dudai, *Science* **317**, 951 (2007).
31. Y. Y. He *et al.*, *Neuropsychopharmacology* **36**, 1972 (2011).
32. Y. Q. Li *et al.*, *J. Neurosci.* **31**, 5436 (2011).
33. R. Sinha, T. Fuse, L. R. Aubin, S. S. O'Malley, *Psychopharmacology (Berl.)* **152**, 140 (2000).
34. M. Eisenberg, T. Kobilo, D. E. Berman, Y. Dudai, *Science* **301**, 1102 (2003).
35. K. Nader, G. E. Schafe, J. E. LeDoux, *Nat. Rev. Neurosci.* **1**, 216 (2000).
36. L. Diergaarde, A. N. Schoffelmeer, T. J. De Vries, *Eur. J. Pharmacol.* **585**, 453 (2008).
37. K. Nader, O. Hardt, *Nat. Rev. Neurosci.* **10**, 224 (2009).
38. R. L. Clem, R. L. Huganir, *Science* **330**, 1108 (2010).
39. M. Soeter, M. Kindt, *Learn. Mem.* **18**, 357 (2011).
40. W. Y. Chan, H. T. Leung, R. F. Westbrook, G. P. McNally, *Learn. Mem.* **17**, 512 (2010).
41. P. J. Hernandez, A. E. Kelley, *Learn. Mem.* **11**, 748 (2004).
42. B. M. Graham, R. Richardson, *Behav. Neurosci.* **124**, 337 (2010).
43. M. Davis, K. Ressler, B. O. Rothbaum, R. Richardson, *Biol. Psychiatry* **60**, 369 (2006).
44. G. J. Quirk, D. Mueller, *Neuropsychopharmacology* **33**, 56 (2008).
45. J. Peters, P. W. Kalivas, G. J. Quirk, *Learn. Mem.* **16**, 279 (2009).
46. J. R. Taylor, P. Olausson, J. J. Quinn, M. M. Torregrossa, *Neuropharmacology* **56** (suppl. 1), 186 (2009).

Orthographic Processing in Baboons (*Papio papio*)

Jonathan Grainger,* Stéphane Dufau, Marie Montant, Johannes C. Ziegler, Joël Fagot

Skilled readers use information about which letters are where in a word (orthographic information) in order to access the sounds and meanings of printed words. We asked whether efficient processing of orthographic information could be achieved in the absence of prior language knowledge. To do so, we trained baboons to discriminate English words from nonsense combinations of letters that resembled real words. The results revealed that the baboons were using orthographic information in order to efficiently discriminate words from letter strings that were not words. Our results demonstrate that basic orthographic processing skills can be acquired in the absence of preexisting linguistic representations.

Reading is a complex process that starts with the extraction of detailed visual information, which is used to access the sounds (phonology) and the meanings (semantics) of words. Before they process the phonological and semantic information, readers of languages that use an alphabetic script must first process the elementary visual features of the word's constituent letters and assign these different letter identities to specific positions in the word. The computation of letter identities and their relative positions is referred to as orthographic processing, and there is a large consensus today that such processing represents the first "language-specific" stage of the reading process that follows the operations involved in the control of eye movements (bringing words into the focus of central vision) and early visual processing (enabling visual feature extraction; Fig. 1A) (*1–4*). In the present study, we examined whether the ability to efficiently process orthographic information can operate in the absence of prior linguistic knowledge.

Orthographic processing lies at the interface between the visual processing and the linguistic processing involved in written language

CNRS and Aix-Marseille University Laboratoire de Psychologie Cognitive, Fédération de Recherche 3C, Brain and Language Research Institute, Aix-Marseille University and CNRS, 3 Place Victor Hugo, 13331 Marseille, France.

*To whom correspondence should be addressed. E-mail: jonathan.grainger@univ-amu.fr

comprehension. The vast majority of research on visual word recognition, however, has ignored the status of printed words as visual objects, focusing mainly on how letter-level information maps onto higher-level linguistic properties (phono-logical, morphological, semantic, and syntactic) (5). The discovery that orthographic processing is achieved by neural structures in the left ventral occipitotemporal cortex (6, 7), a region that is bilaterally associated with object and face pro-cessing, has encouraged a reconsideration of the role of basic object identification processes in visual word recognition. In the light of this finding, Dehaene and colleagues proposed that skilled reading involves an adaptation of general object-

Fig. 1. Teaching baboons to recognize words. (**A**) Skilled readers use an orthographic code to recognize words, mapping elementary visual features, such as lines of different orientation (here features contained in the word "WASP"), onto whole-word orthographic representations via some form of letter-level code (1–4, 9). (**B** and **C**) While maintained in their social group, the baboons had free access to computer-controlled operant conditioning setups with touch screen technology (13). (C) The baboons were trained to recognize four-letter English words and distinguish them from strings of letters that are not English words, such as "STOD." Baboons responded by touching either the cross or the oval shape presented immediately after the word or nonword. After a correct response, a blank screen was presented and baboons received a food reward (dry wheat). A green screen was presented for 3 s after an incorrect response. We asked whether baboons would use an orthographic code, as described in (A), in order to discriminate words from nonwords.
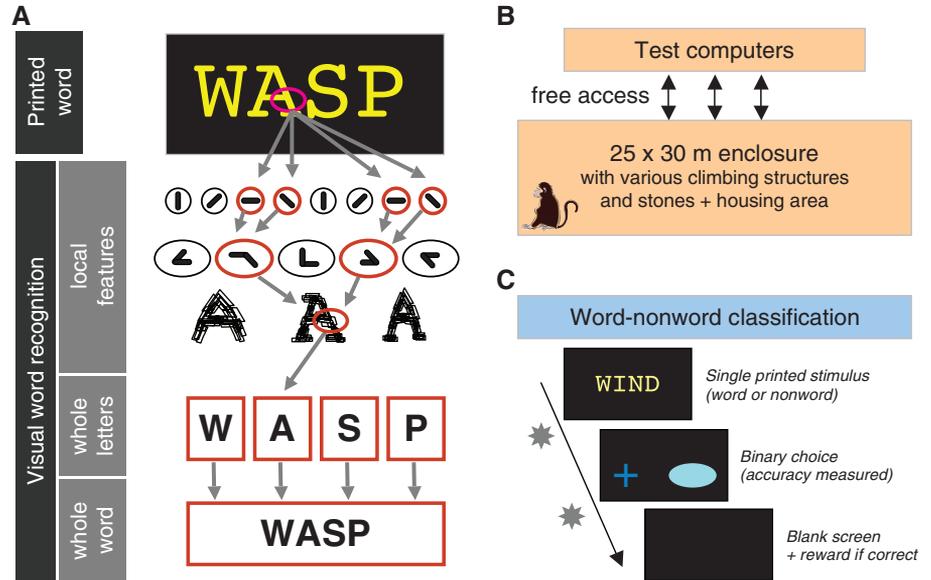
Fig. 2. Successful word-nonword discrimination in baboons. (**A**) Accuracy for words (e.g., DONE, LAND, THEM, VAST) and (**B**) nonwords (e.g., DRAN, LONS, TELK, VIRT) was calculated for blocks of 2000 consecutive trials (except for the last block) separately for each baboon (here and in Figs. 3 and 4, baboons are indicated by their abbreviated names: DAN, ART, CAU, DOR, VIO, and ARI). The results of a signal detection analysis are shown in (**C**) (sensitivity: baboons' ability to discriminate words from nonwords) and (**D**) (bias: baboons' inclination to answer "word" or "nonword"). During the first block of 2000 trials, numerical estimates of bias show that each baboon predominantly chose one of the two possible responses resulting in a "word" or "nonword" bias and low sensitivity. After 2000 trials, the baboons started to perform accurate word-nonword classification by responding "word" to repeated stimuli and "nonword" to novel stimuli as shown by above-zero sensitivities and bias values close to zero. Baboons attained an accuracy level of about 75%. Error bars in (A) and (B) correspond to the 95% binomial confidence interval, which are also displayed in gray for chance performance (see supplementary materials for more details).
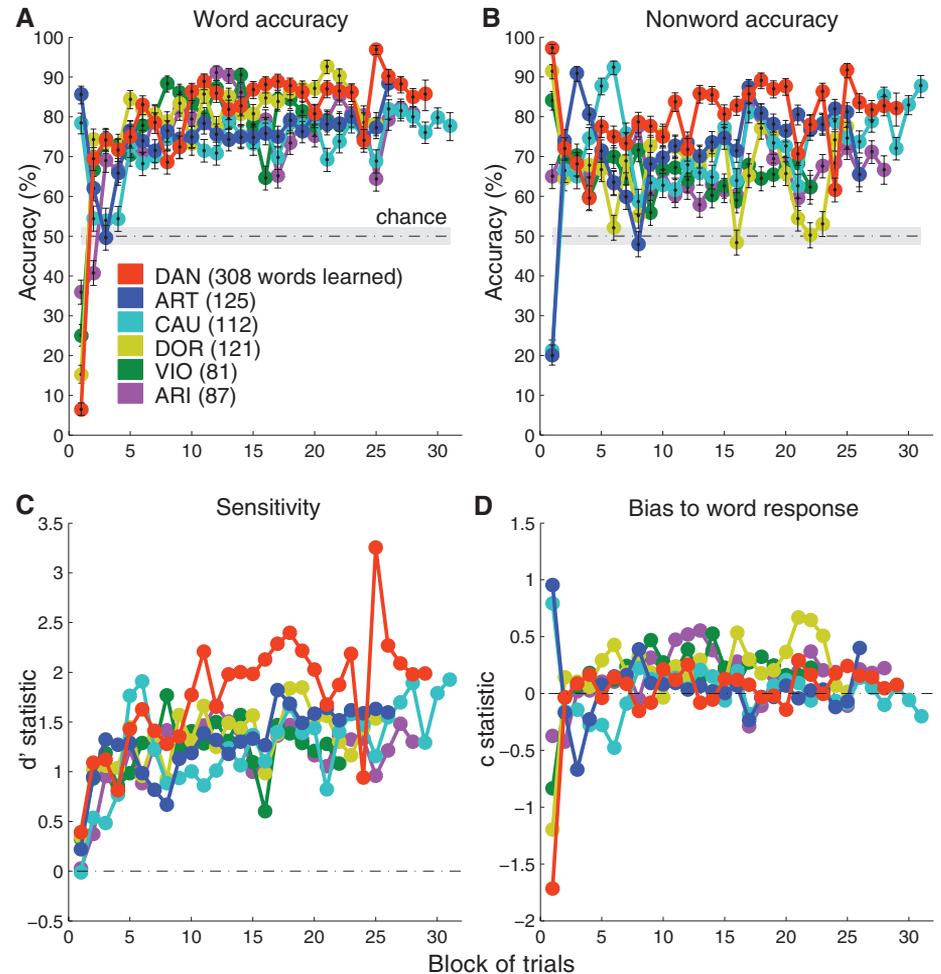
identification processes in ventral occipitotemporal brain areas to the specific characteristics of printed words (8, 9).

However, according to the dominant theories of reading, orthographic processing is still primarily considered to be an extension of already established linguistic skills in the domain of spoken language processing (or sign language processing in hearing-impaired persons). Indeed, the task of learning to read a language with an alphabetic script is facilitated by the fact that individual letters and letter clusters (graphemes) can be associated with the elementary sounds of the language (phonemes) in order to recover a phonological representation of the word being read, and from there to understand its meaning (5). Such phonological recoding operations could well be a major constraint that forces the beginning reader to process individual letters rather than the word as a whole (10). Thus, it has typically been argued that orthographic processing is a predominantly linguistic skill, requiring the same cerebral predisposition as thought to be required for spoken and sign language processing, or at the least, prior exposure to the language in question.

We challenged the hypothesis that learning an orthographic code depends on preexisting linguistic knowledge by investigating whether nonhuman primates can learn this skill. Humans and nonhuman primates from the cercopithecidae family, such as macaques or baboons, have similar visual systems (11). However, the communicative system of cercopithecidae arguably lacks the structural complexity of human language (12) and certainly does not include any phonological representation of English words that could be associated with the printed forms of these words. Using a new testing procedure in which socially housed monkeys had free access to computer-controlled operant conditioning setups with touch screens (Fig. 1B) (13), we trained six baboons to discriminate randomly selected real English words four letters in length from artificially generated four-letter strings of letters that were not real English words [henceforth, nonwords (Fig. 1C)]. All nonwords were formed of a vowel and three consonants and contained letter combinations (bigrams) that occurred in real words. Bigram frequency was minimized in the list of nonwords and maximized in the list of words (14), so that the word versus nonword discrimination could be made implicitly on the basis of statistical dependencies between letters. Words and nonwords were presented randomly in blocks of 100 trials. The 100-trial sessions were composed of 25 presentations of a novel word to learn, 25 presentations of words randomly selected from already learned words, and 50 nonword trials. Each new word was added to the ever-increasing pool of already learned words, once responses to that word exceeded 80% correct within the preceding session. Thus, in terms of explicit information available to the baboons, a word was defined as a string of letters that was repeatedly presented, whereas a nonword was rarely repeated. The baboons responded by touching one of two shapes shown on the touch screen and were given a food reward after a correct response (Fig. 1C) (see the supplementary materials for more details).

Over a period of a month and a half, baboons learned to discriminate dozens of words (the counts ranged from 81 words for baboon VIO to 308 words for baboon DAN) from among a total of 7832 nonwords at nearly 75% accuracy (Fig. 2 and table S1). This in itself is a remarkable result, given the level of orthographic similarity between the word and nonword stimuli. More detailed analyses revealed that baboons were not simply memorizing the word stimuli but had learned to discriminate words from nonwords on the basis of differences in the frequency of letter combinations in the two categories of stimuli
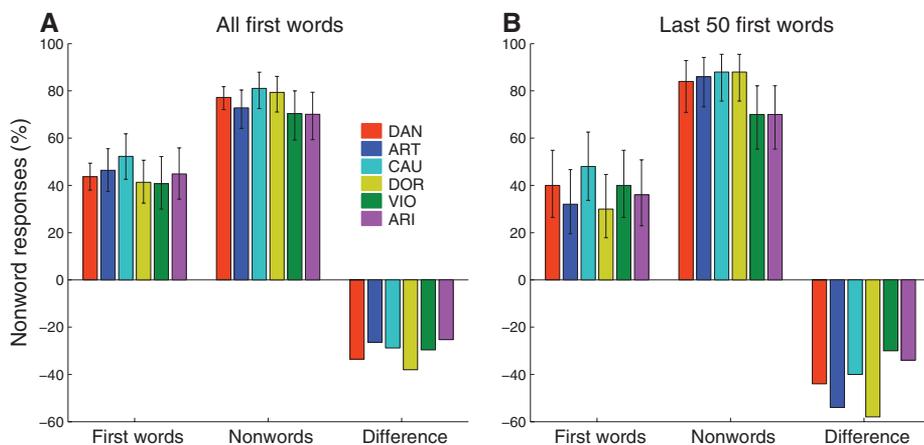
Fig. 3. Percentage of nonword responses on trials corresponding to words seen for the first time as compared to the first nonword stimuli after these particular trials. Performance on trials corresponding to the first presentation of words is of particular interest, because any divergence from performance to nonword stimuli is an indication that the baboons have learned general statistical properties of the two classes of stimuli. All six baboons showed such a divergence for both the total number of first word trials (**A**) and the last 50 first word trials (**B**), as revealed in the differences in the percentage of nonword responses to first words and nonwords (all P values < 0.01).
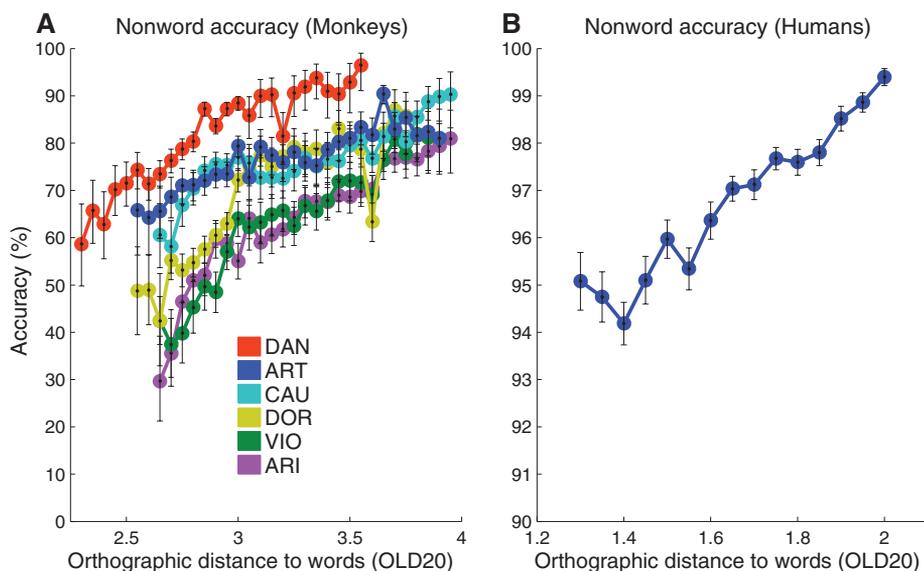
Fig. 4. Performance in response to nonwords depends on their orthographic similarity to learned words for both monkeys and humans. (**A**) For each of the last 20,000 nonword trials, the orthographic Levenshtein distance (OLD20) (15, 16) was computed between the corresponding nonword and each of the words learned at that time, separately for each baboon. The average accuracy corresponding to each unique value of OLD20 was then calculated. The graph shows that baboons responded less accurately to more wordlike nonwords (smaller OLD20 values). Errors bars correspond to the 95% binomial confidence interval. (**B**) For comparison, humans show a similar sensitivity to orthographic distance to known words when responding to nonwords (see supplementary text).

(i.e., statistical learning). Indeed, there was a significant correlation between mean bigram frequency and word accuracy [correlation coefficients (r) ranged from 0.51 for baboon VIO to 0.80 for baboon DAN, all P values < 0.05; see supplementary materials]. More importantly, words that were seen for the first time triggered significantly fewer "nonword" responses than did the nonword stimuli (Fig. 3). This implies that the baboons had extracted knowledge about what statistical properties characterize words and nonwords and used this information to make their word versus nonword decision without having seen the specific examples before. In the absence of such knowledge, words seen for the first time should have been processed like nonwords. Figure 3 shows that this was clearly not the case.

Even more striking is the strong linear relation, shown in Fig. 4, between accuracy in response to nonword stimuli and their orthographic similarity to words that the baboons had already learned. The more similar a nonword was to a known word, the more false positive responses it produced. Orthographic similarity was measured with a standard edit distance used in information theory and computer science (*15*). For each nonword we counted the number of letter insertions, letter deletions, and letter substitutions required to transform the nonword into a known word, and we retained the average of the 20 lowest values as the OLD20 value of that nonword (*16*). Thus, the smaller the OLD20 value, the greater the orthographic similarity between the nonword and the set of known words. This standard measure of orthographic similarity was found to have a quasi-linear relationship with the accuracy of responses to nonwords [explained variance (R^2) ranged from 0.76 for baboon DOR to 0.91 for baboon VIO; see supplementary materials]. Exactly the same quasi-linear relationship was found in an analysis of the accuracy of human responses to nonwords in a large-scale lexical decision experiment (*17*). This finding implies that the baboons were sensitive to the orthographic characteristics of word and nonword stimuli in a way that mimics the sensitivity to orthographic similarity seen in skilled human readers.

Our results indicate that baboons were coding the word and nonword stimuli as a set of letter identities arranged in a particular order. Baboons had learned to discriminate different letters from each other (letter identity) and to associate those letter identities with positional information. Their coding of the statistical dependencies between position-coded letters is reflected in (i) their ability to discriminate novel words from nonwords (i.e., generalization), (ii) the significant correlation between bigram frequency and the accuracy of responses to words, and (iii) the increase in errors in response to nonword stimuli that were orthographically more similar to known words. Thus, our results support the conclusion that the baboons were computing an orthographic code in order to accurately discriminate words from nonwords. Prior linguistic knowledge is therefore not a necessary prerequisite in order to achieve humanlike orthographic processing.

Our findings have two important theoretical implications. First, they suggest that statistical learning is a powerful universal (i.e., cross-species) mechanism that might well be the basis for learning higher-order (linguistic) categories that facilitate the evolution of natural language (*18*, *19*). Second, our results suggest that orthographic processing may, at least partly, be constrained by general principles of visual object processing shared by monkeys and humans. One such principle most likely concerns the use of feature combinations to identify visual objects (*20*), which would be analogous to the use of letter combinations in recent accounts of orthographic processing (*4*, *9*, *21*). Given the evidence that baboons process individual features or their combinations in order to discriminate visual objects (*22*), we suggest that similar mechanisms were used to distinguish words from nonwords in the current study. Our study may therefore help explain the success of the human cultural choice of visually representing words using combinations of aligned, spatially compact, ordered sequences of symbols. The primate brain might therefore be better prepared than previously thought to process printed words, hence facilitating the initial steps toward mastering one of the most complex of human skills: reading.

References and Notes

1. J. L. McClelland, D. E. Rumelhart, *Psychol. Rev.* **88**, 375 (1981).
2. S. Dehaene, *Reading in the Brain* (Penguin Viking, New York, 2009).
3. J. Grainger, A. M. Jacobs, *Psychol. Rev.* **103**, 518 (1996).
4. J. Grainger, J. C. Ziegler, *Front. Psychology* **2**, 54 (2011).
5. J. C. Ziegler, U. C. Goswami, *Psychol. Bull.* **131**, 3 (2005).
6. L. Cohen *et al.*, *Brain* **123**, 291 (2000).
7. S. Dehaene, L. Cohen, *Trends Cogn. Sci.* **15**, 254 (2011).
8. S. Dehaene, L. Cohen, *Neuron* **56**, 384 (2007).
9. S. Dehaene, L. Cohen, M. Sigman, F. Vinckier, *Trends Cogn. Sci.* **9**, 335 (2005).
10. K. Nation, *Q. J. Exp. Psychol.* **61**, 1121 (2008).
11. J. L. Fobes, L. E. King, in *Primate Behavior*, J. L. Fobes, J. E. King, Eds. (Academic Press, New York, 1982), pp. 219–243.
12. R. M. Seyfarth, D. L. Cheney, *Brain Lang.* **115**, 92 (2010).
13. J. Fagot, E. Bonté, *Behav. Res. Methods* **42**, 507 (2010).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. V. I. Levenshtein, *Sov. Phys. Dokl.* **10**, 707 (1966).
16. T. Yarkoni, D. A. Balota, M. J. Yap, *Psychon. Bull. Rev.* **15**, 971 (2008).
17. E. Keuleers, P. Lacey, K. Rastle, M. Brysbaert, *Behav. Res. Methods* **44**, 287 (2012).
18. F. Ramus, M. D. Hauser, C. Miller, D. Morris, J. Mehler, *Science* **288**, 349 (2000).
19. J. R. Saffran, R. N. Aslin, E. L. Newport, *Science* **274**, 1926 (1996).
20. S. L. Brincat, C. E. Connor, *Neuron* **49**, 17 (2006).
21. C. Whitney, *Psychon. Bull. Rev.* **8**, 221 (2001).
22. C. Parron, J. Fagot, *Learn. Behav.* **38**, 374 (2010).

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6078/245/DC1
Materials and Methods
Supplementary Text
Tables S1 and S2
Reference (*23*)

Supplementary Materials for

Orthographic Processing in Baboons (*Papio papio*)

Jonathan Grainger,∗ Stéphane Dufau, Marie Montant, Johannes C. Ziegler, Joël Fagot

*To whom correspondence should be addressed. E-mail: i.jonathan.grainger@gmail.com

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/336/6078/245/DC1)

Supporting online material

I. Materials and Methods

Data collection. The study was conducted in a world unique facility developed by the last author (*13*). The key feature of this testing facility is the fact that baboons have completely free access to the computerized testing booths that are installed in trailers next to their enclosure. The baboons are identified by microchips implanted in their forearms whenever they enter one of the booths, and can thus participate in a given experiment whenever they want, 24 hours a day, 7 days a week. This procedure has numerous advantages over more traditional testing procedures. One advantage is that we do not have to hassle the animals with captures. Second, because the baboons freely participate in the research, they pay much more attention to the tasks. All of our previous studies together with the present study have shown that we can obtain excellent learning performance in these conditions, in fact, much higher than what is commonly reported in more traditional laboratory setups. Finally, the baboons regularly perform between 1000 and 3000 trials/day. This large number of trials explains the sample size of our study (about 50,000 trials/baboon, 300,000 trials altogether for the 6 baboons over one month and half), as well as the performance levels achieved in the task.

Stimuli. A set of 2,235 English four-letter words and their printed frequencies was extracted from the CELEX word-frequency corpus (*23*). Bigram frequencies were calculated for each of the three contiguous bigrams of a word (letters 1&2, letters 2&3, letters 3&4) counting the number of times these bigrams appeared in the corpus at the same position. Mean bigram frequency was calculated for each word by averaging the three positional bigram frequencies. The 500 words with the highest mean bigram frequency were selected as "word" stimuli. A set of 10,091 four-letter nonwords was created using all the bigrams that appeared at a particular position (initial, medial, terminal) in the CELEX English four-letter word list. All nonwords were composed of one vowel letter and three consonants, and the vowel could be at any position. Each nonword was associated with a mean bigram frequency that was calculated using position-specific bigram frequencies as with the word stimuli. We then selected 7,832 nonwords that had a mean bigram frequency that was less than the lowest mean bigram frequency of the word stimuli (mean bigram frequency for words, 3.60×10^{-4}; mean bigram frequency for nonwords, 5.96×10^{-5}). When performing frequency calculations based on the words and nonwords selected for the present experiment, word and nonword stimuli differed significantly in terms of average bigram frequency ($P<0.001$, two-sample Kolmogorov-Smirnov test, n_words = 500, n_nonwords = 7,832), but not in terms of average positional letter frequency ($P>0.1$, two-sample Kolmogorov-Smirnov test, n_words = 500, n_nonwords = 7,832). The word and nonword stimuli were presented on a computer screen using a set of 50x50 pixel images of the 26 uppercase letters of the Roman alphabet. Letters were monospaced, sans-serif, and printed in yellow on a black background. They were 1.8 cm wide and 2 cm high, and separated by 2 pixels. Viewing distance was approximately 30 cm.

Training procedure. Baboons had microchips implanted in their forearms, and a microchip reader detected their identity as soon as one of their hands approached the touch screen of the experimental apparatus. This action triggered a trial that began with the presentation of a stimulus (word or nonword) until the baboon touched the screen. The response screen was then displayed figuring a dark blue cross on the left (correct nonword response) and a light blue oval shape on the right (correct word response) – see Figure 1C. As soon as the baboon touched one of the response zones, the response screen disappeared and was followed by a food reward (dry wheat) when a correct word or nonword response had been given, or followed by a 3 sec delay with a green screen if the response was incorrect. The intertrial interval was minimally 3 sec, but could be longer since the baboons chose when to initiate a trial. Words and nonwords were presented randomly in blocks of 100 trials. The 100 trial sessions were composed of 25 presentations of a novel word to learn, 25 presentations of words randomly selected from already learned words, and 50 nonword trials. Each new word was added to the ever increasing pool of already learned words once response to that word exceeded 80% correct within a session.

Participants. The data obtained from six guinea baboons (*Papio papio*; 3 females) are reported. Prior to this study, baboons had extensive experience with the experimental apparatus used here, but no previous exposure to written materials (words or letters).

II. Supporting text and analysis

Additional information for Table 1. Chi-square values for all of the trials (DAN, 72.25, $P<0.001$; ART, 18.09, $P<0.001$; CAU, 20.76, $P<0.001$; DOR, 36.53, $P<0.001$; VIO, 14.40, $P<0.001$; ARI, 11.38, $P<0.001$). Chi-square values for the last 50 trials (DAN, 20.54, $P<0.001$; ART, 30.14, $P<0.001$; CAU, 18.38, $P<0.001$; DOR, 34.77, $P<0.001$; VIO, 9.09, $P<0.01$; ARI, 11.60, $P<0.001$).

Additional information for Figure 2, panels A and B. Number of stimuli in the last block: DAN, 689, ART, 985, CAU, 1,142, DOR, 1608, VIO, 1,041, ARI, 1,407. **Panels C and D**. For each baboon after the first block of 2,000 trials, d' statistics were all above zero indicating successful word/nonword discrimination (DAN, $t(27)=18.51$, $P<0.001$; ART, $t(24)=23.52$, $P<0.001$; CAU, $t(29)=17.66$, $P<0.001$; DOR, $t(23)=25.04$, $P<0.001$; VIO, $t(20)=21.96$, $P<0.001$; ARI, $t(26)=23.58$, $P<0.001$). Four baboons showed a bias to a "word" response (c statistics above zero; DAN, $t(27)=3.16$, $P<0.05$; DOR, $t(23)=6.45$, $P<0.001$; VIO, $t(20)=7.33$, $P<0.001$; ARI, $t(26)=3.70$, $P<0.05$) while two baboons showed no significant bias (ART, $t(24)=0.15$, $P>0.1$; CAU, $t(29)=-0.55$, $P>0.1$).

Additional information for Figure 4, panel A. For all unique OLD20 values, mean accuracies computed with less than 100 trials were excluded from the analysis. For each baboon, OLD20 data points were regressed with both linear and quadratic models. Accuracy of nonword responses in baboons was strongly related to the orthographic distance to known words (R^2 for linear model, DAN: 0.88, ART: 0.80, CAU: 0.79, DOR: 0.76, VIO: 0.91, ARI: 0.87; R^2 for quadratic model, DAN: 0.93, ART: 0.84, CAU: 0.79, DOR: 0.85, VIO: 0.93, ARI: 0.92). **Panel B**. 1,921 4-letter nonwords derived from English words were selected from the British lexicon project, which provides human lexical decision responses to a large selection of English words and nonwords (18). Items with low accuracy (less than the mean accuracy minus two standard deviations) were excluded from further analysis, leaving 1,849 nonwords. We applied a similar analysis to the one performed for the baboon data. We first calculated the orthographic distance (OLD20) for each of these nonwords, and then calculated mean accuracy for the nonwords at each of the unique values of OLD20 (we excluded groups with less than 50 items; error bars are the standard error). Human OLD20 nonword datapoints were regressed with both linear and quadratic models, showing that nonword accuracy was related to orthographic distance to words (R^2 for linear model, 0.94; R^2 for quadratic model, 0.94).

Additional information for the correlation between bigram frequency and word performance. For each baboon and for each word starting from trial #20000, we computed the frequency of letter pairs (L1L2, L2L3, L3L4) occurring in the pool of words each baboon had already learned. We took the mean of the three frequency values. Each word was therefore associated with a mean bigram frequency. For each monkey, we grouped word trials corresponding to frequency bands from 0 to 1 with a step of 0.01. Groups with less than 100 words were discarded from the analysis. We computed the mean accuracy of each group, and correlated these accuracies with the mean values of the frequency bins. The correlation values were: DAN, $r=0.55$, $P<0.05$; ART, $r=0.77$, $P<0.001$; CAU, $r=0.73$, $P<0.001$; DOR, $r=0.79$, $P<0.001$; VIO, $r=0.51$, $P<0.05$; ARI, $r=0.63$, $P<0.001$.

III. Supporting tables

Table S1. Experimental profiles of the six baboons.

	DAN	ART	CAU	DOR	VIO	ARI
Number of trials	56,689	50,985	61,142	49,608	43,041	55,407
Number of words learned	308	125	112	121	81	87
Number of nonwords presented	7,832	7,832	7,832	7,832	7,832	7,832
General accuracy (%)	79.81	73.41	72.43	73.15	71.55	71.14
Word accuracy (%)	80.01	74.83	73.15	79.26	76.75	75.38
Nonword accuracy (%)	79.61	72.00	71.72	67.06	66.33	66.90

References and Notes

1. J. L. McClelland, D. E. Rumelhart, An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychol. Rev.* **88**, 375 (1981). doi:10.1037/0033-295X.88.5.375

2. S. Dehaene, *Reading in the Brain* (Penguin Viking, New York, 2009).

3. J. Grainger, A. M. Jacobs, Orthographic processing in visual word recognition: A multiple read-out model. *Psychol. Rev.* **103**, 518 (1996). doi:10.1037/0033-295X.103.3.518 Medline

4. J. Grainger, J. C. Ziegler, A dual-route approach to orthographic processing. *Front. Psychology* **2**, 54 (2011). doi:10.3389/fpsyg.2011.00054

5. J. C. Ziegler, U. C. Goswami, Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychol. Bull.* **131**, 3 (2005). doi:10.1037/0033-2909.131.1.3 Medline

6. L. Cohen *et al*., The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* **123**, 291 (2000). doi:10.1093/brain/123.2.291 Medline

7. S. Dehaene, L. Cohen, The unique role of the visual word form area in reading. *Trends Cogn. Sci.* **15**, 254 (2011). doi:10.1016/j.tics.2011.04.003 Medline

8. S. Dehaene, L. Cohen, Cultural recycling of cortical maps. *Neuron* **56**, 384 (2007). doi:10.1016/j.neuron.2007.10.004 Medline

9. S. Dehaene, L. Cohen, M. Sigman, F. Vinckier, The neural code for written words: A proposal. *Trends Cogn. Sci.* **9**, 335 (2005). doi:10.1016/j.tics.2005.05.004 Medline

10. K. Nation, Learning to read words. *Q. J. Exp. Psychol.* **61**, 1121 (2008). doi:10.1080/17470210802034603

11. J. L. Fobes, L. E. King, in *Primate Behavior*, J. L. Fobes, J. E. King, Eds. (Academic Press, New York, 1982), pp. 219–243.

12. R. M. Seyfarth, D. L. Cheney, Production, usage, and comprehension in animal vocalizations. *Brain Lang.* **115**, 92 (2010). doi:10.1016/j.bandl.2009.10.003 Medline

13. J. Fagot, E. Bonté, Automated testing of cognitive performance in monkeys: use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). *Behav. Res. Methods* **42**, 507 (2010). doi:10.3758/BRM.42.2.507 Medline

14. Materials and methods are available as supplementary material on *Science* Online.

15. V. I. Levenshtein, *Sov. Phys. Dokl.* **10**, 707 (1966).

16. T. Yarkoni, D. A. Balota, M. J. Yap, Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychon. Bull. Rev.* **15**, 971 (2008). doi:10.3758/PBR.15.5.971 Medline

17. E. Keuleers, P. Lacey, K. Rastle, M. Brysbaert, The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behav. Res. Methods* **44**, 287 (2012). doi:10.3758/s13428-011-0118-4 Medline

18. F. Ramus, M. D. Hauser, C. Miller, D. Morris, J. Mehler, Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* **288**, 349 (2000). doi:10.1126/science.288.5464.349 Medline

19. J. R. Saffran, R. N. Aslin, E. L. Newport, Statistical learning by 8-month-old infants. *Science* **274**, 1926 (1996). doi:10.1126/science.274.5294.1926 Medline

20. S. L. Brincat, C. E. Connor, Dynamic shape synthesis in posterior inferotemporal cortex. *Neuron* **49**, 17 (2006). doi:10.1016/j.neuron.2005.11.026 Medline

21. C. Whitney, How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychon. Bull. Rev.* **8**, 221 (2001). doi:10.3758/BF03196158 Medline

22. C. Parron, J. Fagot, First- and second-order configural sensitivity for greeble stimuli in baboons. *Learn. Behav.* **38**, 374 (2010). doi:10.3758/LB.38.4.374 Medline

23. R. H. Baayen, R. Piepenbrock, H. van Rijn, *The CELEX Lexical Database* (CD-ROM, Linguistic Data Consortium, Univ. of Pennsylvania, Philadelphia, PA, 1993).