includes gravity, there are other interesting consequences of this line of argument. One of the most striking is the following. In QCD, when the temperature reaches sufficiently high values (above 10^{12} K), a phase transition occurs and quarks and gluons are no longer confined — instead, a 'soup' of free particles is formed, called quark–gluon plasma. In the five-dimensional theory, this transition also corresponds to the formation of a black hole in the interior. Our knowledge of black holes can then tell us something about quark–gluon plasma. In addition, the QCD–string theory provides a simple explanation for an interesting feature of black holes — the Bekenstein–Hawking entropy. This entropy, a measure of the number of possible quantum microstates, arises from the thermodynamic properties of a black hole (which are also at the root of Hawking radiation). Counting these microstates to work out the entropy has proved a major challenge in theories of quantum gravity. But in the five-dimensional theory, the black-hole entropy becomes just the entropy of the plasma of quarks and gluons.

There is an intimate connection between the physics of strong interactions and both string theory and quantum gravity. Hopefully, in the next few years a string-theory description for real-world QCD will emerge, making it possible to perform computations in a relatively simple way. And perhaps, beyond that, we might even arrive at a QCD-like theory that can describe gravity. ■

Juan Maldacena is in the School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA.
e-mail: malda@ias.edu

1. Polchinski, J. & Strassler, M. *J. High Energy Phys.* 05(2003)012.
2. 't Hooft, G. *Nucl. Phys. B* **72,** 461–473 (1974).
3. Maldacena, J. *Adv. Theor. Math. Phys.* **2,** 231–252 (1998).
4. Gubser, S., Klebanov, I. & Polyakov, A. *Phys. Lett. B* **428,** 105–114 (1998).
5. Witten, E. *Adv. Theor. Math. Phys.* **2,** 253–291 (1998).

Cognitive neuroscience

Practice doesn't make perfect

Wilson Geisler and Richard Murray

It may seem counterintuitive, but we are not very efficient at recognizing even the most common words. This finding suggests strict limits on how flexible we are in learning to recognize new patterns.

If there is any truth to the adage that 'practice makes perfect', we should certainly be near perfect at recognizing common words. As young children, we are taught first to recognize letters, and shortly thereafter words, and then for the rest of our lives we practise word recognition every day — the average literate person has read, at a conservative estimate, a hundred million words by the age of 25, thereby practising the recognition of each common word many hundreds of thousands of times. Nonetheless, on page 752 of this issue, Pelli, Farell and Moore[1] show that word recognition is surprisingly inefficient relative to letter recognition. This finding, which is surely not unique to words, implies that the neural learning mechanisms that are involved in pattern and object recognition are severely limited in their capabilities.

To measure the efficiency of word recognition, Pelli *et al.*[1] used the powerful tools and logic of ideal-observer analysis[2–6]. At the heart of this method is the concept of the 'ideal observer' — a theoretical device that achieves the best-possible performance in a perceptual or cognitive task, given the available information and constraints. Perceptual and cognitive tasks generally involve processing noisy physical or neural signals under conditions of uncertainty, and so ideal observers are typically derived using the concepts and methods of Bayesian statistics.

Pelli and colleagues' study elegantly demonstrates the value of this type of analysis. First, derivation of the ideal observer forces one to rigorously specify the task at hand and to determine at least one specific mechanism for achieving optimal performance. Here, the task was to identify words or letters embedded in spatial white noise (top of Fig. 1). In each trial, a word was randomly picked from a set of possible words, given a particular contrast, and then added to a random sample of noise; the contrast determined how well the word stood out against the noisy background.

One type of mechanism that achieves optimal word recognition in this task is the 'template matcher' (Fig. 1, left path). This mechanism stores an exact copy (a template) of each target (such as a word) that could appear in a stimulus. To identify a particular target, the template matcher calculates how well each point on the target corresponds with each point on each stored template, and identifies the target as the template that gives the highest correlation[7]. (If the various targets do not occur with equal probability, then each correlation is adjusted according to the target's probability.) Many neurons that are found in the first stages of the visual pathway can be modelled, to a first approximation, as template matchers.

The second benefit of ideal-observer analysis is that 'ideal' performance provides

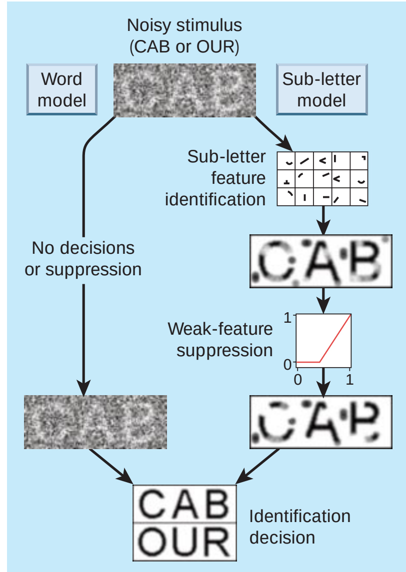

Figure 1 **Why are we so inefficient at recognizing words? An efficient word-recognition mechanism, such as a template matcher (left pathway), might integrate information over the entire stimulus. As more and more letters are added to a word, the template matcher can identify the word at lower and lower contrasts. On the other hand, a mechanism that suppresses weak neural responses to letters or parts of letters (sub-letter features; right pathway) sets a lower limit on the contrast at which words can be identified: below the suppression threshold, the stimulus is simply not visible. This threshold is represented in the graph by the point at which the curve relating input (*x*-axis) to output (*y*-axis) begins to increase. Consequently, the addition of more and more letters does not mean that words can be identified at lower and lower contrasts. Pelli *et al.*[1] have found that our ability to recognize words decreases rapidly as word length increases, implying that the brain does not work as a 'template matcher'. They argue that instead it functions more like the right pathway.**

an appropriate benchmark against which to compare human performance. In Pelli and colleagues' work[1], it makes it possible to rigorously answer the question of whether practice makes perfect, by providing a precise description of 'perfect' performance. The performance of a human observer relative to the ideal is called the efficiency. In this case, the efficiency equals the squared contrast of the target at which it can be identified by the ideal observer with some given level of accuracy (for instance, 70% correct), divided by the squared contrast required for the human observer to identify it with the same level of accuracy. Pelli and colleagues' main finding is that human observers' efficiency at recognizing words declines drastically as word length increases. Specifically, if a person's efficiency at recognizing single letters is *F*, then their efficiency at recognizing *n*-letter words is *F*/*n*. So, for example, common five-letter words are identified with only 1/5 the efficiency of single letters. This relatively poor performance holds for even the most common three-letter words, which often occur many times in a paragraph.

Why are people so inefficient at recognizing even common words? The ideal observer can also be helpful in addressing such questions, by providing a starting point for formulating possible models of human (sub-optimal) performance. We know that, with a little practice, humans can become extremely efficient at recognizing some simple patterns — such as spatially localized bars, edges and spots (see, for example, ref. 4). This is probably possible because the visual cortex in the primate brain contains neurons that act like template matchers for these simple patterns. But the 1/*n* efficiency relationship reported by Pelli *et al.* implies that the brain cannot learn optimal templates for whole words. Rather, their results are quantitatively consistent with a model in which the brain has efficient templates only for parts of words — letters, or even features of letters — and in which only the strong responses of the neurons encoding these parts are passed on to word-identification processes (Fig. 1, right path). The hypothesis that familiar complex objects are initially encoded in smaller parts is consistent with recent psychophysical studies of the identification of objects in noise[8].

Although Pelli and colleagues' results[1] show that word recognition is relatively inefficient, they do not rule out the possibility of specific neural circuits for the recognition of common words or other familiar objects. For example, there could be individual neurons whose receptive fields are precisely tuned to such objects. What the results do imply is that any neural circuits for recognizing complex objects such as words (or letters) cannot integrate neuronal responses as efficiently as can, say, the circuits underlying the receptive fields of neurons in the primary visual cortex (an early stage of the visual pathway), which encode simple patterns such as bars, edges and spots.

Pelli and colleagues suggest that inefficient integration occurs because weak responses to features (or letters) are suppressed ('squelched'). Moreover, they suggest that the brain makes discrete (categorical) decisions about which features (or letters) are present before deciding which word is present. Although this is plausible, there are alternative explanations that are similar in spirit but differ in detail. For example, it might not be necessary to invoke the suppression of weak responses to features. The 1/*n* efficiency relationship could also be a result of increasing noise (neural noise, for instance) or uncertainty at lower target contrasts. The essential thing is that the signal-to-noise ratios of the neuronal responses that are passed on to the word-identification process degrade rapidly as the target's contrast is lowered. Similarly, discrete decisions about features or letters are not required; simulations that we have carried out suggest that the predictions are almost the same if graded responses are used in word identification, as long as the graded responses deteriorate at low contrasts.

Regardless of the specific explanation, Pelli and colleagues' results[1] clearly show that, even with extensive practice over many years, perceptual learning mechanisms are unable to achieve efficient recognition of arbitrary objects. This finding is particularly revealing given that the visual system is very efficient at recognizing certain simple patterns such as bars and edges. It seems that Pelli *et al.* have uncovered, with the help of ideal-observer analysis, a fundamental limit on the neural mechanisms of learning and plasticity. ■

Wilson Geisler and Richard Murray are at the Center for Perceptual Systems and the Department of Psychology, University of Texas, Austin, Texas 78712, USA.
e-mail: geisler@psy.utexas.edu

1. Pelli, D. G., Farell, B. & Moore, D. C. *Nature* 423, 752–756 (2003).
2. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
3. Barlow, H. B. *Vision Res.* 18, 637–650 (1978).
4. Burgess, A. E., Wagner, R. F., Jennings, R. J. & Barlow, H. B. *Science* 214, 93–94 (1981).
5. Geisler, W. S. *Psychol. Rev.* 96, 267–314 (1989).
6. Geisler, W. S. & Kersten, D. *Nature Neurosci.* **5,** 508–510 (2002).
7. Peterson, W. W., Birdsall, T. G. & Fox, W. C. *Trans. Inst. Radio Engineers Prof. Group Information Theory* 4, 171–212 (1954).
8. Gold, J. M., Murray, R. F., Bennett, P. J. & Sekuler, A. B. *Curr. Biol.* 10, 663–666 (2000).

Astrophysics

Superficial resonance

Frits Paerels

Neutron stars are the most poorly understood stellar objects in the Universe. But observations of X-rays emitted from one neutron star have now revealed a clue to the nature of its surface and composition.

Neutron stars are the remnants of massive stars. The density of their cores exceeds that of atomic nuclei, and they might contain exotic states of matter that are not found anywhere else in nature[1]. For decades astronomers have tried to 'see' the surfaces of such stars, in an effort to determine their physical properties. Space-borne telescopes have brought improved X-ray-imaging capabilities, but the first observations of neutron stars with these instruments were rather disappointing, revealing that the X-ray emission from most isolated neutron stars is remarkably featureless. But now observations of one neutron star have finally given us a glimpse of the fundamental properties of these objects: on page 725 of this issue, Bignami *et al.*[2] describe the first physical clue to the nature of this star's X-ray emission.

When massive stars — stars more than eight times heavier than the Sun — run out of thermonuclear fuel, their cores collapse. The resulting configuration typically has a mass about one-and-a half times that of the Sun, but a radius of only 10 km: the density of the material exceeds the nuclear density, and degeneracy (the quantum-mechanical restriction on the occupation of energy levels by particles) provides the support against the gravity of the star's own weight. According to simple astrophysical arguments, a young neutron star is expected to have a magnetic field of the order of 10^{12} gauss at its surface (in comparison, Earth's field is about 0.5 gauss). It should also spin rapidly, appearing as a 'pulsar'.

The first neutron stars were discovered through their emission at radio wavelengths. Radio emission is produced in the envelope of magnetic field that surrounds a pulsar (through processes that are still poorly understood), and although it carries information about the star's magnetic field and spin period, it tells us nothing about the internal properties of the neutron star itself. If, instead, we could observe emission from the star's surface, it would tell us about the surface composition, which in turn would reveal more about how a neutron star forms, and the processes that affect the chemical composition of its atmosphere.

To 'view' the surface of hot young neutron stars requires observations of their X-ray, rather than radio, emission: the emission peaks in the 'soft' X-ray band of photon energies (0.1–1 kiloelectronvolts, or keV). But even in the X-ray band, isolated hot neutron stars are faint. The 1999 launches of NASA's Chandra X-ray Observatory and the European Space Agency's XMM-Newton observatory were eagerly awaited, but the initial observations of a handful of hot neutron stars were universally disappointing. The spectra of radiation from the stars did show the expected characteristic of thermal emission from their surfaces: a gently curving, black-body-like energy distribution. But detailed scrutiny of the X-ray spectrum failed to reveal any structure; the shape of the spectrum in all cases was very close to that of a black body, which by its very nature