Recipe for pattern formation? Schematic diagram of a two-dimensional electron gas in a perpendicular magnetic field (green field lines) subjected to microwave radiation (red). Shown is an example of the type of current pattern that might be associated with the observed zero-resistance state. Domain walls (purple) separate regions of large but counterflowing current density. Net current flow is equal to the applied current, depicted left to right.

terials could even be fabricated.] The key is to consider processes in which electrons in the 2DEG absorb energy from the microwaves and are also scattered by impurities or lattice vibrations. The effect of the microwaves is a type of population inversion (*11*, *12*) similar to that which occurs in a laser. For the microwave frequencies and magnetic fields at which the experiments reveal resistance decreasing until it saturates at zero, these calculations show resistance decreasing, passing through zero, and then becoming negative. Thus, this work explained how the microwaves could cause a decrease in resistance, but left a key question unanswered: Why does the resistance saturate at zero?

Missing from the calculations discussed above is the effect of interactions between the electrons. Such interactions would be exceedingly difficult to include within a

microscopic calculation, but can be considered at the macroscopic level through the solution of Maxwell's equations of electrodynamics. Doing so for the 2DEG system, it was quickly shown that a homogeneous current distribution with negative resistance would be electrodynamically unstable (*13*). To avoid the negative-resistance state, the system can spontaneously arrange itself into an intricate current pattern (see the seccond figure) characterized by domain walls separating regions of large local current density at which the resistance is not negative, but zero.

This is an exciting possibility which, together with the microscopic calculations showing negative resistance, provides a plausible explanation for the phenomenon seen in the experiments. Furthermore, it suggests that the microwave-induced zero-resistance states should be understood in terms of the nonequilibrium physics of pattern formation. However, much work remains to be done before this scenario can be accepted as reality. On the theory side, it

is important to develop a better understanding of what kinds of patterns are optimal, how they are formed, and how their fluctuations can be characterized. Experiments may be able to detect the current patterns, if they do exist, by measuring the local Hall voltages they create, the magnetic field patterns they produce, or the charges built up on the current domain walls. Although experimental work is currently in progress, these measurements are very difficult to perform and we are still waiting for a smoking gun.

References and Notes
1. M. A. Zudov *et al.*, *Phys. Rev. B* **64**, 201311 (2001).
2. R. G. Mani *et al.*, *Nature* **420**, 646 (2002).
3. M. A. Zudov *et al.*, *Phys. Rev. Lett.* **90**, 46807 (2003).
4. Contributions have been made by too many authors to list here. See references in (*8*, *14*).
5. B. J. Keay *et al.*, *Phys. Rev. Lett.* **75**, 4102 (1995).
6. A. C. Durst *et al.*, *Phys. Rev. Lett.* **91**, 86803 (2003).
7. P. W. Anderson, W. F. Brinkman, www.arxiv.org/abs/cond-mat/0302129.
8. M. G. Vavilov, I. L. Aleiner, *Phys. Rev. B* **69**, 35303 (2004).
9. V. I. Ryzhii, *Fiz. Tverd. Tela* **11**, 2577 (1969) [*Sov. Phys. Solid State* **11**, 2078 (1970)].
10. V. I. Ryzhii *et al.*, *Fiz. Tekh. Poluprovodn.* **20**, 2078 (1986) [*Sov. Phys. Semicond.* **20**, 1289 (1986)].
11. S. I. Dorozhkin, *Pis'ma Zh. Eksp. Teor. Fiz.* **77**, 681 (2003) [*JETP Lett.* **77**, 577 (2003)].
12. I. A. Dmitriev *et al.*, www.arxiv.org/abs/cond-mat/0310668.
13. A. V. Andreev *et al.*, *Phys. Rev. Lett.* **91**, 56803 (2003).
14. R. G. Mani *et al.*, *Phys. Rev. Lett.* **92**, 146801 (2004).

NEUROSCIENCE

Matchmaking

Nathaniel D. Daw and Peter Dayan

Sensation is simple, execution easy, according to a burgeoning line of work in neurophysiology, neuroimaging, and neural computation. The difficult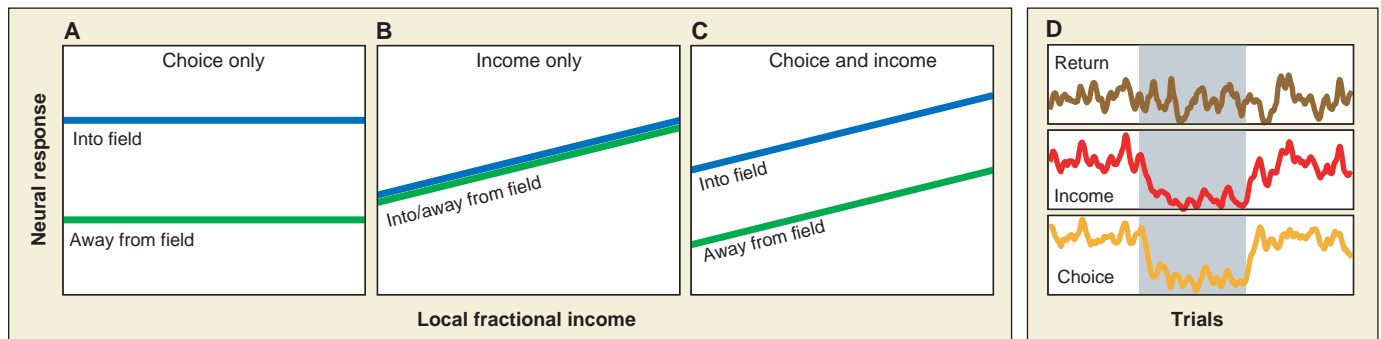y arises in getting them properly married. Situated in the middle between sensation and action is decision-making. Choosing between actions usually requires learning how these actions determine the rewards and punishments that an environment affords, and how this depends on sensory circumstance.

The authors are at the Gatsby Computational Neuroscience Unit, University College London, London WC1N 3AR, UK. E-mail: dayan@gatsby.ucl.ac.uk

Computationally, decision-making is hard to study because these factors are nonstationary: The environment may be dynamic. The subject's knowledge about the environment surely is. Experimentally, the information associated with choice (like most quantities represented at refined levels of the cortical hierarchy) is rather slippery. It is only loosely amenable to experimental control, and only dimly illuminated by behavior. The study by Sugrue *et al.* on page 1782 of this issue (*1*), along with other recent reports (*2*–*4*), exemplifies what is currently the most successful strategy to overcome these obstacles. The idea is to ground a detailed analysis of behavior on a computational theory of near-optimal performance. Key internal variables from al-

gorithms that instantiate the theory can then be correlated with, and thereby elucidate, neural activity.

Sugrue *et al.* (*1*) study a decision task in which, on each trial, a monkey must make an eye movement toward one of two colored targets to indicate its choice. The choice may be rewarded with juice, which becomes available sporadically (and unsignaled) at each target, with different characteristic payoff frequencies. If these frequencies are fixed, then many species adopt a reasonably effective behavioral strategy called matching (*5*). Under this strategy, the probability of choosing a target roughly equals the proportion of the total reward (income) earned from that target (called "fractional income"). In this study, frequencies were constantly changed, producing continual behavioral adaptation. The experiment captures the metrics of behavioral plasticity by relating the ever-mutating choices to the ever-mutating fractional incomes. It also probes the neural

The marriage between sensation and action. (A to C) Hypothetical relationships between the proportion of integrated reward (called fractional income) earned in the recent past for eye movements toward a target and the trial-to-trial activity of a LIP neuron with a response field covering that target. The blue line shows neural activity for trials in which the choice of eye movement was into the response field. The green line shows neural activity for trials in which the eye movement was away from the response field. (**A**) LIP neuron activity correlates with choice, but not fractional income. (**B**) LIP neuron activity correlates with fractional income, but not choice. (**C**) LIP neuron activity correlates with both choice and fractional income. [The third option is consistent with figure 4A of Sugrue *et al.* (*1*), because the abscissa for their green curve is 1 minus the local fractional income.] (**D**) Smoothed return (payoff per choice of target) (brown) during a period when the target payoff frequencies are swapped (gray background). Although smoothed fractional income (red) and choice behavior (orange) change markedly, the return remains roughly constant. Also, local fluctuations in all three quantities are correlated.

mechanics of decision-making by relating dynamic choices and incomes to fluctuating neural activity in an area of the brain called the lateral intraparietal cortex (LIP), a site that helps to control eye movements.

Temporally local estimates of the fractional incomes calculated using a low-pass filter (leaky integrator) fairly closely match changing choice probabilities. The filter's time constant determines over how many trials income is averaged. Intriguingly, the value that best fits the monkeys' eye movement behavior is nearly optimal for harvesting rewards, at least in the restricted class of algorithms considered. In a similar task, a monkey's time constant seemed to evolve systematically with experience (*6*). These findings raise the possibility that animals are engaging in metalearning, adaptively tuning parameters that control learning. Testing this idea would involve manipulating the task parameters that determine the optimal time constant, such as the prevalence of changes to the payoff frequencies.

Many LIP neurons respond preferentially, prior to eye movements, to particular locations (their spatial "response fields"). The monkeys' behavior suggests investigating how this neural activity relates to the local fractional income of targets in the response fields. The figure illustrates various possibilities: correlation of neural activity purely with ultimate target choice (panel A); correlation of neural activity purely with fractional income (panel B), implying that other structures (perhaps subcortical areas associated with habit learning) might control the behavior; or, as actually supported by the data here, correlation of neural activity with a combination of both target choice and fractional income (panel C). There is no evidence in

the Sugrue *et al.* study that choice and income correlation trade off differentially during the delay preceding eye movements, as suggested previously (*2*).

Several questions arise. First, this task has deeper psychological than computational roots. The field of reinforcement learning (*7*) has focused on a different class of task, which allows for choices to have delayed consequences. A key proposal based on this theory is that learning about rewards is driven by errors in predictions of those rewards. These errors are believed to be reported by the neuromodulator dopamine (*8*). It is not clear how (or even whether) dopamine controls the ongoing estimation of fractional incomes in the matching task.

Reinforcement learning and economic decision theory concentrate on expected reward per choice (return), which, unlike income, explicitly factors out choice probability. Return behaves unusually in perfect matching, because it is *equal* for both alternatives, even if the programmed payoff frequencies differ. A temporally local estimate of the return would thus be *unaffected* by swapping the programmed payoff frequencies between targets (see the figure, panel D), even though this change strongly modulates choice behavior and fractional income. An analogous prediction has been investigated in a game-playing task. Early results suggest that LIP responses and behavior are dissociated (*3*).

This raises the critical issue of model validation. Fractional income is strongly correlated with analogous measures defined by innumerable alternative models. Standard statistical model selection procedures could address the question of which theory best explains the LIP responses. Applied to monkeys' behavior on a similar

task, such analysis favors a somewhat different model (*9*). It is not yet clear whether the alternative income estimates of this model (or indeed others) correlate better with neural responses.

Finally, what critical computations, if any, do LIP neurons perform in this task, and how is income involved? As Sugrue *et al.* note, LIP, with its spatially localized response fields, seems an unlikely substrate for tracking the incomes of targets, which, in this task, are determined by color rather than location. Are decision theoretic quantities such as income ubiquitous, with LIP merely being an example? Recent data (*10*) suggest an apparently specific function for LIP in integrating over time noisy sensory information pertaining to a choice, an aspect of decision-making not exercised by the matching task. Tasks combining sensory integration and complex payoff contingencies are under way. If cortical areas like LIP indeed have an intimate role to play in the marriage between sensation and action, then perhaps it will take the theory-driven approach of Sugrue *et al.* to expose it fully.

References and Notes
1. L. P. Sugrue, G. S. Corrado, W. T. Newsome, *Science* **304**, 1782 (2004).
2. M. L. Platt, P. W. Glimcher, *Nature* **400**, 233 (1999).
3. M. C. Dorris, P. W. Glimcher, *Soc. Neurosci. Abstr.* 767 (2003).
4. J. D. O'Doherty *et al.*, *Neuron* **38**, 329 (2003).
5. R. J. Herrnstein, *J. Exp. Anal. Behav.* **4**, 267 (1961).
6. B. Lau, P. W. Glimcher, *Soc. Neurosci. Abstr.* 518.16 (2003).
7. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA, 1998).
8. W. Schultz *et al.*, *Science* **275**, 1593 (1997).
9. B. Lau, P. W. Glimcher, personal communication.
10. J. D. Roitman, M. N. Shadlen, *J. Neurosci.* **22**, 9475 (2002).
11. The authors are funded by the Gatsby Charitable Foundation, the Royal Society, and the EU BIBA project.