work on chirally coupled nanomagnets[7], the authors fabricated a sort of artificial, stationary domain wall in a magnetic cobalt wire interfaced with non-magnetic platinum. The magnetization in the cobalt is perpendicular to the plane of the wire, except in the stationary region. There, it is magnetized in the direction of the wire's long axis, like the region in the middle of an ordinary domain wall, but across a much larger width. This is crucial, because it allows smaller coercivity — that is, the magnetization here is easier to switch.

To picture how the inverter works, consider an input consisting of a domain wall that has left-handed chirality (Fig. 1). This mobile wall is rolled along the wire by spin–orbit torque. When it reaches the fixed artificial boundary, two opposite magnetic moments collide, producing a region of the wire in which the moment changes abruptly. According to theories of magnetism, such an abrupt change has a high energy cost. To lower the energy of the system, one of the moments must be switched, or a new magnetic domain must be generated. In this case, the moment in the low-coercivity fixed wall switches to the same direction as that in the incoming wall.

But a chirality effect now comes into play: this switch of magnetic moment produces a right-handed chirality at the other side of the fixed wall that conflicts with the chirality preferred by the DMI. To resolve this, a new domain wall forms on that side (the system is shaped in such a way as to promote this process) and sets off along the wire. The moments in the resulting outgoing bit thus have the preferred left-handed chirality, rather than the right-handed chirality originally produced at the wall.

By integrating their inverters into junctions, Luo *et al.* designed some simple logic gates (NAND and NOR), as well as more-complicated ones (such as exclusive-OR). Each junction has two inputs, an intrinsic bias towards one magnetic moment and one output. The output is determined by the two inverted inputs and by the bias at the junctions (rather like a 'majority gate'[8]). So, when the inputs and bias are (0,0) and 1, respectively, inverters immediately before the junction invert them to (1,1) and 0 at the junction itself, which consequently outputs 1, acting as a NOT gate. But when the inputs are either (1,0) or (0,1), the value of the bias determines whether the gate behaves as a NOR or a NAND. This majority-gate behaviour mitigates the need to precisely synchronize the two inputs, offering reliable logic operations.

This logic system satisfies key criteria known as cascadability and fan-out. Cascadability means that the output of one gate is produced in the correct form and is strong enough to drive the input of the next gate. And fan-out means that one gate output can be connected to several gate inputs[9]. Moreover, the data can be stored in the absence of an external power source, and evade damage by ionizing radiation.

Challenges remain before chips based on Luo and co-workers' system can reach the market. The operating current will need to be reduced so that it can be accommodated by tiny complementary metal oxide semiconductor (CMOS) transistors, which help to pick up inputs and outputs for use in chips. In theory, current decreases as the size of wires and transistors decreases, and so current density (the charge per unit time that flows through a given cross-section of the wire) remains constant with scaling. A reduction in current density will be needed to increase the speed and reduce the energy consumption of the authors' system. Domain-wall velocity does not scale linearly with current, and so new materials might need to be used to reduce the current density[10,11].

Another issue is that the input and output states of Luo and colleagues' system have to be detected by microscopy, rather than by an electrical method. A different read-out system will be needed for practical applications, but this could be technically challenging. An effect known as tunnelling magneto-resistance might offer one solution[12]. The implementation of a domain-wall logic chip that uses an electrically driven read-out system should be the next goal, following on from Luo and colleagues' exciting discovery.

See-Hun Yang is at IBM Research – Almaden, San Jose, California 95120, USA.
e-mail: seeyang@us.ibm.com

1. Allwood, D. A. *et al. Science* **309**, 1688–1692 (2005).
2. Luo, Z. *et al. Nature* **579**, 214–218 (2020).
3. Parkin, S. & Yang, S.-H. *Nature Nanotechnol.* **10**, 195–198 (2015).
4. Ryu, K.-S., Thomas, L., Yang, S.-H. & Parkin, S. *Nature Nanotechnol.* **8**, 527–533 (2013).
5. Emori, S., Bauer, U., Ahn, S.-M., Martinez, E. & Beach, G. S. D. *Nature Mater.* **12**, 611–616 (2013).
6. Omari, K. A. *et al. Adv. Funct. Mater.* **29**, 1807282 (2019).
7. Luo, Z. *et al. Science* **363**, 1435–1439 (2019).
8. Imre, A. *et al. Science* **311**, 205–208 (2006).
9. Miller, D. A. B. *Nature Photon.* **4**, 3–5 (2010).
10. Yang, S.-H., Ryu, K.-S. & Parkin, S. *Nature Nanotechnol.* **10**, 221–226 (2015).
11. Avci, C. O. *et al. Nature Nanotechnol.* **14**, 561–566 (2019).
12. Julliere, M. *Phys. Lett. A* **54**, 225–226 (1975).

Neuroscience

Research on the cerebellum yields rewards

Jennifer L. Raymond

A brain structure called the cerebellum has mostly been associated with learning from errors. The discovery that the cerebellum is also involved in reward-driven learning in monkeys implies a previously unappreciated role in cognition.

People and organizations alike use rewards, from snacks to salary bonuses and frequent-flyer miles, to shape behaviour through a process called reinforcement learning. For example, if a dog receives a treat for rolling over in response to a verbal command, the likelihood of that behavioural response to the verbal cue will increase. Writing in *Neuron*, Sendhilnathan and colleagues[1] describe neuronal signals that could support such reward-driven learning. What is remarkable is where the authors found these signals — not in the brain areas that have long been implicated in reinforcement learning, but in the cerebellum, a brain structure historically associated with error-driven, rather than reward-driven, learning.

The cerebellum is best known for its role in motor-skill learning — the process by which movements become smooth and accurate through practice. Fifty years of research[2] supports the idea that when you practise a movement, such as your tennis backhand, the cerebellum uses feedback about errors to gradually refine the accuracy of the movement by weakening the neuronal connections that are responsible for those errors. It has been widely assumed that the cerebellum uses a similar, error-correcting learning algorithm to support cognition[3], because the regions of the cerebellum that contribute to cognitive functions such as navigation[4] and social behaviour[5] have the same basic circuit architecture as those that control movement.

In the past three years, however, there has been a flurry of studies showing reward-related neuronal activity in the cerebellum[6–12]. What are reward signals doing in an error-correcting part of the brain? Sendhilnathan *et al.* leveraged the rapid learning abilities of monkeys to gain fresh insights into reward-related signalling in the cerebellum.

In each experimental session, the authors presented a monkey with two visual cues it had never seen before on a computer screen.

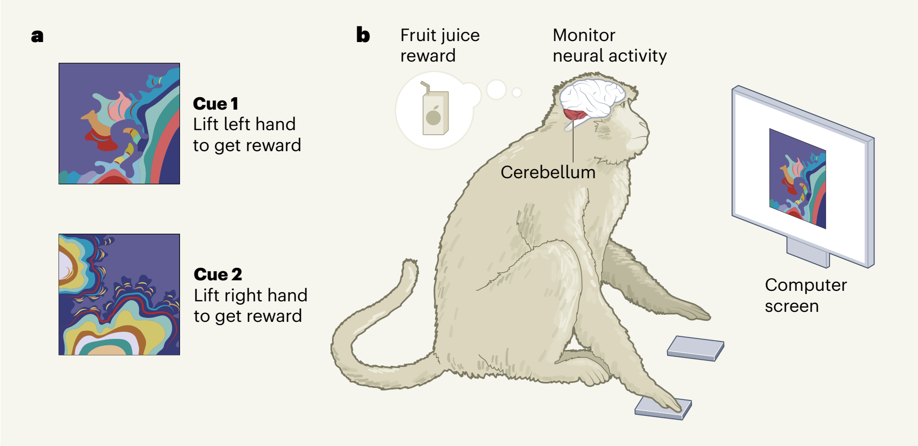

Figure 1 | A rewarding choice. a, Sendhilnathan *et al.*[1] examined neural activity in a brain region called the cerebellum during reward-driven learning. The authors presented monkeys with two visual cues. For one cue, the monkey needed to lift its left hand to receive a reward of fruit juice; for the other, lifting the right hand would lead to a reward. **b**, The monkeys performed a series of trials in which they were presented with one of the two cues. The authors monitored cerebellar neuronal activity while the monkeys learnt, through trial and error, which response would produce a reward for each cue. They found that a subpopulation of neurons carried information about the success or failure of the previous trial until the next trial was completed (not shown). (Figure adapted from Fig. S2 of ref. 1.)

One arbitrarily assigned cue would result in the monkey receiving a reward of fruit juice if the animal responded by lifting its left hand. The other cue would result in a reward if the monkey lifted its right hand. The researchers monitored the activity of neurons called Purkinje cells in the cerebellum as the monkeys learnt, through trial and error, to make the correct response to each visual cue (Fig. 1).

Sendhilnathan *et al.* found that the activity of cerebellar Purkinje cells carried information about the success or failure of the monkey's most recent attempt at the task. One subpopulation showed high activity following a correct response to the cue; another showed high activity following a failed attempt. These signals arose a few hundred milliseconds after the end of a trial and persisted until the next trial was completed. As such, they seemed to provide a working memory that could enable the outcome of one trial to guide the next behavioural choice.

These signals are reminiscent of those carried by neurons in frontal and parietal regions of the brain's cerebral cortex, which encode the 'value' of different behavioural choices on the basis of reward history over multiple trials[13]. In the current study, the cerebellar neurons kept track of only the most recent trial's outcome. But in this task, the outcome of a single trial provides sufficient information for the monkey to infer the correct response for the next trial — if a reward was not given when a monkey lifted its right hand in response to one visual cue, for instance, then the correct response to that cue must be to lift the left hand, and the correct response to the other visual cue would be to lift the right hand. It would be interesting to know whether cerebellar neurons can keep track of a more-extended history of rewards should the task require it, and whether the cerebellum interacts with the cerebral cortex in performing this computation.

Importantly, information about the previous trial's outcome was present in the cerebellum only when a new set of cue–response associations was being learnt. As monkeys improved their performance over trials, the neuronal activity encoding each outcome waned. Moreover, the signal was not present when monkeys earned rewards by responding to a pair of visual cues that they had mastered through several months of training. These observations indicate that cerebellar neurons are not simply carrying information about rewards, predictions about rewards or the movements that animals make when anticipating rewards. Rather, the cerebellum seems to contribute specifically to learning about how to earn rewards in a new situation. The authors speculate that the cerebellum might enhance the rate of learning about rewards, a possibility supported by the recent discovery in rodents of direct, excitatory projections from the cerebellum to neurons in the brain stem that release the reward-associated neurochemical dopamine[14].

> **"The cerebellum seems to contribute specifically to learning about how to earn rewards in a new situation."**

There are several intriguing parallels between the signals found by Sendhilnathan and colleagues and the signals involved in cerebellar control of movement. First, as with reward-driven learning, for some motor skills, cerebellar Purkinje cells contribute selectively to new motor learning and not to performing older motor skills[15,16]. Second, Purkinje-cell activity carries information that could guide both ongoing behaviour and the induction of learning during motor- and reward-based learning[17]. Third, the Purkinje cells carry signals that could support working memory in the form of activity maintained from one trial to the next in reward-based learning, and in the form of activity maintained during a delay period between a cue and the motor response to the cue, which seems to support motor planning[11,18]. Finally, during both types of learning, individual Purkinje cells are active for a specific time period of a few hundred milliseconds, with information seemingly passed from cell to cell over time[19]. These striking parallels raise the possibility that the cerebellum performs a similar function for error-driven motor learning and reward-driven reinforcement learning.

We learn from both our successes and our failures. These two learning schemes were previously attributed to distinct brain structures, but the current results, along with those of others[6–12], blur these mechanistic and conceptual boundaries. As such, the work highlights the need to consider how long-range interactions between brain areas support the shaping of behaviour by experience.

Jennifer L. Raymond is in the Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA.
e-mail: jennifer.raymond@stanford.edu

1. Sendhilnathan, N., Ipata, A. E. & Goldberg, M. E. *Neuron* https://doi.org/10.1016/j.neuron.2019.12.032 (2020).
2. Albus, J. S. *Math. Biosci.* **10**, 25–61 (1971).
3. Schmahmann, J. D. *Neurosci. Lett.* **688**, 62–75 (2019).
4. Rochefort, C., Lefort, J. M., Rondi-Reig, L. *Front. Neural Circuits* **7**, 35 (2013).
5. Wang, S. S., Kloth, A. D. & Badura, A. *Neuron* **83**, 518–532 (2014).
6. Wagner, M. J., Kim, T. H., Savall, J., Schnitzer, M. J. & Luo, L. *Nature* **544**, 96–100 (2017).
7. Heffley, W. *et al. Nature Neurosci.* **21**, 1431–1441 (2018).
8. Heffley, W. & Hull, C. *eLife* **8**, e46764 (2019).
9. Larry, N., Yarkoni, M., Lixenberg, A. & Joshua, M. *eLife* **8**, e46870 (2019).
10. Kostadinov, D., Beau, M., Blanco-Pozo, M. & Häusser, M. *Nature Neurosci.* **22**, 950–962 (2019).
11. Chabrol, F. P., Blot, A. & Mrsic-Flogel, T. D. *Neuron* **103**, 506–519 (2019).
12. Lixenberg, A., Yarkoni, M., Botschko, Y. & Joshua, M. *J. Neurophysiol.* **123**, 786–799 (2020).
13. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. *Nature Rev. Neurosci.* **6**, 363–375 (2005).
14. Carta, I., Chen, C. H., Schott, A. L., Dorizan, S. & Khodakhah, K. *Science* **363**, eaav0581 (2019).
15. Shutoh, F., Ohki, M., Kitazawa, H., Itohara, S. & Nagao, S. *Neuroscience* **139**, 767–777 (2006).
16. Jang, D. C., Shim, H. G. & Kim, S. J. Preprint at bioRxiv https://doi.org/10.1101/513283 (2019).
17. Nguyen-Vu, T. D. B. *et al. Nature Neurosci.* **6**, 1734–1736 (2013).
18. Gao, Z. *et al. Nature* **563**, 113–116 (2018).
19. Li, J. X., Medina, J. F., Frank, L. M. & Lisberger, S. G. *J. Neurosci.* **31**, 12716–12726 (2011).