## 1

# Our Permanently Plastic Brains

**Gina Rippon**

(2019)

One of the most important innovations in brain science in the last 30 years or so is the understanding of just how plastic or moldable our brains are, not only in the early years of development but throughout our lives, reflecting our experiences and the things we do and, paradoxically, the things we don't do.

This is a big change from our early understanding of how our brain developed, which was based on the notion that there were fixed, predetermined patterns of growth and change that unrolled over set time periods, with major deviations arising only via relatively extreme events during these periods.

The discovery of lifelong "experience-dependent neuroplasticity" has drawn attention to the crucial role that the outside world—the lives we live, the jobs we do, the sports we play— will have on our brains. Whereas we used to wonder whether our brains were more a product of "nature" or "nurture," we now realize that the "nature" of our brains is entangled with the brain-changing "nurture" provided by our life experiences.

The most famous example of neuroplasticity is the London taxi-driver studies carried out by University College London neuroscientist Eleanor Maguire and her team. Maguire showed that four years of "doing the Knowledge," the extensive training for taxi drivers that requires memorizing different routes through the 25,000 or so London streets within a six-mile radius of Charing Cross station (and is necessary to qualify for a taxi license), resulted in gray-matter increases in the posterior part of each successful trainee's hippocampus, the part of the brain that underpins spatial cognition and memory. This wasn't because the aspiring cabbies already had bigger hippocampi (Maguire tracked both trainees and retired taxi drivers and mapped increases in the former and decreases in the latter) or because they were having to navigate complex driving routes (bus drivers with fixed routes didn't show the same effect). Maguire also looked at trainees who failed the course and found that they did not show the hippocampal changes that characterized their successful colleagues. There appeared to be a cost to this brain-changing expertise; successful taxi drivers were significantly worse on other tests of spatial memory. However, retired taxi drivers, while showing a return to "normal" gray-matter volume in their hippocampi (and declines in their previous London-specific navigational skills), displayed improved levels of performance in ordinary spatial memory. So this group of studies shows both the ebb and flow of brain plasticity, with shifts in the allocation of brain resources coming and going in the context of acquiring, using, and losing a particular skill.

Understanding neuroplasticity also has implications for understanding individual differences in what might seem to be everyday skills. The taxi-driver studies could be taken as a measure of the plasticity of the brain, but "the Knowledge" is a highly specialized skill acquired from scratch in adulthood. What about more routine skills? Why are some people better at these than others? Is this reflected in brain activation patterns? Can you improve these kinds of skills, and does this change the brain?

There is certainly evidence that more experience with activities related to certain skills can both improve your performance and change your brain. In 2005, psychologists Melissa Terlecki and Nora Newcombe showed that computer and video-game usage was a powerful predictor of certain spatial skills. It also explained most of the gender differences that had been reported for this particular skill—there was a much higher level of computer use and video-game playing among the male participants, and it appeared to be this that was driving their better spatial skills.

It seems this kind of behavioral plasticity is reflected in structural brain changes as well. In 2009, psychologist Richard Haier and colleagues measured structural and functional brain images in a group of girls before and after a three-month stint of playing *Tetris* for, on average, one and a half hours a week. Compared to a matched group who didn't play *Tetris*, the girls' brains showed enlargement in cortical areas associated with visuospatial processing. There were also changes in the *Tetris*-induced bloodflow measures. In a different study, 30 minutes a day of playing *Super Mario 64* over a period of 2 months also proved to be a brain-changing experience, with increases in gray-matter volume in the hippocampus, as well as the frontal areas of the brain. Interestingly, such brain and performance changes are not task-specific. One study showed that 18 hours of origami training improved mental rotation performance and changed the brain correlates associated with it.

Recognizing lifelong brain plasticity and the role of external factors such as experience and training means that we will need to revisit past certainties about fixed, hardwired, biologically determined differences. Understanding any kind of differences between the brains of different people means we will need to know more than what sex or age they are; we will need to consider what kind of lifetime experiences are embedded in these brains.

This state of lifelong neuroplasticity offers a much more optimistic view of our brains' futures. But it can also offer insights into what is happening to our brains in the present—how our brains can and will be changed by what they encounter in our world, how our brains can get diverted and derailed. Knowing more about how our brains engage with the world means we have to pay much more attention to what is in that world.

## Your Brain as a Predictive Satnav

The plastic and changeable nature of our brains suggests that they are not just rather passive (though hugely efficient) information processors but instead are constantly reacting and adjusting according to the huge swathes of information that are fired at them every day—we now think of the brain as a proactive guidance system, continuously generating predictions as to what might be coming next in our worlds (known in the business as "establishing a prior"). Our brains monitor the fit between these predictions and the real outcome, passing back error messages so that the prior is updated, and we're guided safely through the unremitting streams of information with which we are constantly bombarded. The core aim of this system is to minimize "prediction error" by speedily and continuously generating and updating priors based on the normal course of events. These will draw on pretty minimal amounts of information to estimate the next step and ensure no surprises, efficiently reducing the need for cognitively wasteful rechecking or "overthinking." In the light of feedback about a mismatch, a quick reconstruction of a new prior will follow. So, our brain navigates us through the world via a combination of predictive-texting-like skills and high-end satnav guidance.

If you ever visit Hanoi, you'll see a traffic-based version of predictive coding at work. The roads are filled with what seems like a never-ending, never-stopping stream of motor scooters, packed wheel to wheel across the width of the road. On my first visit there, I hovered hopelessly on the pavement, waiting for the gap that never came. At last, a tiny old Vietnamese lady took pity on me, took me by the arm and signed for me to come with her, adding instructions to "NOT STOP." Fixing a glare on a spot on the other side, she led me into the stream of scooters and steadily walked through. The scooters smoothly swirled round us and we made it across. It was later explained to me that the "NOT STOP" is the crucial ingredient—the scooter drivers appear to have an uncanny instinct of knowing just where in their path you are likely to be as they approach you (establishing their prior) and adjust their trajectories to steer round you accordingly. If you stop, you aren't where they expect you to be and you become an instant "prediction error"—with bruising and undignified consequences.

It is claimed that our brain's "predictive coding" power is not only applied to the most basic sights and sounds and movements but also allows us to engage with higher-level processes such as language, art, music, and humor, as well as the often hidden rules of social engagement, underpinning our ability to predict the actions and intentions of other people and interpret their behavior accordingly. The guidelines we employ are extracted from our outside world, the "data in" side of things, and used to generate rules to determine the next most likely outcome in life's rich pattern, what behaviors are associated with what facial or verbal expressions, what intention is being flagged up by what action. The rules that are extracted can range from "this kind of smell usually results in finding something good to eat" to "that kind of facial expression usually means that someone is happy" or to even more abstract and hard-to-define rules of social engagement, such as understanding turn-taking in conversations.

Most of the time, of course, our brains are indeed hyperefficient—their best guesses, with just the right amount of precision behind them, almost always provide the winning ticket. But the fact that the system is not infallible is revealed by phenomena such as visual illusions, where we might see a triangle where there isn't one, just because a particular configuration of shapes is normally associated with the presence of a triangle. The system can be tricked by "misdirecting" the establishment of priors. If the brain is busy with solving a very specific problem, it can overlook information that tells it that something else is going on at the same time and miss this key prediction error. Our attention to what is going on around us can be very, very selective, and we can easily miss something that is in plain sight but unexpected.

But sometimes the speedy shortcuts can let us down more seriously. The brain's templates or "guide images" can be over-general, lumping several varieties of information into a single category in order to cut down on the amount that has to be scrutinized and sorted, especially if that is what is on offer in the outside world. Our brains are, in fact, the ultimate stereotypers, sometimes drawing very rapid conclusions based on very little data or based on strong expectations, arising from personal past experience or from the cultural norms and expectations of our surroundings. In a 2015 opinion for *The New York Times*, psychologists Lisa Feldman Barrett and Jolie Wormwood described the phenomenon of "affective realism," where your feelings and expectation affect the prediction process and your perception. You, quite literally, see things differently. Barrett and Wormwood used the example of newly released statistics on shootings of unarmed individuals by police, where officers, in the context of challenging a suspect, had misidentified a mobile phone, wallet, or other object in the suspect's hand as a gun. The authors also reported studies in which a neutral face, when viewed in parallel with a subliminally presented scowling face, was perceived as less trustworthy, unattractive, and more likely to commit a crime. So external data and expectations can divert and distract our otherwise helpful predictive guidance system. Stereotypes can and do change how we see the world.

It is also the case that the system may not distort what is happening in the outside world but may, all too accurately, exactly reflect it. In 2016, Microsoft launched a chatbot named Tay, based on an interactive conversation-understanding program, which was to be trained online to engage in "casual and playful conversation" by interacting with Twitter users. Within 16 hours, Tay had to be shut down: starting off tweeting about how "humans are super cool," it quickly became a "sexist, racist asshole" thanks to the multiple prejudice-laden tweets that were being input. Although some of Tay's responses were just imitations, there was evidence of general rules being extracted from common themes, resulting in statements that had never specifically been made, such as "feminism is a cult," which Tay had "learned" by putting together what it knew about the characteristics of cults with the statements it was receiving about feminism.

The process behind this experiment is modeled on a system of training computers called "deep learning." Computers are programmed to extract patterns from information and to "self-train," to achieve ever more nuanced representations of the outside world, rather than be programmed to carry out specific tasks. This is at the heart of today's developments in computer-based artificial intelligence and has parallels in contemporary models of how the brain learns. And, just as poor old Tay found out, if the world our brains are getting their data from is sexist, racist, or rude, then the priors that guide our experience of the world may well be the same.

In terms of trying to understanding the emergence of sex differences and the role of brain-environment interactions, neuroscientists have been fascinated to see that one of the problems that these deep learning systems are having is that if the data being input are intrinsically biased, then this is the rule that the system will learn. If a system is trying to generate a rule associated with images of kitchens, it will link these to women because that is what it finds in the outside world it is exploring. When one computer program was asked to complete the statement "*Man* is to *computer programmer* as *woman* is to *X*," it supplied the response "homemaker." Similarly, a request to characterize business leaders or CEOs produced lists and images of white men. A recent study showed that simply inputting language data into a system that was learning to recognize images not only revealed significant gender bias, but also magnified it. So while in actuality "cooking" might be more likely to involve women than men 33% of the time, the computer model cheerily learning to tag images of cooking might label it as a female activity up to 68% of the time, due to the imbalance on the web of examples of who "did" cooking.

The researchers "training" this model checked out other language examples from the internet that might be input into such learning systems and discovered that 45% of verbs and 37% of objects showed some kind of gender bias of more than two to one; that is, it was twice as likely that certain verbs or certain objects would be associated with one gender rather than another. They then went on to show how you could constrain the model to more accurately reflect the bias. They made no comment as to its existence in the first place (although they did call their paper "Men Also Like Shopping").

So, in today's understanding of the brain, we are appreciating more and more that what our brain does with our world very much depends on the information it has extracted from that world, and the rules it has generated for us are based on this information. To establish its priors, our brain will act like an eager "deep learning" system. If the information it soaks up is biased in some way, perhaps based on prejudice and stereotypes, then it is not hard to see what the outcome might be. Just like the outcomes of overreliance on a misinformed satnav, we may find ourselves steered down unsuitable pathways or taking unnecessary detours (or we may even give up the journey altogether).

The key issue here is that how our brains determine the way in which we respond to our world, and how that world responds to us, is much more entangled with that world than we used to think. Brain differences (and their consequences) will be as much determined by what is encountered in the world as by any genetic blueprint or hormonal marinade, so understanding these differences (and their consequences) will necessitate a close look at what is going on outside our heads as well as inside.

Another shift in focus in the 21st century has been on what aspects of human behavior we neuroscientists are trying to explain. Much of the speculation about the evolution of the human brain has concentrated on the emergence of high-level cognitive skills (such as language, mathematics, abstract reasoning, and the planning and execution of complex tasks) and how these contributed to the success of *Homo sapiens*. But there is an increasing focus on the idea that human success is actually based on the fact that we have learned to live and work cooperatively, to decode the invisible social rules that are signaled by facial expression and body language or that just appear to be understood by "in-group" members. We need to understand which people our group includes and how we should behave in order to be accepted by that group. We also need to spot those who are *not* group members and why. We need to read our fellow human beings' minds and understand their beliefs, intentions, hopes, and wishes; see things from their perspective and predict how this might make them behave; and adjust our own behaviors to encompass, or perhaps thwart, the goals of others.

Exploring how and when we humans use our brains to become social beings has led to a new branch of cognitive neuroscience—social cognitive neuroscience—and a new model of the brain: the "social brain." Social cognitive neuroscientists explore the neural real estate behind our drive to be a member of the many social and cultural networks that surround us and, further, show how the entanglement of our brains with these networks will come to shape our brains themselves.

## References

Barrett, L. F., & Wormwood, J. (2015, April 17). Opinion: When a gun is not a gun. *The New York Times Sunday Review*. https://www.nytimes.com/2015/04/19/opinion/sunday/when-a-gun-is-not-a-gun.html

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). *Man is to computer programmer as woman is to homemaker? Debiasing word embeddings*. https://arxiv.org/pdf/1607.06520.pdf

Groen, D. (2018, June). How we made AI as racist and sexist as humans. *The Walrus*. https://thewalrus.ca/how-we-made-ai-as-racist-and-sexist-as-humans

Haier, R. J., Karama, S., Leyba, L., & Jung, R. E. (2009). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Research Notes*, *2*, 174. https://doi.org/10.1186/1756-0500-2-174

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., & Gallinat, J. (2014). Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a commercial video game. *Molecular Psychiatry*, *19*(2), 265–271. https://doi.org/10.1038/mp.2013.120

Maguire, E. A., Gadian, D. G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R. S., & Frith, C.D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4398–4403. https://doi.org/10.1073/pnas.070039597.

Simonite, T. (2017, August 21). Machines taught by photos learn a sexist view of women. *Wired*. https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women.

Terlecki, M. S., & Newcombe, N.S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles*, *53*, 433–441. https://doi.org/10.1007/s11199-005-6765-0

Vincent, J. (2016, March 24). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge*. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's layout drives structural brain changes. *Current Biology*, *21*(24), 2109–2114. https://doi.org/10.1016/j.cub.2011.11.018

Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K.-W. (2017). *Men also like shopping: Reducing gender bias amplification using corpus-level constraints*. https://arxiv.org/pdf/1707.09457.pdf