



An advert for Pelmanism, a brain-training technique that became popular in the early twentieth century.

base, researchers usually turn to meta-analysis to assess the evidence. Unfortunately, in this field, even meta-analyses are producing divergent conclusions. Some researchers have suggested that the inconsistencies stem from the use of inadequate control groups and measures of outcome. But most agree that the field needs bigger, better studies — and a return to basic science.

THE TROUBLE WITH TRAINING

Numerous studies purport to show benefits from cognitive training, but delve deeper and there is less than meets the eye. Many of these studies show little more than improved performance on tasks closely related to those that participants trained for — known as near transfer. “When you practise something, of course you get better at it,” says Jaeggi. “The real question is whether there is far transfer”, referring to when training benefits different cognitive abilities. This is why a lot of training research has focused on working memory. Working-memory capacity predicts everything from reading ability to academic achievement, and correlates highly with fluid intelligence. Increasing this capacity might, therefore, have a broad impact on cognition.

Among the most prolific early sceptics of far-transfer effects were Randall Engle and his fellow psychologists at the Georgia Institute of Technology in Atlanta. They pointed to two main recurring problems in working-memory training studies, the first being inadequate control groups. Many findings, they said, could be due to ‘Hawthorne effects’, referring to the fact that people change their behaviour when they know that they are being observed. To discount this effect, Engle and colleagues recommended that studies should use active control groups in which participants perform activities that are identical to the training in every aspect except the main task. For example, although Jaeggi and her team² tested their control group before and after training, they did not give them additional exercises; they accounted for ‘test–retest’ effects (people do better the second time around), but not phenomena such as the Hawthorne effect.

The second issue concerns the use of only one measure of an outcome. Participants can develop strategies during training that aid their performance on a task — for example, inner vocalization of a visual stimulus — without any change in their underlying ability. And, because no task taxes only one cognitive ability, practising one exercise can lead to improvements on seemingly unrelated tasks if there is overlap between the abilities that they engage. To avoid this possibility, the consensus among cognitive psychologists is that researchers should use a range of tasks that cover, for example, numeric, verbal and visuospatial abilities. Most studies have not done this.

There have now been several attempts to clarify the picture by pooling study results. In 2013, psychologists Monica Melby-Lervåg of Oslo

BRAIN TRAINING

Memory games

Conflicting results are expected in a young field, but what do you do when even the meta-analyses do not agree?

BY SIMON MAKIN

“New Minds for Old in 12 Weeks!” proclaimed adverts for Pelmanism, a brain-training technique that swept the United Kingdom in the early part of the twentieth century. Promotional material claimed that this system could combat such troubling mental phenomena as “lack of ideas” and “brain fag”. It became so successful that the Pelman Institute established offices in Australia, South Africa and the United States. Pelmanism was still being promoted as late as the 1960s, but has since sunk into obscurity, becoming merely a curious chapter in the history of psychology.

Except that almost 100 years later, brain training is back — this time with scientific backing. The first apparent breakthrough came in 2002 when a group of researchers in Sweden showed that training children with attention-deficit hyperactivity disorder on adaptive working-memory tasks — which test a person’s ability to retain and manipulate information over a short period of time

and that increase in difficulty to match performance — improved their attention and reasoning¹. But the real excitement came in 2008 with research led by psychologist Susanne Jaeggi, now at the University of California, Irvine. Jaeggi’s study seemed to show that healthy young adults who practised an adaptive working-memory task, which the authors called dual *n*-back (see ‘A test of sight and sound’), showed increases in the unrelated ability of fluid intelligence (the ability to reason in novel situations)². Furthermore, there was a dose effect: the more that people trained, the ‘smarter’ they became.

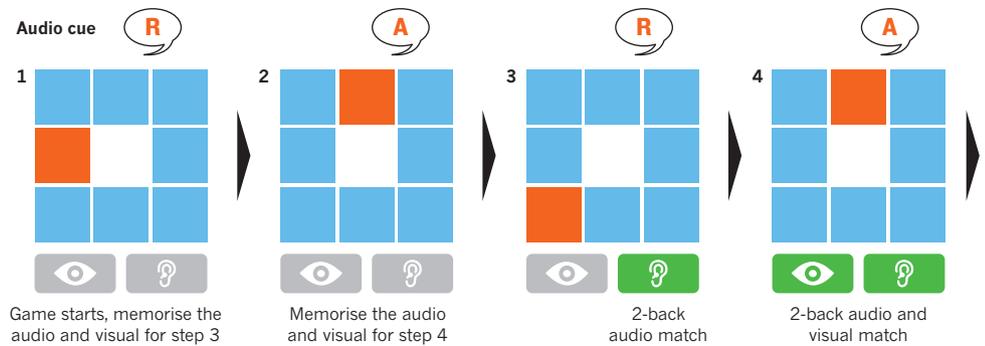
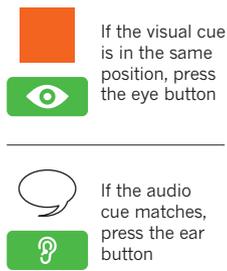
And, as with Pelmanism, a lucrative industry has grown up around the idea that cognitive performance can be enhanced by training. But the claim that mere hours spent playing a memory game can increase intelligence is an extraordinary one, and sceptics soon started voicing objections. Negative studies appeared, and the field is now awash with conflicting results. When faced with a large but uncertain evidence

➔ NATURE.COM

To watch an animation of the dual *n*-back test visit: go.nature.com/dofbmd

A TEST OF SIGHT AND SOUND

In the dual n -back test, participants are presented with an audio and a visual cue simultaneously. Participants must remember both and determine whether they match n turns later. In this example, $n = 2$.



University and Charles Hulme at University College London produced a meta-analysis³ that included 23 studies of adaptive working-memory training, each lasting at least 2 weeks. They found a small far-transfer effect on non-verbal reasoning — but only in studies that used passive control groups. In another meta-analysis⁴ of 20 studies that only looked at n -back training, psychologist Jacky Au of the University of California, Irvine, Jaeggi and their colleagues found a statistically significant positive effect on fluid intelligence — albeit small. “There is very good evidence that training on working-memory tasks does improve performance on tests of fluid intelligence,” says Au. But, he adds, it is not yet clear whether this translates to a meaningful, real-world increase in intelligence.

The analyses did little to settle the matter and the two groups have since exchanged further critiques^{5,6}. Hulme and Melby-Lervåg have gone so far as to call for journals to stop publishing studies that use passive controls⁵. “No analysis can get over having a passive control group,” says Hulme. “The only way round it is to have a control group that’s doing something very similar, with the same expectations that it’s beneficial.” But training studies are time-consuming and expensive; doubly so when using active controls. “Best practice is to use active controls,” Au admits, “but our work suggests the cost–benefit profile is not always clear, especially given the current state of funding.”

Meta-analysis is a powerful tool for finding a consensus in a confusing picture. However, it cannot make up for the design flaws of the constituent studies, says psychologist Claudia von Bastian of the University of Colorado, Boulder. And many brain-training studies have weaknesses. For example, the average number of participants per training group in Au’s meta-analysis was 20, meaning that most studies did not have enough participants for their results to be reliable — particularly when it comes to detecting small effects. This low statistical power not only increases the risk of finding spurious effects, but it also tends to inflate the size of any effect found.

Neither can meta-analysis adjust for studies that use a single (and therefore inadequate) measure of an outcome. Using multiple measures can enhance the validity of results, but some researchers advocate going further.

“There are statistical methods that allow you to analyse change in cognitive abilities instead of changes in test scores,” says cognitive neuroscientist Ulman Lindenberger of the Max Planck Institute for Human Development in Berlin. These ‘latent variables’ are inferred from the analysis of a battery of observed measures, and represent shared variance in those measures — in other words, changes in some underlying ability, such as reasoning or memory. Psychologist Florian Schmiedek of the German Institute for International Educational Research in Frankfurt, and his co-authors found that less than one-quarter of brain-training studies used multiple outcome measures, with only 7% using latent variables⁷. Schmiedek urges researchers to use more robust methods. “I don’t think this case will be closed by just running more of the kinds of studies that went into the recent meta-analyses,” he says. “The few studies that went to such effort paint a modest picture of the effectiveness of cognitive training.”

BACK TO BASICS

What is clear is that the effect of working-memory training on fluid intelligence lies somewhere between zero and very small, depending on whose analysis you trust. Taking the optimistic view, the question becomes one of opportunity cost: how much time and effort is needed to produce a lasting effect, and how does this compare to other uses of that time? Other activities purported to improve cognition include exercise, musical training and learning a new language. Unfortunately, studies of these activities often have similar problems to brain training, says von Bastian, from choosing valid controls to difficulties in untangling causality — and results are similarly contested.

Other researchers are examining the problem from a biological perspective. Proponents of brain training are fond of pointing out that the brain remains plastic throughout life. But, a degree of stability is also essential. Lindenberger and cognitive neuroscientist Martin Lövdén of the Karolinska Institute’s Aging Research Center in Stockholm, argue that the brain strikes a balance between plasticity and stability, and that this shifts as we age. From this perspective, changing something as fundamental as intelligence in an adult is a big deal. “Everything from a theoretical perspective suggests that to move

intelligence you would need to do something massive,” says Lövdén.

Some studies have looked for evidence of such changes at a neural level. The problem is that neuroscientists do not yet fully understand how everyday experience affects the brain. During learning, both activity and volume in parts of the brain initially increase, but then decrease — and the time course is not clear, says Lövdén. “We don’t have the link yet between experience, the brain and behavioural change,” he says. “We need to take a step back and try to understand the basic science.”

Von Bastian also advocates a return to the fundamentals. She is focused on developing a greater understanding of working memory — what its components are, the extent to which each might be malleable and how these components might affect other cognitive abilities. “That might be very interesting as a way to experimentally look at the relationship between working memory and intelligence,” she says. Research using this kind of theory-driven design, however, has been lacking.

To facilitate the sharing of training tasks, protocols and data, von Bastian has developed a web-based, open-source software package, Tatoon. The aim is to stimulate “methodologically rigorous research” with huge sample sizes that is unbiased by commercial products, she says. So far, more than 100 researchers are signed up, which, she says, should bring some new ideas and methodology to help explore the questions around whether brain training really works. “If we keep doing bad studies with small samples, we’ll never know.” ■

Simon Makin is a freelance science writer based in London.

1. Klingberg, T., Forssberg, H. & Westerberg, H. *J. Clin. Exp. Neuropsychol.* **24**, 781–791 (2002).
2. Jaeggi, S. M., Buschkuhl, M., Jonides, J. & Perrig, W. *J. Proc. Natl Acad. Sci. USA* **105**, 6829–6833 (2008).
3. Melby-Lervåg, M. & Hulme, C. *Dev. Psychol.* **49**, 270–291 (2013).
4. Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M. & Jaeggi, S. M. *Psychon. Bull. Rev.* **22**, 366–377 (2015).
5. Melby-Lervåg, M. & Hulme, C. *Psychon. Bull. Rev.* <http://doi.org/bb9d> (2015).
6. Au, J., Buschkuhl, M., Duncan, G. J. & Jaeggi, S. M. *Psychon. Bull. Rev.* <http://doi.org/bb9f> (2015).
7. Noack, H., Lövdén, M. & Schmiedek, F. *Psychol. Res.* **78**, 773–789 (2014).