Visual memory

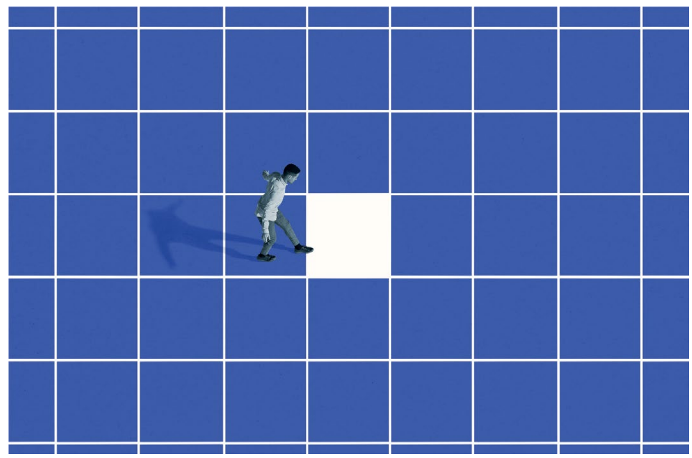
https://doi.org/10.1038/s41562-023-01565-1

Scaling up behavioural studies of visual memory

Jordan W. Suchow

Check for updates

A century of experiments on human visual memory have catalogued the many determinants of what people remember about their visual environments. In a massive experimental study of visual memory, Huang leverages mobile gaming to collect a dataset of 35 million behavioural responses that reveals how the mechanisms of visual spatial memory fit together.



CREDIT: KLAUS VEDFELT/DIGITALVISION/GETTY

When Sir Frederic Bartlett ran some of the early laboratory experiments on human memory, the only technology in the room was pen and paper[1]. Times changed. Slowly but surely, behavioural and social scientists retooled their laboratories with each decade's advances in hardware and software – from cathode-ray tube monitors for generating images to wearable devices for measuring biosignals. Then, laboratories moved online. These 'virtual lab' experiments have enabled behavioural and social scientists not only to scale up traditional experiment designs, but also to reconsider the fundamental structure of the experiments that they perform. In an article published in this issue of *Nature Human Behaviour*, Huang[2] reports the results of a study on human visual memory that uses a massive dataset collected through a mobile game to create a highly predictive model of visual spatial memory.

Visual memory supports the ability to remember visual stimuli and is critical to everyday activities such as driving a car or recognizing a familiar face[3]. Psychophysicists and cognitive psychologists have conducted thousands of carefully controlled laboratory experiments to identify the factors that determine what people remember about their visual environment[3]. Yet in most of these experiments, only a handful of factors can be studied at a time, which leaves many questions about how all the factors fit together unanswered. Developing a unified model of visual working memory that accurately predicts what will be remembered of an arbitrary visual stimulus remains a grand challenge of the cognitive sciences.

Huang responded to this challenge by leveraging scale through a partnership with the technology company Entelligence, which runs one of China's most popular puzzle-game mobile apps (with several million active users). The partnership afforded Huang the opportunity to collect an experimental dataset on human visual memory with 35 million behavioural responses across 4 million experimental trials and, in doing so, to rethink the role of experiment design in studying visual memory.

In the present study[2], Huang did just this while adopting the 'scientific regret minimization' method that was first introduced in a study of moral decision-making[4].

The method begins by constructing a design space of experimental conditions: here, spatial patterns made by arranging 8 items among cells of a 6 × 6 grid. Next, 80,000 of these spatial patterns are selected as stimuli in a spatial memory task. In the task, participants briefly view one of the patterns and are then asked to recreate it from memory by selecting the corresponding locations on an empty grid. Next, a convolutional neural network is trained to take as input the pattern and, for each cell, output the probability that a person would recall that cell as having contained one of the eight items.

Finally, candidate models of visual memory are developed to account for the predictions of the trained convolutional neural network. What motivates this seemingly odd choice to use the predictions of the machine-learning model, instead of the behavioural data, as the data to guide the development of the models of visual memory? At a sufficient scale of data collection, the residuals against the predictions of a machine-learning algorithm can better approximate the mismatch with human behaviour than can the residuals against the behavioural data collected at the chosen design points[4].

Which candidate models of visual memory are then evaluated? Here we find another innovation. Huang crafted an integrative model that includes many candidate factors that have previously been considered in the literature, from so-called 'chunking' to lateral inhibition. However, the back catalogue of experimental psychologists' accountings for visual spatial memory is vast, and Huang's integration is therefore incomplete. One wonders what happened, for example, to Gestalt principles[5] such as closure, figure versus ground, or focal points. Or to temporal effects such as decay. Or interference. Or interactions with other memory systems, such as iconic or long-term memory. What is missing here, perhaps, is an ontology of the determinants of what people remember about their visual environments and a complete accounting of how the mechanisms of visual memory all fit together.

Huang's ambitious study provides a fascinating glimpse of where experimental studies in the behavioural and social sciences may be headed. Indeed, in recent years, entire fields (such as crowdsourcing[6] and human computation[7]) have sprung up and then blossomed, in part to rethink the future of science and the future of work in a world in which crowd workers, game players and web surfers can be recruited to perform short tasks that consist of as little as a single, atomic decision. Some teams have used these capacities to perform new kinds of collaborative work[8], others have scaled up traditionally structured behavioural experiments to probe how cognitive abilities change over the lifespan[9], and still others have run elaborately structured networked games[10] and experimental evolutionary simulations of social cognition[11].

Off in the distance, one sees the scientific method turned on its head. No longer do experimentalists collect bespoke datasets that are narrowly tailored to distinguish between competing hypotheses. Rather, they collect the data first and ask questions only later. In such a future, the very presence of massive, complex open datasets that serve as benchmarks of human behaviour encourages researchers to craft integrative theories and models that make concrete and accurate predictions. The experimentalist, having completed one leg of the race, passes the baton to the theorist or the modeller, praying it is not dropped. Bartlett watches on in amusement.

Jordan W. Suchow
Stevens Institute of Technology, Hoboken, NJ, USA.
e-mail: jws@stevens.edu

Published online: 23 March 2023

References

1. Bartlett, F. C. *Remembering* (Cambridge Univ. Press, 1932).
2. Huang, L. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-023-01559-z (2023).
3. Luck, S. J. & Hollingworth, A. *Visual Memory* (Oxford Univ. Press, 2008).
4. Agrawal, M., Peterson, J. C. & Griffiths, T. L. *Proc. Natl Acad. Sci. USA* **117**, 8825–8835 (2020).
5. Köhler, W. *Psychol. Forsch.* **31**, XVIII–XXX (1967).
6. Howe, J. *Wired Magazine* **14**, 1–4 (2006).
7. Law, E. & von Ahn, L. *Human Computation* (Morgan & Claypool, 2011).
8. von Ahn, L., Maurer, B., McMillen, C., Abraham, D. & Blum, M. *Science* **321**, 1465–1468 (2008).
9. Hartshorne, J. K. & Germine, L. T. *Psychol. Sci.* **26**, 433–443 (2015).
10. Rand, D. G., Arbesman, S. & Christakis, N. A. *Proc. Natl Acad. Sci. USA* **108**, 19193–19198 (2011).
11. Morgan, T. J. H., Suchow, J. W. & Griffiths, T. L. *Phil. Trans. R. Soc. Lond. B* **375**, 20190504 (2020).

Competing interests

The author declares no competing interests.