



## Should Confidence Be Trusted?

Hakwan Lau and Brian Maniscalco

*Science* **329**, 1478 (2010);

DOI: 10.1126/science.1195983

*This copy is for your personal, non-commercial use only.*

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of April 18, 2013):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/329/5998/1478.full.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/329/5998/1478.full.html#related>

This article **cites 11 articles**, 4 of which can be accessed free:

<http://www.sciencemag.org/content/329/5998/1478.full.html#ref-list-1>

This article appears in the following **subject collections**:

Neuroscience

<http://www.sciencemag.org/cgi/collection/neuroscience>

On the experimental side, quantum walks have been performed in a number of different systems, including trapped ions, atoms in optical lattices, and photons. A distinctive feature of the method described by Peruzzo *et al.* is their use of two walkers, i.e., two photons. They measured quantum correlations between the photons that are stronger than those that can be obtained with phase-averaged classical light. There have been only a few investigations of quantum walks with two walkers, and most of these have focused on the case in which the walkers can undergo statistical correlations, as in the study of Peruzzo *et al.*, but do not interact through forces (for example, as

charged particles might do). Recently, Gamble *et al.* (8) found that two interacting walkers are more successful at distinguishing non-isomorphic graphs (ones that connect vertices differently) than are noninteracting walkers.

There is still a great deal to be learned about quantum walks. For a single walker, walks on more complicated graphs or simple graphs with defects are possible areas of investigation. Walks with multiple walkers, both the noninteracting and interacting cases, are relatively unexplored. Finally, it is possible that more quantum algorithms will emerge from a better understanding of quantum walks that will enable new ways to

speed up computation.

#### References

1. A. Peruzzo, *Science* **329**, 1500 (2010).
2. D. Aharonov, A. Ambainis, J. Kempe, U. Vazirani, in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2001), pp. 50–59.
3. E. Farhi, S. Gutman, *Phys. Rev. A* **58**, 915 (1998).
4. N. Shenvi, J. Kempe, K. B. Whaley, *Phys. Rev. A* **67**, 052307 (2003).
5. M. Hillery, D. Reitzner, V. Buzek, *Phys. Rev. A* **81**, 062324 (2010).
6. A. Ambainis, *SIAM J. Comput.* **37**, 210 (2007).
7. E. Farhi, J. Goldstone, S. Gutman, *Theory Comput.* **4**, 169 (2008).
8. J. K. Gamble, M. Friesen, D. Zhou, R. Joynt, S. N. Copper-smith, <http://arxiv.org/abs/1002.3003> (2010).

10.1126/science.1195446

## NEUROSCIENCE

# Should Confidence Be Trusted?

Hakwan Lau<sup>1,2</sup> and Brian Maniscalco<sup>1</sup>

Imagine two witnesses in a courtroom. One is absolutely sure of her testimony; the other gives opposing testimony, but is less confident. Who would you trust more? All else being equal, we would tend to trust the former, because we believe that judgments made with high confidence are more accurate. This correlation between confidence and accuracy, though often true, unfortunately is not infallible. On page 1541 in this issue, Fleming *et al.* (1) report a relationship between the brain scans of people obtained by magnetic resonance imaging (MRI) and how seriously we should take their expressed level of confidence.

When analyzing how confidence predicts accuracy, it is desirable to account for the effect of “response bias.” For instance, perhaps the witness’s high confidence is driven by a brash personality rather than a genuinely accurate memory. The problem of response bias is traditionally addressed by methods such as “signal detection theory,” which is an analytic tool that allows us to separate efficacy from bias (2). By applying such techniques (3), one can characterize how well the subject’s expressed level of confidence distinguishes between correct and incorrect responses, independent of response bias. This measure of the efficacy of confidence ratings has been called “type 2 performance” to distinguish it from “type 1 performance,” which measures how accurately a subject actually identifies a stimulus.

<sup>1</sup>Department of Psychology, Columbia University, New York, NY 10027, USA. <sup>2</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands. E-mail: hakwan@psych.columbia.edu



**Confidence ratings.** Why does a witness’s expressed level of confidence, when giving testimony, affect our judgement of its accuracy?

In other words, high type 2 performance indicates a close relationship between the confidence of the subjects and how accurately they identify the stimulus.

But there is another problem: Type 2 performance is influenced by type 1 performance (3). The intuition is simple: Suppose a subject has great difficulty making accurate judgments about stimuli (such as the orientation of a figure), making many incorrect judgments as well as “fluke” judgments that are correct only by chance. The subject cannot distinguish incorrect judgments from those that are correct only by chance; they all seem like guesses. Thus, some variation in type 2 performance may be attributable merely to the quality of “lower-level” stimulus processing in the brain (i.e., type 1 performance). To isolate this confounding factor, Fleming *et al.* controlled for type 1 performance by programming a computer to give harder

The ability to monitor the efficacy of one’s own perception is associated with differences in the structure of specific brain regions.

trials to the better observers, and easier trials to the poorer observers. The authors still found substantial variability in type 2 performance across 32 observers. Also, structural MRI brain scans revealed that those observers with a high type 2 performance had higher gray matter signal intensity (which implies greater volume or density) in the frontal lobe than the low type 2 performers. The difference was most prominent in the frontal polar areas, but also was apparent in the dorsal lateral prefrontal cortex and anterior cingulate cortex. Furthermore, neuronal fibers connecting these regions showed higher signal intensity in the MRI scans.

The results speak to the debated issue (4) of whether type 2 performance reflects genuine metacognition (5)—that is, cognition about another cognitive process, rather than about an external stimulus. For instance, one may argue that confidence ratings in a perceptual task may be made by tracking the strength of the external stimulus, rather than by introspection of the efficacy of the perceptual process (metacognition). However, metacognition is one plausible way in which such confidence ratings can be made, as has been demonstrated by computational modeling (6). The correlation of type 2 performance with structures in the frontal polar region of the brain seems to support the metacognitive account, because this area is at the top of the

information-processing hierarchy, receiving input from other cognitive regions rather than early sensory areas (7).

Metacognition is perhaps particularly controversial in nonhuman animals, such as dolphins and monkeys (4, 8). Such animals can make responses that seem to reflect confidence. For instance, when given the option to abort a trial quickly instead of trying for an answer, they took the option when their accuracy was low, as if they were expressing “uncertainty,” i.e., a lack of confidence (4). Do the same brain structures identified by Fleming *et al.* govern the “uncertain responses” in these animals? There is considerable behavioral evidence in favor of the metacognitive account of uncertainty judgments for these animals (4). Or do they use non-metacognitive mechanisms to generate “uncertain responses,” thus recruiting different brain structures? It would be interesting to determine whether lesions to the prefrontal cortex would affect these responses.

One might expect type 2 performance of nonhuman animals to be considerably poorer than that of humans, because their prefrontal cortices are not as developed. One difficulty in testing this possibility is that we cannot easily control for observers’ type 1 performance across studies and species. How-

ever, the mathematical relationship between type 1 and type 2 performance has recently been mapped out (3), and a method is now available (9, 10) to estimate type 2 performance even when we cannot control for type 1 performance. Future studies can use this method to test the hypothesis that across species, or across different developmental stages in humans, type 2 performance may be correlated with structural development in prefrontal regions.

Fleming *et al.* were cautious in interpreting their results in relation to sensory awareness. Nonetheless, the close conceptual relationship between confidence and sensory awareness has been discussed for at least a century (11). Given that type 1 performance can be shown to dissociate from sensory awareness in some cases (9, 12, 13), perhaps we should not equate the two, as is commonly done (14). Rather, perhaps awareness arises when the observer’s brain introspectively “recognizes” that the perceptual signal was actually strong rather than weak, regardless of the underlying type 1 performance (6, 14, 15). Although this does not mean that sensory awareness is the same as type 2 (metacognitive) performance, both may depend on shared neural mechanisms that support the same kind of introspective monitoring of perceptual certainty.

Indeed, although the sensory signal itself may be represented by activity in posterior brain regions, visual awareness may depend on prefrontal regions similar to those reported by Fleming *et al.* (9, 12, 13, 16).

#### References and Notes

1. S. M. Fleming, R. S. Weil, Z. Nagy, R. J. Dolan, G. Rees, *Science* **329**, 1541 (2010).
2. N. A. Macmillan, C. D. Creelman, *Detection Theory: A User’s Guide* (Erlbaum, Hillsdale, NJ, 2005).
3. S. J. Galvin, J. V. Podd, V. Drga, J. Whitmore, *Psych. B. Rev.* **10**, 843 (2003).
4. J. D. Smith, *Trends Cogn. Sci.* **13**, 389 (2009).
5. J. Metcalfe, A. P. Shimamura, *Metacognition: Knowing about Knowing* (MIT Press, Cambridge, MA, 1996).
6. A. Cleeremans, B. Timmermans, A. Pasquali, *Neural Netw.* **20**, 1032 (2007).
7. N. Ramnani, A. M. Owen, *Nat. Rev. Neurosci.* **5**, 184 (2004).
8. N. Kornell, L. K. Son, H. S. Terrace, *Psychol. Sci.* **18**, 64 (2007).
9. E. Rounis, B. Maniscalco, J. C. Rothwell, R. E. Passingham, H. Lau, *Cognit. Neurosci.* **1**, 165 (2010).
10. For freely shared code online, see [www.columbia.edu/~bsm2105/type2sdt/](http://www.columbia.edu/~bsm2105/type2sdt/).
11. C. S. Peirce, *J. Jastrow, Mem. Natl. Acad. Sci.* **3**, 73 (1884).
12. H. C. Lau, R. E. Passingham, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18763 (2006).
13. L. Weiskrantz, *Consciousness Lost and Found* (Oxford Univ. Press, New York, NY, 1997).
14. H. Lau, *WIREs Cognit. Sci.* **10.1002/wcs.93** (2010).
15. H. C. Lau, *Prog. Brain Res.* **168**, 35 (2007).
16. A. Del Cul, S. Dehaene, P. Reyes, E. Bravo, A. Slachetky, *Brain* **132**, 2531 (2009).

10.1126/science.1195983

## GEOPHYSICS

# Seismic Images of the Biggest Crash on Earth

Rainer Kind and Xiaohui Yuan

In plate tectonics, the upper layer of Earth (the lithosphere) consists of rigid plates that shift over geological time scales. Data from global positioning surveys have estimated that the entire Indian subcontinent has moved over the past 50 years about 2 m to the north, diving slowly underneath Tibet. This giant collision has been ongoing for 50 million years and has thrown up the highest mountains as well as the largest and highest plateau on Earth. Not only is the world climate strongly influenced by this massive plateau (average elevation of 5000 m), but the collision also causes catastrophic earthquakes in southern, central, and eastern Asia.

There is an ongoing international effort to record seismic waves in Tibet and use them to study the deeper structure beneath Tibet. The Sino-American experiment Hi-CLIMB (Himalayan-Tibetan Continental Lithosphere during Mountain Building) (1–6), the multinational experiment INDEPTH (International Deep Profiling of Tibet and the Himalayas) (7–10), and others (11) have covered the main part of the plateau. Different seismic techniques have been combined in these studies: (i) Seismic tomography, which is sensitive to smooth variations of material properties, can locate the lithosphere and asthenosphere by their higher and lower seismic velocities, respectively (2, 12, 13); (ii) the analyses of converted waves (where the propagating seismic waves change from shear to compressional waves, or vice versa, also called receiver functions) (1,

Several large seismic experiments are providing a large-scale and detailed picture of the Indian tectonic plate diving underneath Tibet.

5–8, 10, 11) or of internally reflected waves (3), which are sensitive to sharp boundaries, can locate the crust-mantle boundary (Moho) and the lithosphere-asthenosphere boundary (LAB) with high resolution; and (iii) seismic anisotropy studies can provide an indication of mantle deformation (4, 9, 11). These new seismic studies cover the entire lithosphere, especially the LAB, whereas earlier studies focused on the more accessible crust.

The main results of the various seismic field campaigns in Tibet can be summarized graphically (see the figure). Different tomography results (2, 12, 13) indicate a broad Indian lithosphere of 100 to 200 km thickness reaching farther north than the northern boundary of the stable Indian plate at the surface (see the figure, top panel). A clear boundary between the Indian and Asian

Deutsches GeoForschungsZentrum, Telegrafenberg, 14473 Potsdam, Germany. E-mail: kind@gfz-potsdam.de