Obsessing about Uncertainty?

Jacqueline Scholl[1,*] and Matthew F.S. Rushworth[1,2]
[1]Wellcome Centre for Integrative Neuroimaging (WIN), Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK
[2]Wellcome Centre for Integrative Neuroimaging (WIN), Centre for Functional MRI of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK
*Correspondence: jacqueline.scholl@cantab.net
https://doi.org/10.1016/j.neuron.2017.09.045

A striking observation in obsessive-compulsive disorder is that patients know that their obsessions and compulsions are excessive, but their symptoms nevertheless persist. Drawing on computational models from basic neuroscience, Vaghi and colleagues (2017) suggest a quantitative account of this clinical finding.

There has been great interest in understanding the neural mechanisms mediating decision-making and the subjective sense of confidence that a decision taken is the correct one. In this issue, Vaghi and colleagues (2017) build on ideas derived from human neuroimaging, animal models, and computational studies to provide a detailed characterization of a human clinical condition: obsessive-compulsive disorder (OCD). Normally, people adjust the way in which they make decisions as a function of their confidence in those decisions. In OCD, however, this is no longer the case.

To behave adaptively, humans and animals learn from experience (for example, learning to predict where your opponent in tennis is likely to serve a ball). However, in the real world, which is changing and uncertain, this is not easy. If a serve is in a different place than expected, was this a chance event or does it indicate a change in the opponent's strategy? Inspiration for how the brain might solve this has come from Bayesian approaches to information representation (O'Reilly, 2013). Beliefs are seen as distributions, not as point estimates (Figure 1A). For example, you don't only know where the serve is most likely to land but also how likely it is to land in other places. This meta-belief, or confidence, should affect how much you update your beliefs about the most likely serving target when there is an unexpected serve. If you are not confident that your belief is true, you should be very influenced by a surprising serve, while if you are already quite certain about your belief, it is more likely that a surprising outcome is due to chance and therefore you should update your belief less.

Ideas about confidence can also be combined with canonical learning theories in which belief updates are driven by prediction errors (differences between predictions and subsequent actual outcomes). Vaghi and colleagues take such an approach (Vaghi et al., 2017). In this framework, confidence determines a learning rate that in turn determines how much impact prediction errors have on belief updates. The learning rate is faster when you are uncertain (you are more ready to change your beliefs) but slower when you are confident. This Bayesian framework has informed our understanding of how the brain computes uncertainty and uses it in diverse ways such as weighing up and integrating sources of evidence (Franklin and Wolpert, 2011) or driving exploration for information (Zajkowski et al., 2017).

In this Bayesian view, information is represented in an inherently meta-cognitive way: there are beliefs about beliefs. As a result, updating of beliefs, using Bayes' rule, automatically takes the meta-beliefs into account so that learning is informed by confidence. What Vaghi and colleagues realized, however, is that the behavior of OCD patients might provide an intriguing window into this process. An important feature of OCD is that it is *ego-dystonic*: patients know their obsessions (e.g., germ contamination) and compulsions (e.g., hand washing) are excessive, but their symptoms nevertheless persist. Vaghi and colleagues therefore examined the extent to which OCD patients hold metacognitions about their beliefs and the degree to which they influence adaptation of behavior, i.e., behavioral measures of learning. In a computer game task, participants learned where a "particle" was likely to land so they could place a bucket to catch it (a similar problem to the one confronted by the tennis player trying to predict where the opponent will serve). They then rated their confidence in how well they had placed the bucket. Just like the tennis serves, the particles were likely to land close to one another for a while (following a normal distribution with some random noise), but sometimes a "change point" occurred: the most likely particle position changed just as a tennis player might change her tactics. Uncertainty varied throughout the task: immediately after change points, uncertainty was high. Then, as participants observed more and more particles in the same vicinity, uncertainty decreased. Importantly, two measurements were made on every trial: explicit report of the participants' confidence in their bucket positioning (prior to outcome observation) and the actual bucket position participants chose. Trial-by-trial changes in bucket position could then be used to estimate behavioral change or learning rate. As expected, healthy controls reported both higher confidence and exhibited lower learning rates the more samples they had seen since a change point. However, a striking dissociation between confidence and behavioral adaptation was revealed in OCD. OCD patients did not differ in their explicit confidence ratings from healthy controls. However, they did not adapt their behavioral learning rates in the same way. They made large behavioral adjustments even when their explicit reports had indicated high confidence. A similar result was found by analyzing the data more formally

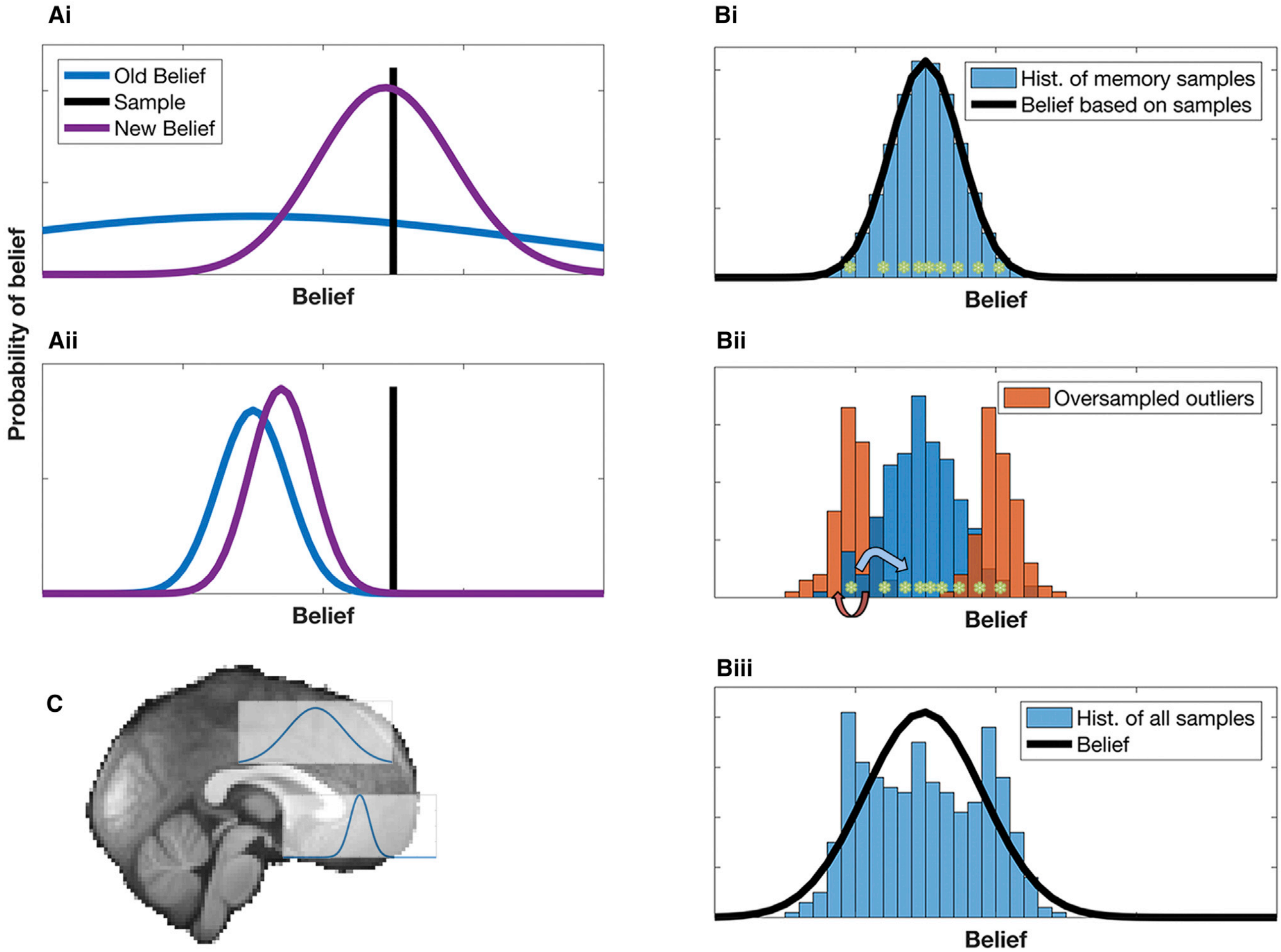

Figure 1. Bayesian Framework for Representing Beliefs

(A) In a Bayesian framework, beliefs are represented as probability distributions (blue). That is, a probability is assigned to each possible belief. The uncertainty of a belief (i.e., how wide the distribution is) is important for determining the impact of new evidence in updating beliefs. If the belief is uncertain or broad (Ai), a new piece of evidence (black) has a strong impact on the belief (purple). In contrast, if the belief has a high certainty (Aii), a new piece of evidence only has a small impact on the belief.

(B) One way the brain might create a distribution of beliefs is by drawing multiple (noisy) samples from memory (for example, at each of the green stars). The resulting belief distribution is indicated by the histogram and fitted line in black. If the sampling is proportional to the distribution of memories (Bi), then the resulting distribution will be correct. (Bii) However, if there is potentially a bias to over-sample certain memories (red arrow)—for example, because they are particularly salient—before also sampling the correct distribution (blue arrow), the resultant combined distribution will be skewed (Biii). Different ways of probing the distributions, such as asking participants for explicit reports or requiring them to implicitly use the knowledge to guide behavior, may lead to the distribution being sampled in different ways.

(C) There is evidence that there are parallel systems in the brain that represent partially overlapping information. In this framework, it is possible that there is a problem with a distribution in one brain area, without this affecting the representation in other areas. Information from different areas might be read out depending on how participants are probed.

by regressing measures of uncertainty derived from a Bayesian model against either type of measure. To test directly whether participants used their belief about uncertainty as reported in the ratings to guide behavior, Vaghi and colleagues measured the impact of these ratings on behavioral adaptation. For control participants, there was a clear relationship between the reported uncertainty and the trial-by-trial behavioral learning rates. However, in OCD there was a disjunction between the possession of intellectual knowledge and its use in behavioral adaptation.

The study is an exemplary demonstration of the increasingly popular computational psychiatry approach: applying knowledge about cognitive and neural processes from basic research to psychiatric disorders. This leads to objective and quantifiable accounts of a disorder that go beyond symptoms, and instead tackle the fundamental mechanistic changes associated with the disorder, sometimes even revealing aspects of a disorder that were not previously captured by clinical intuition. The finding here of a difference between intellectual knowledge and use of the knowledge to guide behavioral adaptation is in line with previous experimental findings that OCD patients do not learn in the same manner as healthy controls (Gillan and Robbins, 2014), regardless of whether or not the learning task concerns the content of their obsessional thoughts. However, this emerging picture of OCD contrasts with the description of OCD in the DSM-V clinical manual that sees obsessions as primary and compulsions simply as consequences—something patients do to reduce the distress caused by the obsessive thought. These findings thus highlight how careful behavioral investigation can reveal new insights into the nature of even well-known conditions that could not have been gleaned from patients' subjective reports given in the clinic. For the future, it will be interesting to explore how this new understanding could inform treatments for OCD. If patients' problems reflect a divergence between uncertainty estimates in different representations—an accurate representation of uncertainty that can be read out by explicit report and an inaccurate one that guides behavioral change—then a psychological therapy that "synchronizes" these representations might be beneficial. Notably, the focus of mindfulness-based treatments for other psychological illnesses such as depression is on transforming the way in which intellectual knowledge is linked to actual behavior (Teasdale, 1999).

Stepping back from OCD, however, the results have implications for basic cognitive neuroscience because they shed light on how Bayesian beliefs might be stored in the brain. The results suggest the brain may not (or not only) represent beliefs as explicit probability distributions because if this were the case, then updating should inevitably take into account uncertainty. This is not the case in OCD, yet OCD patients have a sense of uncertainty. One way in which such a finding might be explained is if the brain stores, as heuristics, point estimates of both the belief and the uncertainty; the quasi-optimal Bayesian learning model that Vaghi and colleagues employ actually operates in a similar manner. In this scenario, the deficit in OCD simply results from a problem in making the right information accessible for actions.

In addition, as proposed more recently, brains might be better described as "Bayesian samplers," rather than representing explicit probability distributions (Bornstein and Norman, 2017; Sanborn and Chater, 2016). According to this view, the brain represents a probability distribution when needed by "drawing samples" from memory (Figure 1B). Importantly, how exactly the memory is probed could change what is retrieved. When retrieving the distribution to make a confidence judgement, patients might sample the distribution correctly, leading to a correct confidence judgment. Similarly, they can report that their obsessive beliefs are inappropriate. In contrast, when sampling the distribution to generate action, OCD patients might show a bias in sampling; for example, showing "stickiness" in their sampling and repeatedly sampling the same unlikely (outlier) memory again and again (Figure 1B). This would be analogous to having obsessive thoughts that return to the same content again and again.

Given that there are parallel systems for decision-making, another possibility is that different systems may have independent representations of uncertainty (Figure 1C). For example, during decision-making some brain structures such as anterior cingulate cortex simultaneously hold multiple estimates of response evidence that are based on different timescales of experience (Meder et al., 2017; Wittmann et al., 2016). Co-activation of such representations may provide one substrate for an estimate of uncertainty. Other brain structures, such as ventromedial prefrontal cortex, that possess quite different representations of response evidence, may construct uncertainty estimates on the basis of a quite distinct process. Changes in just one system's representation, while another system remains intact, might underlie the diverging uncertainty effects identified by Vaghi and colleagues and some of the striking features of OCD.

ACKNOWLEDGMENTS

We would like to thank Nils Kolling for helpful discussions of this manuscript. This work was funded by the MRC (J.S. holds a Skills Development Fellowship, MR/N014448/1) and the Wellcome Trust (M.F.S.R. holds a senior investigator award, WT100973AIA).

REFERENCES

Bornstein, A.M., and Norman, K.A. (2017). Nat. Neurosci. *20*, 997–1003.

Franklin, D.W., and Wolpert, D.M. (2011). Neuron *72*, 425–442.

Gillan, C.M., and Robbins, T.W. (2014). Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, http://dx.doi.org/10.1098/rstb.2013.0475.

Meder, D., Kolling, N., Verhagen, L., Wittmann, M., Scholl, J., Madsen, K.H., Hulme, O., Behrens, T., and Rushworth, M. (2017). bioRxiv. http://dx.doi.org/10.1101/195842.

O'Reilly, J.X. (2013). Front. Neurosci. *7*, 105.

Sanborn, A.N., and Chater, N. (2016). Trends Cogn. Sci. *20*, 883–893.

Teasdale, J.D. (1999). Clin. Psychol. Psychother. *6*, 146–155.

Vaghi, F., Luyckx, F., Sule, A., Fineberg, F., Robbins, T.W., and De Martino, B. (2017). Neuron *96*, this issue, 348–354.

Wittmann, M.K., Kolling, N., Akaishi, R., Chau, B.K., Brown, J.W., Nelissen, N., and Rushworth, M.F. (2016). Nat. Commun. *7*, 12327.

Zajkowski, W.K., Kossut, M., and Wilson, R.C. (2017). eLife. Published online September 15, 2017. http://dx.doi.org/10.7554/eLife.27430.