Precision and the Bayesian brain

Daniel Yon[1,2] and Chris D Frith[3,4]

Posted to *PsyArXiv* on 21/07/21. Accepted for publication in *Current Biology.*

Scientific thinking about the minds of humans and other animals has been transformed by the idea that the brain is *Bayesian*. A cornerstone of this idea is that agents set the balance between prior knowledge and incoming evidence based on how reliable or 'precise' these different sources of information are — lending the most weight to that which is most reliable. This concept of precision has crept into several branches of cognitive science and is a lynchpin of emerging ideas in computational psychiatry — where unusual beliefs or experiences are explained as abnormalities in how the brain estimates precision. But what precisely is *precision*? In this Primer we explain how precision has found its way into classic and contemporary models of perception, learning, self-awareness, and social interaction. We also chart how ideas around precision are beginning to change in radical ways, meaning we must get more precise about how precision works.

Precise and imprecise percepts

Imagine you are walking a particularly disobedient dog. After being let off the leash he leaps into the bushes, and you have to dive in to fetch him out. But you are not entirely sure where he is. You hear the sound of twigs cracking to the left, but you see the leaves shake to the right. Where should you jump in to catch him?

Locating your dog based on a combination of sight and sound is an example of a general class of *multisensory integration* problems where our perceptual systems have to triangulate different sensory signals. In our example, the visual signal (shaking leaves to the right) and the auditory signal (cracking twigs to the left) both tell us something about one feature of the environment (the dog's location), and so it makes sense to combine them. But how? A simple approach could be for our brain to average them together — if sight says right and sound says left, we should dive in straight ahead.

But simple averaging turns out to be suboptimal when some signals are more reliable than others. For example, the spatial acuity of vision is much greater than that of hearing, meaning visual estimates of location are considerably more precise than auditory ones. This insight was formalised in Marc Ernst and Martin Banks' Bayesian model of multisensory integration, which assumes that our perceptual systems combine different signals according to their reliability or uncertainty. Agents are thought to achieve this by keeping track of the noise or variance in different sensory modalities — with low noise taken as an index of high precision — and affording a higher weight to those channels that are more precise.

This idea of 'precision-weighting' provides a good account of near-optimal cue integration seen in humans and other animals. Typically, we won't jump straight to catch the dog, but will veer off to the right as our brains give more credence to the more precise visual signal. Importantly, this idea can also explain why perception sometimes errs: such as when we are fooled by a ventriloquist's dummy. It is common to say that the ventriloquist 'throws their voice' so it appears to be coming from the silent puppet. In fact, the illusion of the speaking doll emerges because our perceptual systems infer that the visual and auditory signals come from a common source, but give more weight to what we *see* than what we *hear* as we try to pinpoint where this source is located. The perceptual experience is false — the voice is not coming from the dummy — but this can still be thought of as an optimal inference from the brain's perspective, given that coincident sensory signals often do come from a common source, and visual information about the location of these sources is typically so much more precise.

Precision and prediction

So far, we have considered how a system might precision-weight different sources of incoming information. However, influential Bayesian models of perception suggest that our experience of the world around us is constructed by combining bottom-up sensory signals with top-down expectations about what the world contains. For example, if we are trekking through the desert we are much more likely to encounter a camel than a polar bear, and our

sensory systems can make use of this knowledge by biasing perceptual inferences towards the most probable alternatives (see Figure 1).

But how much should we rely on our predictions and how much on evidence? This combination problem can also be optimised by precision-weighting: giving relatively more weight to bottom-up evidence or top-down predictions when one is relatively more precise. One upshot of this idea is that we should depend more strongly on our prior beliefs when the sensory world is most ambiguous. If a blustering sandstorm disrupts our desert trek, and the swirling sand makes it difficult to identify a creature in the distance, it is optimal for our perceptual system to rely strongly on our prediction that it's a camel — as relying instead on the noisy and indecisive sensory evidence will only corrupt our perceptual inferences. But if viewing conditions are perfect we should trust the reliable evidence from our senses and minimise any contribution of top-down knowledge.

It is, however, also important that we estimate the precision of our predictions. This can be achieved by tracking the stability of our environment and the degree to which the present resembles the past. At the beginning of our desert trek, we may have very narrow predictions about the kinds of animals we will encounter (for example, camel yes, polar bear no). Every camel we encounter on our trip will strengthen this prediction, furnishing in turn a stronger belief about the predictability of the desert landscape. Encountering a single surprising polar bear, however, will cause us to update our predictions and also alter our more global belief about the stability of the environment: perhaps the next time we see a creature on the horizon we should expect to encounter something unexpected — like a penguin or a chimpanzee? As we entertain a wider range of possibilities our top-down predictions become less *precise* and we should give relatively more weight to bottom-up signals.

A long-standing hypothesis suggests that our brain uses specific neuromodulators to achieve this delicate weighting, altering the synaptic gain afforded to top-down predictions and bottom-up evidence based on how precise they are estimated to be. Indeed, recent work by Rebecca Lawson and colleagues has combined drug interventions with

computational modelling to test and support this conjecture. The authors focused on noradrenaline, a neuromodulator previously implicated in signalling the volatility of the world around us, and thus our relative (in)ability to make predictions about what will happen next. In unstable environments our current beliefs are poor predictors of what will happen in the future. This makes volatility a form of second-order precision estimate, reflecting the reliability of our expectations. One hypothesis suggests that, when we estimate our environment to be more volatile, noradrenergic neuromodulation increases the gain (signal-to-noise) of incoming signals. This has the effect of upweighting incoming information — enabling rapid learning about potential changes — while also downweighting the impact old expectations have on inferences, since predictions are less reliable in a world that changes often.

Lawson and colleagues investigated this hypothesis in an experiment where volunteers made perceptual decisions about noisy visual stimuli — is that picture a face or a house? — while tracking fluctuating probabilistic cues that predicted what the upcoming stimulus would be. Importantly, one group of participants completed the task after taking propranolol, a beta-blocker which antagonises the noradrenaline system. If noradrenaline does indeed encode environmental volatility, supressing this system should artificially enhance the apparent predictability of the environment — causing agents to give more weight to top-down expectations (and equivalently less weight to incoming signals). In line with this idea, the authors found that those who receive the drug rely more heavily on their expectations when making perceptual judgements and are slower to update these predictions in the face of new evidence — as though they believe their models of the environment are especially reliable or *precise*.

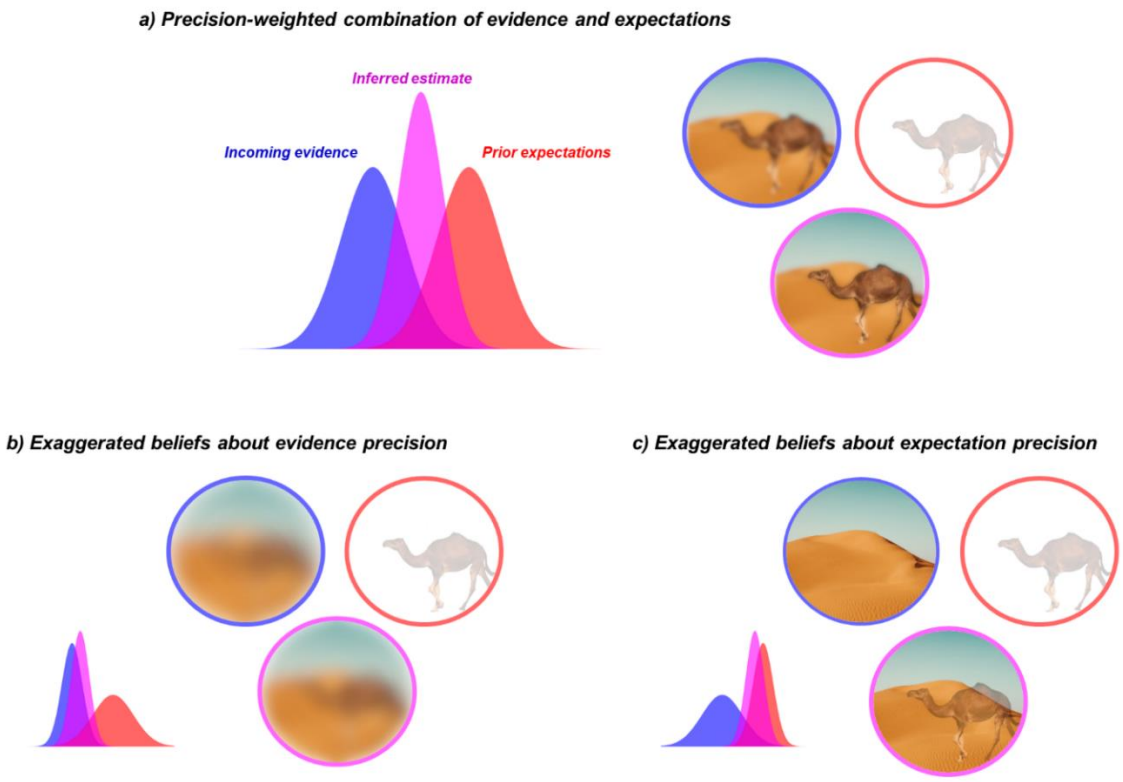

Figure 1. Precision-weighted inference and 'false beliefs' about precision.

(A) In an ambiguous sensory environment (e.g., a desert in a sandstorm), you can generate more reliable inferences by incorporating prior knowledge into perception (e.g., that shape is probably a camel). The weight given to these two sources of information depends on how reliable or 'precise' we estimate them to be. Recent predictive processing models suggest agents can entertain false beliefs about precision that can lead to maladaptive inferences (see main text 'The changing face of precision'). For example, an exaggerated belief in the reliability of incoming evidence (B) can leave agents with noisy inferences when signals are actually weak or ambiguous. In contrast, exaggerated beliefs about the reliability of predictions (C) may lead agents to falsely infer the presence of expected but absent events (i.e. hallucinate).

Some prediction errors are more important than others

A complementary perspective on the relationship between precision and learning is offered by research on reward. Prevailing models of reward learning suggest that humans and other animals form and update their beliefs about valuable outcomes by tracking prediction errors.

If you find yourself in a grimy cafe where the tables are dirty and the staff are rude you may not expect to enjoy your cappuccino very much — but if you receive a delicious coffee you will be pleasantly surprised and may use this positive prediction error to update your future beliefs (for example, you might visit the café again).

Over the past few decades, neuroscientists have found that the reward prediction error signals implied by these models are readily detected in the dopaminergic midbrain and striatum of humans and other animals. But it has only more recently been appreciated that not all error signals are created equal.

The key insight here is that the world is a stochastic place, and fluctuations in our environment from time-to-time do not always signal a need to change our minds. For example, if we visit our new favourite café and are served a foul cup of coffee, we are likely to be disappointed — but does this negative prediction error represent a reliable change, meaning we should never come back?

Contemporary accounts of reward learning suggest this kind of puzzle can be solved by scaling prediction error signals according to our uncertainty about the distribution they come from. This means that agents must represent not only the outcome they expect *on average,* but also the variance in the outcomes they are likely to encounter. The logic here is that when our environments are less variable, and thus our expectations are more precise, a small deviation from our predictions might signal an important change: if the coffee has been consistently good every day for years, the single disappointing cup is newsworthy, and might lead you to think they have moved to a worse supplier or fired the best barista. In contrast, if our environment is more variable — sometimes the coffee is excellent, sometimes it is average — small fluctuations should not surprise us. This is just noise. It is not informative.

Evidence has begun to emerge that prediction errors are indeed precision-weighted in this way. In one indicative study, Kelly Diederen and colleagues examined how neural signatures of reward prediction error changed when values were more or less certain. Volunteers had their brain activity recorded in an MRI scanner while they completed a reward learning task: on every trial predicting how much money they would receive from a

probabilistic lottery. Different lotteries could have the same expected pay-out *on average* but have wider or narrower distributions of possible pay-outs. Consistent with previous work, the authors found reward prediction error signals can be detected in the midbrain and the striatum — signalling whether a lottery pay-out was better or worse than expected. Crucially though, these error signals were augmented when lotteries were reliable (high precision) and attenuated when they were more uncertain (low precision) — and differences in this level of neural adaptation correlated with task performance.

This precision-weighting of prediction errors also appears to depend on specific neuromodulators. If participants complete the same task after taking sulpiride — a drug that antagonises dopaminergic function — prediction error signals are still elicited in the same brain regions when the lottery gives an unexpected pay out. However, the adaptive scaling is markedly attenuated, and similar error signals are elicited from certain and uncertain lotteries. This suggests that — under dopamine blockade — agents are less able to incorporate information about the reliability or precision of their predictions into the signals they use for learning.

There is an intriguing contrast between these studies looking at the role of dopamine in reward learning, and the work mentioned above examining noradrenaline's contribution to sensory prediction. In the sensory prediction case, noradrenaline appears to encode the reliability of predictions — and more precise predictions lead to slower learning and a stronger reliance on expectations when making perceptual decisions. In the reward prediction case, however, dopamine also appeared to play a key role in enabling agents to track the reliability of the environment. When agents could make more precise predictions, unexpected outcomes elicited stronger error signals — leading to more rapid belief updating away from the prior rather than inferences biased towards it.

It is possible this disjunct reflects a difference between domains: perhaps precision-weighting optimises prediction and learning about perceptual signals and rewards in fundamentally different ways. However, another tantalising possibility is that agents separately estimate their uncertainty about *how the world* is and uncertainty about their

models of it. These two kinds of uncertainty may appear to go hand-in-hand, but can in principle decouple. For example, if you toss a fair coin you know perfectly well the outcomes that can ensue even if it is impossible to predict whether a specific toss will yield heads or tails. Thus, we are certain in our model of how the coin works, even if this model entails we should be uncertain about the outcomes, and these two kinds of uncertainty may be encoded at different levels of our cognitive hierarchy (see Figure 2).

Knowing the reliability of ourselves and others

The foregoing discussion illustrates how uncertainty or precision plays a key role in what are typically thought of as low-level processes like perception and reward. It is possible that the uncertainty involved in these cases is being computed by 'subpersonal mechanisms' — mental processes that operate beneath the level of subjective awareness. For example, while there is good evidence that our brain estimates the precision of incoming signals when combining data across our senses, or when combining inputs and expectations, we are typically only aware of the resulting combined percept rather than the tacit uncertainty estimates involved in producing it.

Uncertainty estimation is, however, also at the core of high-level mental abilities such as metacognition: our ability to subjectively monitor the reliability of our own minds. A paradigmatic example of metacognition is the explicit feeling of confidence we have in our decisions, and these confidence computations enable particularly sophisticated forms of cognitive control — allowing us to slow down, gather more information or seek advice when we are uncertain.

Growing evidence suggests that feelings of subjective confidence are generated by mechanisms that 'read out' the precision of other representations. One important line of work comes from the lab of Janneke Jehee, who have developed a pioneering multivariate modelling approach for neuroimaging data that measures the objective uncertainty in neural representations. In a recent study, Laura Geurts and colleagues used magnetic resonance imaging (MRI) to record the brain activity of observers while they made perceptual

judgements about tilted visual patterns, recording explicit confidence ratings about their choices. Using their innovative analysis approach, the authors could decode trial-by-trial variability in the population responses evoked by these stimuli across early visual cortex. Critically, they found that this objective measure of 'neural precision' predicts an observer's subjective confidence rating on that trial — consistent with the idea that metacognitive mechanisms directly track the reliability or precision of other mental states.

A role for precision estimation in metacognition in turn suggests a role for precision in social cognition, as confidence plays a key role in how we interact with others. Consider collective decisions: if we're baking a cake and I think we need a tablespoon of salt and you think we only need a pinch, how much should we add? It may not be a good idea just to split the difference. Bahador Bahrami and colleagues have noted that this problem of combining information across different minds shares a similar structure to the problem a single agent faces when combining signals from different sensory modalities. Indeed, like the multisensory integration case, combined estimates across multiple agents are most accurate when we give more weight to the most reliable opinions. If you are certain that we only need a pinch of salt, and I'm merely guessing that we need a spoonful, we should give more weight to your estimate in our combined decision.

But while we can monitor the uncertainty in our own beliefs, we do not have direct access to uncertainty in the minds of others. Thus, for this kind of precision-weighted combination to work we must explicitly communicate our confidence with others to work out whose opinion to weight most. Experimental work shows that when agents share their confidence — rather than just their beliefs — collective decisions can indeed become more reliable, and this suprapersonal information integration is well fit by the exact same kinds of precision-weighting models developed in multisensory cue combination.

The changing face of precision

Precision thus seems central to how psychologists and neuroscientists think about perception, learning, metacognition and social interaction (see Figure 3). However, the

concept of ‘precision’ in cognitive science has begun to change in radical ways. In particular, influential predictive processing models suggest that all aspects of perception, action and cognition can be explained within a single computational framework. Unlike classic models, these new theories suggest that our beliefs about precision can become divorced from objective reality. Under this way of thinking, agents might believe that they can make ‘precise predictions’ about events that are objectively unpredictable or believe their senses provide ‘precise evidence’ even when signals are corrupted or the system is noisy.

Entertaining nonveridical beliefs about precision in this way is key to how predictive processing models account for functions like attention. Psychologists have classically distinguished exogenous attention, where perceptual resources are captured by strong, salient events in the sensory world, from endogenous attention, where we strategically deploy resources to upweight less salient events that are relevant to our current goals. Bayesian models of attention elide this distinction by suggesting that all kinds of attention reduce to *precision*. When a bright, loud or otherwise strong event appears in our environment, such signals do indeed provide *precise* (high signal-to-noise) information that should be weighted accordingly. But this is not true for the non-salient, task-relevant information that guides endogenous attention — when you start looking for your keys, their jangle does not become physically louder and their metal surface does not become physically brighter. However, Bayesian models suggest that endogenous attention can be explained by assuming that agents entertain fictitious (counterfactual) beliefs about sensory precision — *as if* useful information in the world really is brighter and louder – rendering us particularly sensitive to the right kind of incoming signal.

These models also suggest that false beliefs about precision are key to explaining action. Bayesian models of ‘active inference’ suggest that actions are generated by strong self-fulfilling predictions. In these accounts, our motor systems are thought to generate counterfactual predictions about the states of our bodies, for example, predicting that your hand *is* grasping a door handle, even if it is currently in your pocket. These counterfactual signals generate a prediction error because there is a mismatch between the actual and

predicted state of your body. Peripheral reflexes are thus engaged to minimise the error, reconfiguring the body in line with the prediction so that the expectation comes true.

Curiously, for this scheme to work agents must hold onto the counterfactual prediction even when it conflicts with incoming evidence – you must really believe your hand is grasping the door handle, since revising this belief to match the actual evidence from your senses (e.g., you feel your hand truly remains in your pocket) would abolish the error signal needed to drive the reflexes. Active inference models suggest this problem can be finessed by assigning fictitiously high precision to these imperative predictions and corresponding low precision to incoming sensory signals — ignoring the true state of our body so that we can retain a false belief that it is somewhere else (though this idea remains controversial, see Yon *et al.* 2020).

Divorcing beliefs about precision from reality also gives predictive processing accounts enormous scope to model cognition in health and disease. For example, the hallucinations that characterise illnesses like psychosis can be cast as an 'optimal inference' given overly-strong beliefs about the reliability of our expectations (see Figure 1C). Conversely, characteristics of autism, such as a preference for stable and repetitive environments, can be cast as a consequence of overly-strong beliefs about the precision of incoming evidence, such that every fluctuation in our sensory systems seems to signal the need to change our models of the environment (and the world thus seems unstable).

The success of this kind of unshackled precision-weighting is unsurprising to many who have developed these models. Some note a mathematical proof known as the *complete class theorem,* which guarantees that it is always possible to specify a set of beliefs (including beliefs about precision) that would make the behaviour of a participant in an experiment or a patient in the clinic seem 'optimal'. While this may be good news for modellers, it presents a problem for experimentalists, as any empirical result is compatible with the framework.

With this in mind, we believe that it is essential for cognitive scientists to develop mechanistic hypotheses that are precise about why, where and when precision estimation

may go awry. Below we outline one possibility: that beliefs about precision at higher levels of a cognitive system (e.g., metacognition) are more likely to be inaccurate than precision estimates at lower levels (e.g., perception).

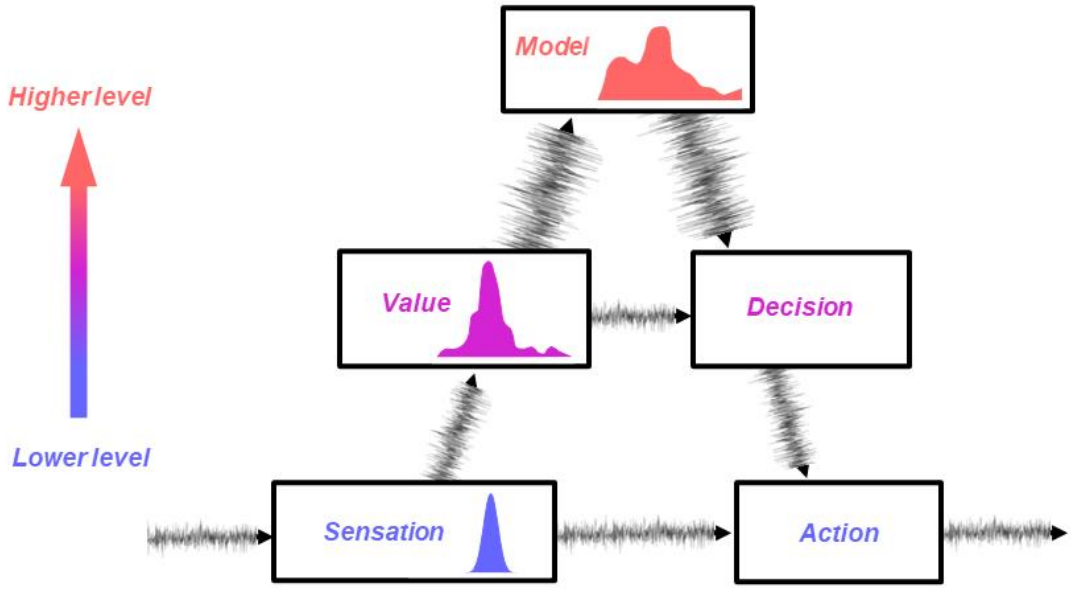

Figure 2. Varieties of precision representation.

Agents can estimate the precision of different cognitive states by tracking properties of their internal representations, such as their variance. For example, if a population of neural units encodes some property of the outside world our 'best guess' may be represented by the mean of this distribution, while the precision of this estimate is reflected in the variance (how much the units 'agree'). The significance of a precision estimate depends on what the distribution encodes. In sensory systems where units are tuned to different perceptual features, higher precision may be associated with stronger incoming signals and reduced environmental noise. In reward circuits where units encode expected value, higher precision indicates a narrower range of predicted outcomes. At higher levels of abstraction, the same statistical quantity can convey more complex information. For example, if a distribution forms a 'contingency space', with each unit representing a possible relationship between different events in our environment, variance reflects how certain we are that our current models of the world are correct.

Imprecision in high level cognition

Recent Bayesian models suggest that meta-level systems generate feelings of confidence via 'second order inference', independently integrating evidence that is also available to other low-level systems. Empirically, metacognitive introspection is often suboptimal such that confidence reports are less accurate than first order decisions. These models account for this information loss by assuming that higher-level systems are corrupted by an independent source of noise that may be greater than that afflicting lower levels.

Such noise could in principle have a physiological origin (for example, leakier neural circuits) but may also arise because of the kinds of computations performed at higher levels. If a system takes many noisy but unrelated signals as inputs, the uncertainty in its outputs can grow multiplicatively. For example, when computing confidence in a decision we might want to take account of multiple independent sources of information — such as the difficulty of the choice and the quality of the evidence we received — but if both our 'difficulty' signal and our 'evidence' signal are corrupted, our resulting confidence estimate can be even noisier than either input was to begin with. Because higher level computations (such as metacognition) are more likely to draw on diverse inputs, relative to lower level systems (such as perception), this kind of multiplicative noise may particularly interfere with high level precision estimates.

Even if higher cognitive levels are not intrinsically noisier, however, there may be other limits on their fidelity. As they ascend through the hierarchy of the mind and brain, representations abstract over increasingly diverse inputs to support global aspects of cognition. By definition, global estimates require more information, often integrated over longer timescales. For example, we need to experience more 'data-points' to estimate the stability of our environment than we need to estimate what is occurring right now. The increasing input dimensionality at higher levels presents a computational challenge, as the resulting internal models can be too complex to perform straightforward operations with (for example, the *Sensation* distribution in Figure 2 can easily be approximated by computable summary statistics like mean and variance, while the *Model* distribution cannot). As a result,

it may become intractable for high levels of a cognitive system to integrate all the evidence potentially available at lower levels.

In the face of such 'information overload' agents may instead rely on heuristic computations that only integrate a subset of the information at hand. There is evidence for this kind of data-reduction in metacognition, because confidence estimates are sometimes best modelled by assuming that agents discard potentially useful information. For example, an effective way to compute decision confidence could be to compare our uncertainty about the options we have chosen *versus* the options we did not — but agents have been found to ignore the latter, relying more on the absolute evidence they have for the option they picked when forming a confidence estimate.

In general, "coarse-graining" high-level representations can be adaptive, as it permits more effective transmission of macroscopic information within and between minds. For example, our internal model of how to behave on a first date could contain fine-grained details (for example, "don't talk about your first ex-girlfriend, or your second, or your third…") or contain a coarser policy (for example, "don't talk about your past relationships") that supports more efficient action planning and is easier to communicate to others. However, limiting and coarse-graining the data transmitted to higher levels also makes it more likely representations will depart from objective reality, enabling false beliefs about precision to emerge.

High-level precision estimates may also be strongly shaped by expectations. Bayesian accounts argue it is rational to rely on prior knowledge when estimating the precision or accuracy of our cognitive systems. For example, if I buy a new pair of glasses with a stronger prescription, I may expect my visual system to provide more accurate information — inflating my estimates of sensory precision. Relying on prior knowledge will lead to inaccurate estimates when expectations and reality diverge — if I pick up the wrong pair of glasses, I may expect to see more clearly, but become more myopic than before. If high-level systems (like metacognition) receive noisier inputs, precision estimates at these levels may be strongly swayed by expectations (see Figure 1) while low-level systems (like

perception) may remain closely tied to the true reliability of incoming signals. Thus, I may have the false metacognitive belief that my vision has improved, even if my perceptual systems are not misled. Indeed, from a Bayesian point of view, a stronger reliance on prior beliefs also entails a stubborn insensitivity to environmental feedback. If the signals reaching metacognitive mechanisms are relatively noisier, existing beliefs in these systems will be updated more slowly, allowing false beliefs about precision to persist even when we encounter contradictory evidence.

A final reason to suspect that high-level precision estimates are less veridical concerns the social role these mechanisms play. As noted above, explicit metacognition allows us to share our confidence with others, and in group decisions we give more weight to those expressing greater certainty. But agents can strategically distort the confidence they express: if we are being ignored, exaggerating confidence secures the attention of others, while expressing caution protects our reputation when we already hold high status. Agents could maintain separate representations of 'private' and 'public' confidence, but if we track our own actions to infer how confident we should be, a habit for exaggerating our confidence to others may ultimately bias how we represent the reliability of our cognitions to ourselves.

Explicit communication about certainty and uncertainty also makes it possible for us to quickly acquire beliefs about precision in the absence of direct experience. This kind of explicit learning can be very useful: if we receive trustworthy gossip that a colleague is unreliable or that a funding body makes decisions at random, we can acquire accurate beliefs about these features of the world without having to experience these disappointments firsthand. There is, however, evidence that false beliefs acquired from reliable sources can be particularly resistant to change. For example, when volunteers play economic trust games and receive gossip that they have a trustworthy partner (someone predictably cooperative), they fail to revise this initial belief when their expectations are betrayed. This slowed learning is accompanied by attenuated midbrain signatures of prediction error — consistent with the idea that the precision of prediction errors has been down weighted relative to the precision of the belief about trustworthiness. Thus, while explicit

communication can provide a fast track to good beliefs, quickly acquiring beliefs about precision from others may make us prone to persistent error - especially if we have learned to trust the wrong kind of gossip.

Conclusion

Here we have seen how estimating uncertainty is central to the way cognitive scientists currently think about many aspects of the mind. Prevailing models that explain a host of cognitive abilities — including perception, learning, metacognition and social cognition — assume that our success critically depends on our ability to track the precision of the beliefs we hold and the reliability of the evidence that we receive.

While psychologists and neuroscientists have begun to unpick the computational and neural mechanisms that support precision-weighted inference across diverse domains, there remain open questions about how precision works. In particular, while recent models suggest that our beliefs about precision can decouple from reality, they are currently silent on how this decoupling occurs. Here we have sketched out one hypothesis that could explain why precision estimates at higher levels of a cognitive system become inaccurate, and thus diverge from the ground truth. This conjecture may prove false, but nonetheless the need will remain for mechanistic hypotheses that get more precise about how precision is estimated in the mind and brain.

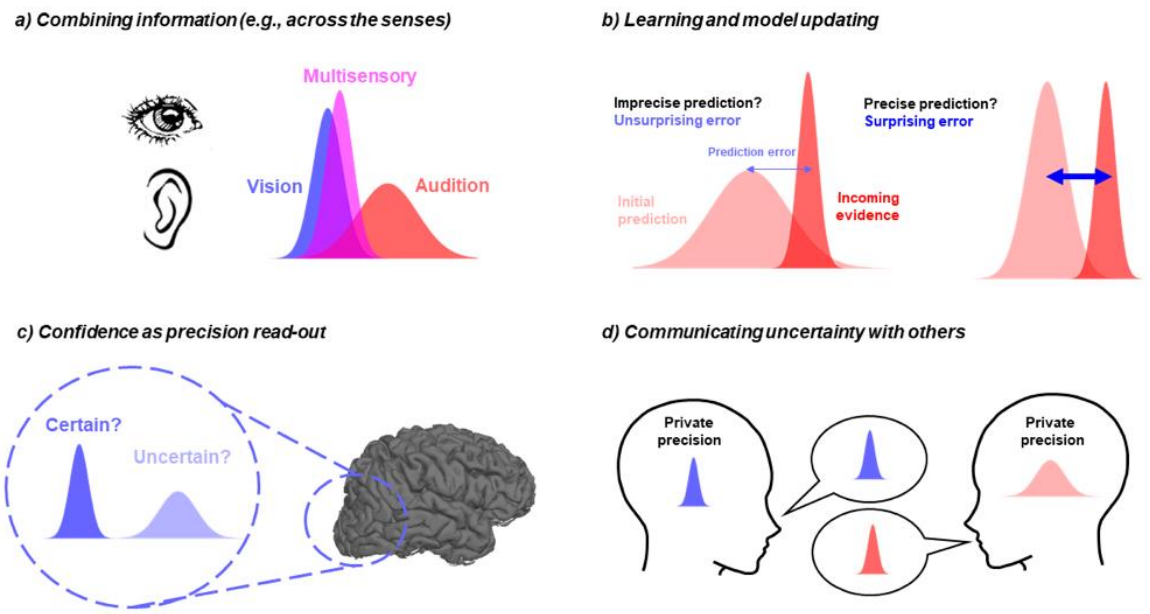

Figure 3. Four of the many faces of precision.

The concept of precision is central to current thinking about the mind and brain in a variety of domains. (A**)** Models of precision-weighted inference suggest that agents combine different sources of information (for example, signals from different senses) according to their estimated reliability or precision. This means estimates are biased toward more precise sources (for example, vision tends to dominate spatial perception because the spatial precision of vision tends to be higher). (B**)** Models of learning suggest that we update our beliefs about our environment using 'prediction error' signals that express the mismatch between our initial expectations and the evidence we sample. However, it is crucial to estimate the precision of our predictions to establish how surprised we should be. (C**)** Bayesian models of metacognition suggest that feelings of subjective confidence are constructed by 'reading out' the precision of relevant low-level representations (for example, our certainty about our visual percepts reflects a read out of the noise in our visual system). (D**)** This kind of explicit metacognition allows us to communicate our uncertainty with others, allowing us to share precision across individual minds to optimise group decisions (or strategically influence others).

Further reading

Aitchison, L., Bang, D., Bahrami, B., and Latham, P.E. (2015). Doubly Bayesian analysis of confidence in perceptual decision-making. PLOS Comput. Biol. *11*, e1004519. DOI:10.1371/journal.pcbi.1004519

Bahrami, B., Olsen, K., Latham, P.E., Roepstorff, A., Rees, G., and Frith, C.D. (2010). Optimally interacting minds. Science *329*, 1081–1085. DOI: 10.1126/science.1185718

van Bergen, R.S., Ji Ma, W., Pratte, M.S., and Jehee, J.F.M. (2015). Sensory uncertainty decoded from visual cortex predicts behavior. Nat. Neurosci. *18*, 1728–1730. DOI: 10.1038/nn.4150

Corlett, P.R., Horga, G., Fletcher, P.C., Alderson-Day, B., Schmack, K., and Powers, A.R. (2019). Hallucinations and strong priors. Trends Cogn. Sci. *23*, 114–127. DOI: 10.1016/j.tics.2018.12.001

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. Nat. Neurosci. *8*, 1611–1618. DOI: 10.1038/nn1575

Diederen, K.M.J., Spencer, T., Vestergaard, M.D., Fletcher, P.C., and Schultz, W. (2016). Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. Neuron *90*, 1127–1138. DOI: 10.1016/j.neuron.2016.04.019

Diederen, K.M.J., Ziauddeen, H., Vestergaard, M.D., Spencer, T., Schultz, W., and Fletcher, P.C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. J. Neurosci. *37*, 1708–1720. DOI: 10.1523/JNEUROSCI.1979-16.2016

Ernst, M.O., and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature *415*, 429–433. DOI: 10.1038/415429a

Fleming, S.M., and Daw, N.D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. Psychol. Rev. *124*, 91–114. DOI: 10.1037/rev0000045

Friston, K.J., Lawson, R., and Frith, C.D. (2013). On hyperpriors and hypopriors: comment on Pellicano and Burr. Trends Cogn. Sci. *17*, 1. DOI: 10.1016/j.tics.2012.11.003

Friston, K.J. (2017). Precision psychiatry. Biol. Psychiat. *2*, 640–643. DOI: 10.1016/j.bpsc.2017.08.007

Geurts, L.S., Cooke, J.R.H., Bergen, R.S. van, and Jehee, J.F.M. (2021). Subjective confidence reflects representation of Bayesian probability in cortex. bioRxiv, 2021.04.10.439272. DOI: 10.1101/2021.04.10.439272

Hertz, U., Palminteri, S., Brunetti, S., Olesen, C., Frith, C.D., and Bahrami, B. (2017). Neural computations underpinning the strategic management of influence in advice giving. Nat. Commun. *8*, 2191. DOI: 10.1038/s41467-017-02314-5

Lawson, R.P., Bisby, J., Nord, C.L., Burgess, N., and Rees, G. (2021). The computational, pharmacological, and physiological determinants of sensory learning under uncertainty. Curr. Biol. *31,* 163-172. DOI: 10.1016/j.cub.2020.10.043

Pothos, E.M., Lewandowsky, S., Basieva, I., Barque-Duran, A., Tapper, K., and Khrennikov, A. (2021). Information overload for (bounded) rational agents. P. R. Soc. B. *288*, 20202957. DOI: 10.1098/rspb.2020.2957

Shea, N., and Frith, C.D. (2019). The global workspace needs metacognition. Trends Cogn. Sci. *23*, 560–571. DOI: 10.1016/j.tics.2019.04.007

Yon, D. (2021). Prediction and learning: understanding uncertainty. Curr. Biol. *31*, R23–R25. DOI: 10.1016/j.cub.2020.10.052

Yon, D., de Lange, F.P., and Press, C. (2019). The predictive brain as a stubborn scientist. Trends Cogn. Sci. *23*, 6–8. DOI: 10.1016/j.tics.2018.10.003

Yon, D., Heyes, C., and Press, C. (2020). Beliefs and desires in the predictive brain. Nat. Commun. *11*, 4404. DOI: 10.1038/s41467-020-18332-9

In brief

In this Primer, Daniel Yon and Chris Frith explain ‘precision’ – a key concept in Bayesian models of the mind and brain. The idea of precision is central to current thinking across the cognitive sciences, but in recent years ideas about precision have begun to change. This raises important questions about precisely how precision works.

Author affiliation and contact

1. Department of Psychology, Goldsmiths, University of London, Lewisham Way, London, SE14 6NW, UK.
2. Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, UK.
3. Institute of Philosophy, School of Advanced Study, University of London, Malet Street, London, WC1E 7HU, UK. 4.
4. Wellcome Centre for Human Neuroimaging, University College London, 12 Queen Square, London, WC1N 3AR, UK.

Correspondence: d.yon@bbk.ac.uk

Acknowledgements

We are grateful to Nick Shea and Steve Fleming for helpful comments on an earlier draft of this manuscript.