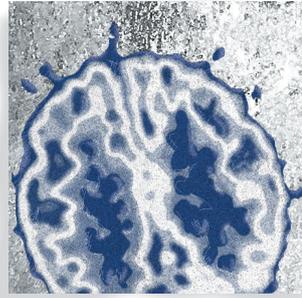


Translational research

Emotor control: computations underlying bodily resource allocation, emotions, and confidence

Adam Kepecs, PhD; Brett D. Mensh, MD, PhD



Introduction

The functions of most organs are understood in mechanistic detail. When their components fail, they usually generate physical symptoms and signs that are fairly objective and rationally relatable to the failed component. These relationships allow physicians to make accurate diagnoses for a wide range of bodily disorders. Psychiatry enjoys none of these conveniences.

Emotional processes are central to behavior, yet their deeply subjective nature has been a challenge for neuroscientific study as well as for psychiatric diagnosis. Here we explore the relationships between subjective feelings and their underlying brain circuits from a computational perspective. We apply recent insights from systems neuroscience—approaching subjective behavior as the result of mental computations instantiated in the brain—to the study of emotions. We develop the hypothesis that emotions are the product of neural computations whose motor role is to reallocate bodily resources mostly gated by smooth muscles. This “emotor” control system is analogous to the more familiar motor control computations that coordinate skeletal muscle movements. To illustrate this framework, we review recent research on “confidence.” Although familiar as a feeling, confidence is also an objective statistical quantity: an estimate of the probability that a hypothesis is correct. This model-based approach helped reveal the neural basis of decision confidence in mammals and provides a bridge to the subjective feeling of confidence in humans. These results have important implications for psychiatry, since disorders of confidence computations appear to contribute to a number of psychopathologies. More broadly, this computational approach to emotions resonates with the emerging view that psychiatric nosology may be best parameterized in terms of disorders of the cognitive computations underlying complex behavior.

© 2015, AICH – Servier Research Group

Dialogues Clin Neurosci. 2015;17:391-401.

Keywords: *computational psychiatry; decision confidence; decision making; model-based neuroscience; RDoC; Research Domain Criteria*

Address for correspondence: Adam Kepecs, PhD, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY, 11724, USA (e-mail: kepecslab.cshl.edu)

Author affiliations: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA (Adam Kepecs); Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, USA (Brett D. Mensh)

Translational research

The workings of the offending organ—the brain—are not understood in mechanistic detail, and when it fails, the symptoms generated are highly subjective. Diagnostic reliability is an intrinsic challenge for mental disorders, which can affect language, thought, emotion, sense of self, and other difficult-to-objectify fruits of mentation. Can we trust patients' reports of their own mental states? How can we overcome psychiatrists' own subjective evaluations of patient reports and interpretations of their behavior? *Prima facie*, it appears that there is no way around the issue of subjectivity unless we find biological diagnostic criteria that are physical in nature. Unfortunately, the core functional components of the mind, and how they interact to mediate mentation, are still largely unknown.

This problem of biological psychiatry—how the mind can be understood in terms of brain processes—is a problem that systems neuroscience has begun to tackle. The approach is to view subjective behavior as the result of mental computations instantiated in the brain. A network of neurons performs a computation by transforming the representations arriving on its inputs into different, useful representations that it outputs to other networks. Employing quantitative models of these computations, which can be tested and refined, allows systematic and rigorous study of the relationships between brain and behavior. This approach has recently enabled concepts such as subjective “valuation” and “confidence” to be quantified in computational models and mapped to specific neural substrates.¹⁻³ These successes have given the field hope that this approach will be broadly applicable in elucidating core cognitive processes and what goes wrong with them.⁴⁻⁶

The experience of emotions is innate,^{7,8} automatic, and central to behavior, yet has defied definitional consensus.⁹⁻¹² Here, we argue that progress in the study of emotions is possible if they are approached from a computational standpoint, as is being done for other forms of cognition. We will outline a new conceptualization of emotions as the product of neural computations that have a particular feature in common: their outputs serve to reallocate bodily resources that are mostly controlled by smooth muscles.

To illustrate the computational framework, we describe research about confidence as an example of a formally defined mental computation that is fundamental to behavior and psychiatry.¹³⁻¹⁹ Although familiar as a feeling, confidence is also an objective statistical quan-

tity. By understanding the computational algorithms for estimating confidence that account for behavioral confidence reports, we can look for neural correlates. This computational process can then be thought of as underlying a behavioral trait, which has important implications for psychiatry. Disorders of confidence appear to contribute to a number of psychiatric diseases: schizophrenia, anxiety disorders, and personality disorders.

Cognition is myriad brain computations

Relating mind stuff (eg, perceptions, thoughts, emotions) to brain stuff (eg, action potentials, synaptic transmission, neuromodulation) is an ancient problem of philosophy. The critical insight into this problem came with the dawn of computer science: cognition is the result of computations performed by the brain.²⁰ This view has been widely embraced in the fields of artificial intelligence, cognitive psychology, and more recently by neuroscience.^{21,22} While there is some disagreement about what precisely is meant by “computation,”²³ here we use the term in its most generic sense, encompassing not only logical, but also statistical, analog, and other modes of information processing—anything that can be mathematically formalized. This definition avoids committing to a particular framework for how the brain might compute, while providing an empirical path forward. Computational models are viewed as hypotheses that are concrete enough to be physically instantiated and hence provide a tool for linking neural machinery and cognitive functions.

In systems neuroscience, our current understanding of how perception and action are supported by sensory processing and motor control is fundamentally computational in nature.^{24,25} Neural computations can be conceptually separated into the identification of what inputs a network receives and how these are transformed into outputs. For instance, in visual processing, early parts of the visual system represent small spots of light that are transformed into representations of edges. The outputs of these edge-detection computations serve as inputs to higher-order networks, which eventually generate complex object representations in the temporal lobe. Determining the precise neural circuitry is an active area of investigation. Nevertheless, the underlying computational principles are understood: retinal information is transformed across layers into more and more complex features by filtering, pooling, amplification,

and normalization at each stage, computations that enable the detection of more and more complex features, while being invariant to irrelevant dimensions.^{24,26-28} Importantly, models of these neurally derived computations can be used to build real-world systems that can recognize objects.²⁹

The ability to link neural activity to external variables that can be directly manipulated (in sensory systems) or observed (in motor systems) has been key to neuroscientific progress. Until recently, it has been unclear how we can go about understanding the intervening processes—cognition—that enable us to adaptively engage with the world. In recent years, a new breed of computational approaches has risen to this challenge.^{1,30} The key idea is that by constructing computational models that predict behavior, we can identify variables in the model that serve as proxies for unobservable variables that guide decision-making. We can then go on to identify these variables in neural representations. Although they differ in their particulars, the principles of these models are the same as those described for sensory and motor processing. They are more challenging to deploy and interpret because the relevant variables are internal and hence not amenable to direct manipulation. However, through indirect manipulations, such as varying behavioral contingencies, this approach has been successful in formalizing concepts like “utility” and “reward prediction error.” Additional successes have given us hope that these problems may be neuroscientifically tractable, for example, the identification of brain structures performing computations such as integrating evidence across time, statistical inference of hidden belief states, or evaluating the unexpectedness of outcomes.³¹⁻³⁵

Note that computational models are never the ground truth, but rather serve as rigorous interpretational tools. Models can be thought of as the formalization of hypotheses, which can be reformulated, adjusted, and tested until they work, ie, account for the data parsimoniously. An exciting aspect of this approach is that it yields experimentally testable predictions, thus rendering squishy psychological concepts neurobiologically accessible.

The brain evolved in order to control movement in an adaptive manner. Thus, the “end result” of the myriad cognitive computations is the triggering of a motor program (or an internal thought that can influence a future motor program). This triggering reflects a

“decision” made by the cognitive networks—the selection of one out of a multitude of options. Such decisions are special in that they commit the entire organism to a course of action. Motion is a commitment that evolution takes seriously.

Emotions are a special class of brain computation

To execute an action, motor cortex, cerebellum, basal ganglia, and other regions coordinate skeletal muscle activation patterns to pull the bones just so, producing movements like walking, jumping, reaching, grabbing, swimming, and slithering. To accomplish this, these neural systems employ internal models of body mechanics, world contents, and physics to maximize the effectiveness of the motor program. The familiar senses provide ongoing feedback about what is being physically encountered, to enable compensatory actions. Whether the body is moving or not, we *feel* these sensations.

The brain exerts its influence over the body not only via skeletal muscle, but also through smooth and cardiac muscle. Skeletal and nonskeletal muscle activation needs to be coordinated. For example, a sudden run requires an increase in blood flow to the muscles, provided by a faster heartbeat and the vasodilation resulting from smooth muscle relaxation in arteriolar walls. While some of this coordination is provided by local homeostasis (eg, muscle contraction-induced carbon dioxide accumulation induces vasodilation), much of this coordination is provided by the autonomic nervous system. Having autonomic actions controlled centrally enables anticipatory resource reallocation (eg, tachycardia in preparation for intense action). Visceral afferents provide ongoing feedback about the status of body organs, to enable compensatory actions.³⁶ We feel these sensations as well.

What then of emotions? Functions of the autonomic nervous system have long been recognized to be closely associated with familiar emotions: flushing with anger, blushing with shame, defecating with fear, and fainting with surprise.^{7,37-39} Within the computational framework we advocate, emotional systems can be viewed as performing computations that support autonomic control. The term “motor control” typically refers to the systems that control the spatiotemporal patterns of *skeletal* muscle activation. In the present framework, emotional systems modulate the spatio-

Translational research

temporal patterns of *nonskeletal* muscle activation—“emotor control” (Figure 1).

The spatiotemporal patterns of nonskeletal muscle activation regulate and allocate bodily resources such as oxygen, nutrients, and heat, through alterations in ventilation, blood flow, metabolism, piloerection, etc. Like the (“voluntary”) decisions to initiate a skeletal movement, these (“involuntary”) decisions to reallocate bodily resources are distinguished by how they pertain to what the whole organism is about to do. (To be more complete, nonskeletal muscle effectors include not only smooth and cardiac muscle, but also excretory and secretory cells, all of which are under both autonomic and neuroendocrine control. And since neural systems tend to be integrated, autonomic and endocrine systems also affect skeletal muscle, eg, blood epinephrine stimulates glycogenolysis in skeletal muscle.)

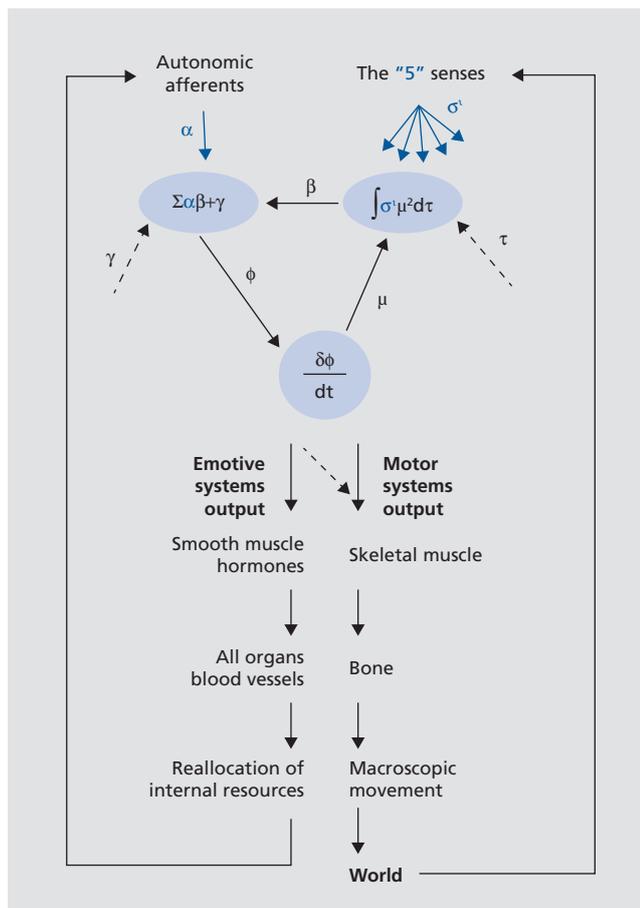


Figure 1. Cartoon of computations underlying emotional and motor control.

What might a computational model of an emotion look like? As an example, consider a brain network that computes the likelihood of a very negative outcome. This could be written down as the distribution of the utility of potential outcomes⁴⁰ multiplied by a negatively biased weighting function. A good computational model, in addition to being specific about mathematical form, needs to propose how its outputs serve as the inputs to other networks that compute/mediate other processes, and so on until it generates something behaviorally quantifiable. When the output of this hypothetical “potential for doom” computation is high, it should generate a trigger signal to prepare the body for negative outcomes. Specifically, this computation of “potential for doom” may induce increases in blood pressure (arteriolar smooth muscle contraction), and increases in pulmonary air flow (bronchiolar smooth muscle relaxation) to prepare for intense skeletal muscle action. Above a threshold, its output representation may be relayed to subcortical areas such as the hypothalamus and periaqueductal gray. These control autonomic functions and have the ability to prepare the body for the doom scenario. The autonomic afferents in turn sense these bodily changes, which together with the original computation, results in the experience of “anxiety.” The feeling associated with such a computation, generating a trigger signal and thus appropriate levels of anxiety, can be adaptive. When the “potential for doom” computation produces a signal that rises above a mere thought and triggers a bodily action, it must be significant for the organism. This significance is realized as a reafferent sensation that manifests itself as an emotion. Importantly, such a model can be tested experimentally and refined, both mathematically and in terms of its relationship to associated brain systems.

Note that we are focusing on a computational account of the brain processes that trigger emotional output and the initial sensory experience they produce. Once triggered, the ensuing dynamics can be complex, and because emotions are often coordinated with other cognitive processes, we perceive them as an integrated unit. With time, emotional experiences tend to become more complex, in the same way the taste of wine cannot be fully appreciated with the first sip. Similarly, emotional experiences are elaborated and evolve in time in a way that is beyond the scope of our framework.

Another important issue is that although we are asserting a primary role for emotional systems in auto-

onomic/endocrine control, these systems are of course highly integrated with skeletal muscle systems; for example, smiling can be also an output of joy.⁴¹ Interestingly, it is easy to fake a smile—virtually all skeletal muscles are under “voluntary” control by nonemotional motor systems.⁴² However, it is difficult to think one’s way into generating tachycardia without experiencing an emotion.³⁹

Why are emotions treated as a special computation by an organism? The impending and ongoing reallocation of bodily resources represents a major investment.^{36,39} Such a highly salient signal merits being broadcast throughout the brain so that diverse networks can be engaged in order to influence particularly important decisions about what the skeletal muscles should do next. “Reafferenting” of autonomic motor signals might serve as such a brain-wide signal, using the body as an intermediary. As in skeletal motor systems, efference copies of these autonomic signals may also be broadcast intracranially, for example via the major neuromodulatory systems arising from the midbrain and hindbrain.

The easy and hard problems of emotions

It appears as though we have glossed over the “magic” of emotions. If their mechanisms are so similar to other brain processes, then why do emotions *feel* different from other mental processes? How a phenomenal experience arises from mentation is the “hard problem” of consciousness.⁴³ Addressing phenomenal experience itself is the domain of philosophy. To approach the issue scientifically, the question becomes what is the best way to operationalize the phenomenal experience (which cannot be directly measured) so that its existence can be inferred? Such an operationalization would allow it to be studied as a neuroscientific question and may be called the “easy problem” of emotions.

Our proposal is to approach the problem of emotions in terms of the cognitive computations that mediate their initiation, motor control, and the ensuing sensations from the periphery. Whether this computational approach will lead to deeper insights remains to be seen; nonetheless, we emphasize that the most important aspect of our proposal is that this operationalization of emotions leads to empirically testable theories. To illustrate this approach, we next turn to the question of subjective confidence, a mental state that has been

recently tackled by systems neuroscience from a computational perspective.

Is confidence a feeling? An emotion?

In facing a world that is often ambiguous and unpredictable we possess a remarkable ability to form beliefs about our environment. An intimate sense of confidence (whether high or low) tends to accompany our beliefs and appears to play a critical role in behaviors from the mundane to the most complex.⁴⁴ Confidence in the likelihood that the adjacent lane is empty, given only a quick glance in the mirror, informs the decision to change lanes. Confidence that a sector of the economy will thrive, given current market data, informs financial investment decisions. This sense of confidence is clearly adaptive and has both cognitive and emotional aspects.

Often, confidence can also be felt, somewhat like an emotion. But without a consensus definition of emotion it is impossible to say whether confidence is one. Confidence seems intimately related to a bona fide emotion, anxiety⁴⁵: as a personality trait, systematically high or low confidence levels are correlated with the degree of anxiety.⁴⁶ Importantly, individuals can report how confident they “feel” about something on a trial-by-trial basis. Regardless of whether confidence constitutes an emotion, its quantifiability makes it a useful case study in relating feelings to computations.

Confidence as a statistical computation

The concept of confidence leads a double life. Besides subjective confidence discussed above, it is widely studied in computational sciences as an entirely objective mathematical quantity. Indeed, it is at the heart of statistical decision theory and machine learning.⁴⁷ This raises the possibility that we can define confidence from first principles in statistics to provide a formal foundation for the scientific inquiry into subjective confidence and its neural basis.³

As a statistical quantity, we can define confidence as the probability estimate that the chosen hypothesis is correct, given the available evidence. Here, evidence can be any source of data contributing to a decision: perceptual, memory, or otherwise. For instance, when driving in the fog, the hypothesis may be a turn to the left, given a barely visible road sign (ie, “I *feel* this is the

Translational research

right way to go”). This example also raises a question: can the experimenter make use of the above definition without having direct access to the subject’s percept (or phenomenal evidence) that was used to arrive at the decision? The surprising answer is yes, at least in principle, if we can manipulate the evidence (eg, by varying the fog density and thus the visibility of the sign) and observe the resulting choice patterns. In addition, we need many repeated events to understand the relationship between the observable variables, the degree of fog, choices, and outcomes. Then it is possible to construct

an ideal observer model and make predictions about the optimal confidence levels.

Using this definition of statistical decision confidence, our group has derived several general properties of statistical confidence without making assumptions about the precise implementation.⁴⁸ First, statistical confidence is a prediction; the degree of confidence predicts the expected fraction of correct choices (*Figure 2-A1*), as intuitively expected. Second, statistical confidence increases with the discriminability of evidence for correct choices, but counterintuitively, for incorrect

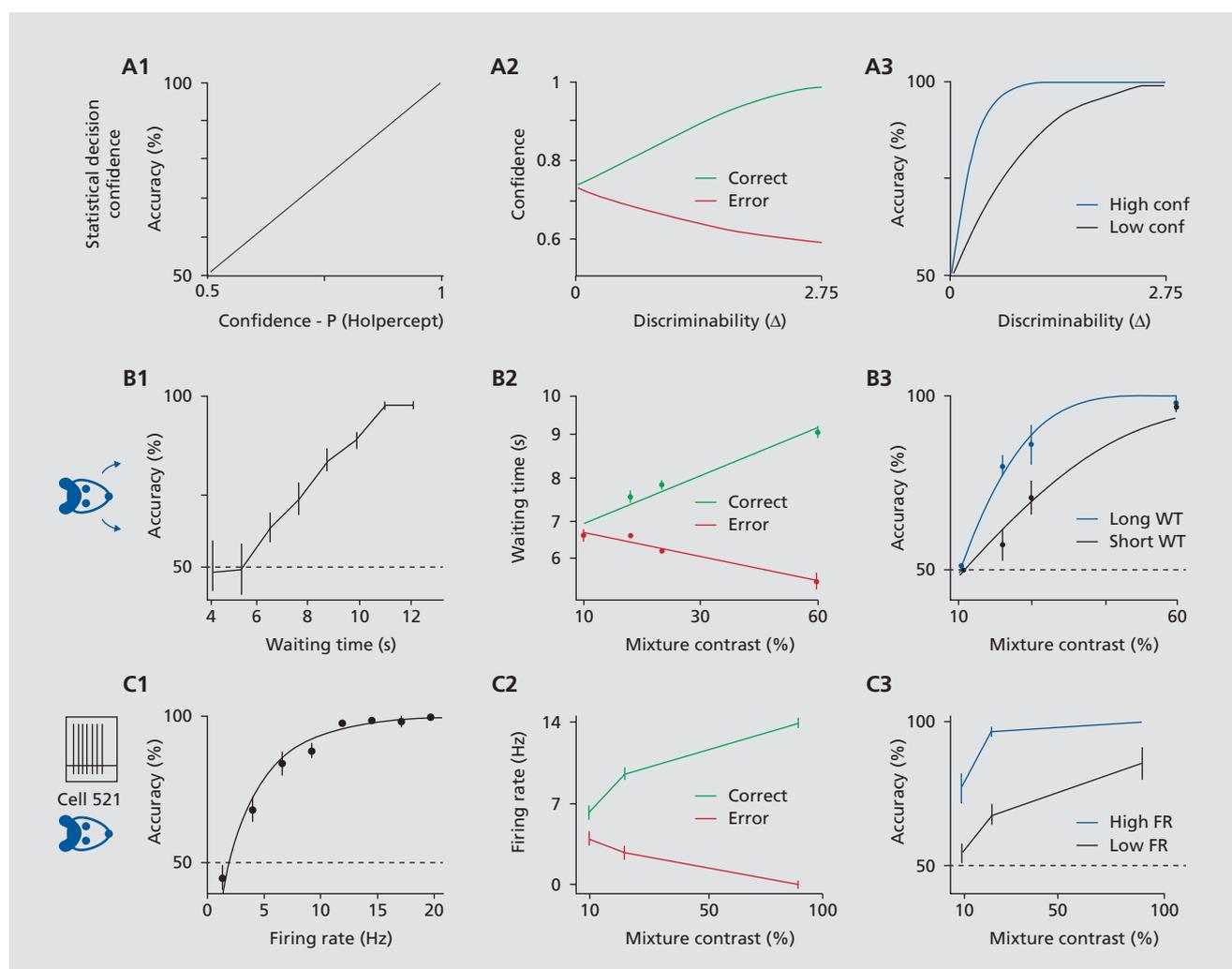


Figure 2. Decision confidence. Monte-Carlo simulations of the normative definition for confidence reveals three types of interrelations between evidence, outcome, and confidence that can be proved mathematically (**A1-A3**).⁴⁸ Post-decision time investment (waiting time) shows all the hallmarks of a confidence estimate (**B1-B3**).¹⁸ Single orbitofrontal cortex neurons can encode confidence (**C1-C3**).¹³ conf, confidence; FR, firing rate; Hz, hertz; S, seconds; WT, waiting time

choices, confidence decreases with increasing discriminability (*Figure 2-A2*). Note that for choices that are not discriminable (eg, complete fog), the mean decision confidence is precisely 0.75 (*Figure 2-A2*). Finally, while discriminability determines accuracy (a property referred to as the psychometric function), confidence provides further information improving the prediction of accuracy for any given level of discriminability (*Figure 2-A3*).

Does statistical confidence map onto subjective confidence in humans? This was recently tested in the context of a perceptual decision task.⁴⁹ The critical observation was that the relationship between human confidence, discriminability, and choice correctness reveals robust patterns (as in *Figure 2A*), which could also be quantitatively predicted by statistical decision confidence. These results are the first step toward providing a link between the objective and subjective notions of confidence, but we want to emphasize that much further work will be required to further strengthen this link and understand the conditions under which it applies.

Of course, many confidence judgments are not entirely veridical: the accuracy of confidence reports is often not a good predictor of outcomes. In fact, there are many laboratory studies that show a strong divorce between confidence and accuracy,⁵⁰⁻⁵² although the generality of these observations has been questioned since in many everyday decisions when people are experts, their probability judgments seem accurate.⁵³ A more subtle but commonly observed effect reveals miscalibration: people tend to overestimate their confidence in outcomes with small probabilities, and often simultaneously underestimate their confidence in the context of high probabilities.⁵⁴ Thus, statistical computations may only provide an initial confidence estimate for human decision-makers. We suspect that in many circumstances, this internal value is further modified by context (eg, social factors), accounting for a range of reported assays where confidence can be divorced from choice accuracy.

How to spot confidence in the brain

Framing confidence as a statistical estimation process yields insight into the behavioral process, but how does it help us in looking for the underlying neural mechanisms? The core issue of course remains that confidence, as a subjective experience, is only accessible via

introspection. This often leads to an approach that takes human self-reports literally and searches for their neural correlates. Beyond the concern that human reports can often contain components unrelated to confidence (framing effects, biases, etc) this approach cannot be applied to nonhuman animals, where we can more powerfully measure and manipulate neural activity.

Recent work in several labs, including in one of ours (AK), has placed confidence on solid neurobiological footing.¹³⁻¹⁹ Here, we review these studies of confidence built on two pillars: well-designed behavioral tasks to incentivize appropriate reports of confidence, and the use of a computational framework to validate the behavior and search for neural correlates.

How could animals possibly consider their thoughts and report their confidence? We argue that underlying confidence is an evolutionarily ancient computation that enables organisms to optimize future behavior amid uncertainty. In the laboratory, establishing a confidence reporting behavior requires one to incentivize animals to use confidence, for instance by enabling animals to collect more reward or seek out valuable information based on confidence. In the above driving example, imagine that after the turn, the driver is looking for a restaurant, but it does not appear. Eventually, the driver will give up and turn around. The time invested before turning around ought to reflect the degree of confidence in the original decision to make the turn. Naturally, in a complex situation such as this there may be many other factors contributing, and because of individual differences it is impossible to infer the exact degree of confidence from the time spent. However, improved inference is possible with better behavioral control and many repeated trials available in animal studies.

In a recent study, we trained rats in an analogous situation in which we could assess their confidence by measuring their willingness to wait.¹⁸ Rats were trained to insert their snouts into an odor port that delivered a binary mixture of odors. Rats then had to determine the dominant odor in the mixture and respond to one of the choice ports located on either side of the odor port. Correct choices were rewarded with a drop of water. For easy mixtures (say 90% odor A and 10% odor B), rats nearly always responded to the correct side A, while they were challenged for difficult mixtures (say 55% odor A and 45% odor B). To infer their confidence, we introduced a random delay between the choice and the outcome (reward or not)

Translational research

and measured how long rats were willing to wait for the water reward. We reasoned that an animal waiting for an uncertain reward is more likely to wait longer if it is confident in ultimately receiving the reward. We formalized this intuition with a model and showed that waiting time should be proportional to confidence. In agreement with theoretical predictions (*Figure 2-A1*), waiting time (WT) is correlated with decision accuracy (*Figure 2-B1*) and jointly reflects stimulus difficulty and outcome (*Figure 2-B2*). Thus, the timing of these leaving decisions in a psychometric perceptual decision-making task provides a graded behavioral measure of decision confidence.¹⁸

Can we identify neural representations of confidence? We assessed whether the orbitofrontal cortex (OFC), an area involved in representing and predicting decision outcomes, carries neural signals related to confidence. We recorded the firing of OFC neurons while rats performed a task similar to the one described above and focused our analysis on the reward anticipation period.¹³ This is the period of waiting at the reward port, after a choice was made but before any feedback was received about the outcome of a choice. *Figure 2C* shows an example neuron whose firing rate during this anticipation period signals decision confidence based on the criteria outlined above. For instance, one dramatic signature of confidence is that the firing rates of many neurons predict the accuracy of the rat's choice before feedback is provided about choice correctness (*Figure 2-C1*). It is important to note that these neurons may also be correlated with “anxiety,” “arousal,” or “exploration.” Indeed, these concepts can be related to different uses of uncertainty. However, our interpretation does not hinge on these terms, but rather on the notion that confidence estimation was the only class of computational mechanisms that we found could successfully explain the observed firing patterns. Next, we inactivated OFC pharmacologically and found that it specifically degrades the accuracy of trial-to-trial confidence reports behaviorally, without changing decision accuracy.¹⁸

The results leave open the question whether OFC locally computes confidence or receives the constituent signals from other areas. We suspect that there are other brain regions that compute choice and confidence together and relay it to OFC. Indeed, two studies identified neuronal correlates with confidence in visual decisions upstream of OFC, in the parietal cortex and pul-

vinar nucleus of the thalamus.^{14,17} Thus, it may be that OFC is a central confidence-monitoring region, combining information from various sources independent of sensory modality, a hypothesis that remains to be tested.

Disorders of confidence

This computational perspective on confidence, and behavior in general, has important implications for psychiatry. Psychopathologies, by definition, are behavioral in nature and hence subjective, making diagnostic reliability an intrinsic challenge to clinical psychiatry. Importantly, even when a specific configuration of symptoms can reliably describe a disease category, it does not point to mechanisms.

We believe that reconceptualizing confidence as a computational process makes it possible to quantitatively study even its subjective aspects. Underlying any expression of confidence is a behavioral process that evolved to serve as a calibration tool for arriving at correct beliefs about our actions and states of the world. It endows us with a sense of when to persevere and when to quit and contributes to a range of other processes. The underlying computational process is not necessarily optimal in the way it uses internal evidence or in the way in which it is calibrated to yield the correct accuracy. For instance, people may be systematically miscalibrated (*Figure 3*).

When confidence computations are more severely disrupted, they can contribute to a range of psychiatric disorders.⁴⁵ For instance, if confidence is systematically

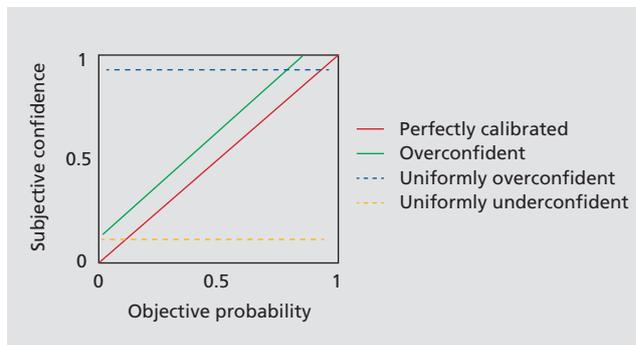


Figure 3. Calibration of confidence. Since confidence reports are estimates of an underlying objective probability, a critical issue is how good a predictor they are. While small systematic overconfidence may even be adaptive, uniformly over- or underconfidence can be signs of psychopathologies.

low, even when the objective probability of outcomes is high, it unnecessarily fuels anxiety. Indeed, there is evidence for underconfidence in anxiety disorders, such as obsessive-compulsive disorder (OCD).^{46,55} Patients with OCD are often inflicted with pathological doubt; this has been linked with various measures of confidence.^{46,56,57} Interestingly, the OFC has been strongly implicated in OCD, which may be related to the results from our rat studies described above. The opposite situation, uniformly high confidence independent of evidence, might contribute to other psychopathologies. Interestingly, although narcissists overtly express high confidence, the overt behavioral manifestations may in fact hide systematic underconfidence.⁵⁸ Besides these somewhat intuitive links between confidence and mental disorders, there is also increasing evidence that schizophrenics have reduced ability to estimate confidence.^{59,60} For instance, the mixed findings of a study showing that schizophrenic patients are less confident in their correct choices, but overconfident when they make errors, would be a simple consequence of a degraded statistical confidence estimate⁶¹ (Figure 2). Another line of investigation proposed that depressed individuals make more realistic judgments about beliefs.⁶² Following up on this “depressive realism” hypothesis, some studies found that depressive patients have attenuated capacity for confidence reporting,⁶³ while others found evidence against this.⁶⁴ A more complete computational framework might be able to resolve some of these controversies and also distinguish between different types of abnormal confidence: miscalibration, degraded confidence computation, or maladaptive postestimation adjustments to confidence, so-called framing effects.

Outlook for psychiatry and application to Research Domain Criteria

The computational perspective on cognition and emotion we advocate strongly resonates with the Research Domain Criteria (RDoC) initiative of the National

Institute of Mental Health.⁶⁵⁻⁶⁷ The RDoC framework seeks to generate a new classification of mental disorders based on dimensions of observable behavior and neurobiological measures. The hope is that these dimensions will cut across current disease categories and map more directly to mechanisms, thus better integrating neuroscience and psychiatry. One critical difference from the RDoC as currently conceived is that in the framework discussed above we are specifically concerned with dimensions that are the results of computations, as opposed to arbitrary behavioral measures. Thus, it may be possible to make progress by focusing on behavior alone, the primary manifestation of psychiatric disorders. The critical challenge will be to identify relevant behavioral metrics using a data-driven, iterative approach.

There are already a number of research efforts focused on “computational psychiatry.” The important feature of this research program is that it enables the development and validation of quantitative behavioral dimensions through iterative improvements of computational models. Moreover, such model-based approaches are particularly powerful because they can also be used to rigorously link neural circuit mechanisms with behavioral observations—thus, automatically furthering the RDoC’s initial goals to link pathophysiology with neuroscience insights. Emotions are particularly challenging to quantify and categorize; thus, a data-driven computational approach to their measurement is expected to provide new insights.

More broadly, the ideas reviewed here have significant implications for biological psychiatry. Even as our knowledge of neuroscience expands at a rapid pace, linking biological insights to clinical observations—observations based on subjective behavior—remains an unbroken challenge. Thus, the hope is that when applied to an appropriate behavioral task, the interpretation of behavioral measures, such as confidence reports via computational models, will lead to quantitative dimensional measures of personal traits that can inform research and, ultimately, clinical diagnoses.⁴⁻⁶ □

REFERENCES

1. Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci.* 2005;6(5):363-375.
2. Kable JW, Glimcher PW. The neurobiology of decision: consensus and controversy. *Neuron.* 2009;63(6):733-745.
3. Kepecs A, Mainen ZF. A computational framework for the study of confidence in humans and animals. *Philos Trans R Soc Lond B Biol Sci.* 2012;367(1594):1322-1337.
4. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn Sci.* 2012;16(1):72-80.
5. Stephan KE, Montague R, Dolan RJ, Friston KJ. Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry.* 2014;1(2):148-158.

Translational research

Control emotor: cálculos que están a la base de la distribución de los recursos corporales, las emociones y certidumbre

Los procesos emocionales son centrales para la conducta, pero su naturaleza intensamente subjetiva ha sido un desafío para el estudio neurocientífico y el diagnóstico psiquiátrico. En este artículo se exploran las relaciones entre los sentimientos subjetivos y los circuitos cerebrales subyacentes desde una perspectiva computacional. Para el estudio de las emociones se aplican conocimientos recientes de la neurociencia de sistemas, planteándose la conducta subjetiva como el resultado de cálculos mentales que se ejemplifican concretamente en el cerebro. Se desarrolla la hipótesis que las emociones son el producto de cálculos neurales cuyo papel motor es redistribuir los recursos corporales regulados por los músculos lisos, análogo a los cálculos del control motor que coordinan los movimientos del músculo esquelético. Para ilustrar este modelo se revisa la investigación reciente sobre la "certidumbre". Aunque ésta es familiar como un sentimiento, la certidumbre es también una magnitud estadística objetiva: una estimación de la probabilidad de que una hipótesis sea correcta. Este enfoque basado en el modelo ayuda a revelar las bases neurales de la certidumbre en las decisiones en mamíferos y construye un enlace al sentimiento subjetivo de certidumbre en humanos. Estos resultados tienen importantes implicancias para la psiquiatría, ya que los trastornos de los cálculos de certidumbre parecen contribuir a numerosas psicopatologías. En un sentido más amplio, este enfoque computacional de las emociones se hace eco de la visión emergente de la nosología psiquiátrica que puede ser mejor parametrizada en términos de trastornos de los cálculos cognitivos subyacentes a las conductas complejas.

Contrôle moteur : computations sous-tendant l'affectation des ressources corporelles, les émotions et la confiance

Les processus émotionnels sont au cœur du comportement, pourtant leur nature subjective profonde s'est révélée être une difficulté pour les études neuroscientifiques comme pour le diagnostic psychiatrique. Nous examinons ici les relations entre les sentiments subjectifs et leurs circuits cérébraux d'un point de vue computationnel. Nous appliquons les connaissances récentes de la neuroscience des systèmes - en considérant le comportement subjectif comme résultat de computations mentales générées dans le cerveau - à l'étude des émotions. Nous développons l'hypothèse que les émotions sont le produit de computations neuronales dont le rôle moteur est de réattribuer les ressources corporelles principalement contrôlées par les muscles lisses. Ce système de contrôle moteur émotionnel est semblable aux calculs plus familiers du contrôle moteur qui coordonnent les mouvements musculaires du squelette. Pour illustrer cette perspective, nous examinons la recherche récente sur la « confiance ». Bien que connue comme sentiment, la confiance est aussi une quantité statistique objective : une estimation de la probabilité qu'une hypothèse est exacte. Cette approche fondée sur un modèle permet de révéler les bases neuronales de la confiance décisionnaire chez les mammifères et fournit une passerelle au sentiment subjectif de confiance chez les humains. Les implications de ces résultats en psychiatrie sont importantes, puisque des troubles de computations de confiance semblent contribuer à de nombreuses psychopathologies. Plus largement, cette approche computationnelle des émotions fait émerger la possibilité d'une nosologie psychiatrique pouvant être mieux paramétrée en termes de troubles des computations cognitives sous-tendant des comportements complexes.

6. Wiecki TV, Poland J, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clin Psychol Sci*. 2015;3(3):378-399.
7. James W. II. What is an emotion? *Mind*. 1884;os-IX(34):188-205.
8. Anderson DJ, Adolphs R. A framework for studying emotions across species. *Cell*. 2014;157(1):187-200.
9. Dubois J, Adolphs R. Neuropsychology: how many emotions are there? *Curr Biol*. 2015;25(15):R669-R672.
10. LeDoux J. Rethinking the emotional brain. *Neuron*. 2012;73(4):653-676.
11. Russell JA. Introduction to special section: on defining emotion. *Emotion Rev*. 2012;4(4):337-337.

12. Izard CE. Basic emotions, relations among emotions, and emotion-cognition relations. *Psychol Rev*. 1992;99(3):561-565.
13. Kepecs A, Uchida N, Zariwala HA, Mainen ZF. Neural correlates, computation and behavioural impact of decision confidence. *Nature*. 2008;455(7210):227-231.
14. Kiani R, Shadlen MN. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*. 2009;324(5928):759-764.
15. Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. Relating introspective accuracy to individual differences in brain structure. *Science*. 2010;329(5998):1541-1543.
16. Rolls ET, Grabenhorst F, Deco G. Choice, difficulty, and confidence in the brain. *Neuroimage*. 2010;53(2):694-706.

17. Komura Y, Nikkuni A, Hirashima N, Uetake T, Miyamoto A. Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nat Neurosci.* 2013;16(6):749-755.
18. Lak A, Costa GM, Romberg E, Koulakov AA, Mainen ZF, Kepecs A. Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron.* 2014;84(1):190-201.
19. Fetsch CR, Kiani R, Newsome WT, Shadlen MN. Effects of cortical microstimulation on confidence in a perceptual decision. *Neuron.* 2014;83(4):797-804.
20. Turing AM. Computing machinery and intelligence. *Mind.* 1950;59:433-460. doi:10.2307/2251299.
21. Gershman SJ, Horvitz EJ, Tenenbaum JB. Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science.* 2015;349(6245):273-278.
22. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND. How to grow a mind: statistics, structure, and abstraction. *Science.* 2011;331(6022):1279-1285.
23. Hanson PP. Review of Zenon W. Pylyshyn 'Computation and cognition.' *Dialogue: Canadian Philosophical Review.* 1986;25(4):811-814.
24. Heeger DJ, Simoncelli EP, Movshon JA. Computational models of cortical visual processing. *Proc Natl Acad Sci U S A.* 1996;93(2):623-627.
25. Wolpert DM, Ghahramani Z. Computational principles of movement neuroscience. *Nat Neurosci.* 2000;3(suppl):1212-1217.
26. Niebur E, Wörgötter F. Design principles of columnar organization in visual cortex. *Neural Comput.* 1994;6:602-614.
27. Gollisch T, Meister M. Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron.* 2010;65(2):150-164.
28. Zhang Y, Meyers EM, Bichot NP, Serre T, Poggio TA, Desimone R. Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci U S A.* 2011;108(21):8850-8855.
29. Serre T, Poggio T. A neuromorphic approach to computer vision. *Commun ACM.* 2010;53(10):54-61.
30. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature.* 2006;441(7095):876-879.
31. Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature.* 1999;400(6741):233-238.
32. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci.* 1996;16(5):1936-1947.
33. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997;275(5306):1593-1599.
34. Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol.* 2001;86(4):1916-1936.
35. Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science.* 2004;304(5678):1782-1787.
36. Saper CB. The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. *Annu Rev Neurosci.* 2003;25:433-469.
37. Darwin C. *The Expression of the Emotions in Man and Animals* (with an introduction, afterword and commentaries by Paul Ekman). 3rd ed. New York, NY: Oxford University Press; 1872/1998.
38. Damasio AR. *The Feeling of What Happens: Body, Emotion and the Making of Consciousness.* New York, NY: Random House; 2000.
39. Kreibitz SD. Autonomic nervous system activity in emotion: a review. *Biol Psychol.* 2010;84(3):14-41.
40. Weber EU, Shafir S, Blais AR. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychol Rev.* 2004;111(2):430-445.
41. Ekman P, Friesen WV, Ellsworth P. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings.* Burlington, MA: Elsevier; 2013.
42. Ekman P, Sorenson ER, Friesen WV. Pan-cultural elements in facial displays of emotion. *Science.* 1969;164(3875):86-88.
43. Chalmers DJ. Facing up to the problem of consciousness. *J Consc Stud.* 1995;2(3):200-219.
44. Metcalfe JE, Shimamura AP. *Metacognition: Knowing About Knowing.* Cambridge, MA: MIT Press; 1994.
45. Cacioppo JT, Berntson GG, Larsen JT, Poehlmann KM, Ito TA. The psychophysiology of emotion. In: Lewis M, Haviland-Jones JM, eds. *Handbook of Emotions*, 3rd ed. New York, NY: Guilford Press; 2000.
46. Barahmand U, Tavakolian E, Alaei S. Association of metacognitive beliefs, obsessive beliefs and symptom severity with quality of life in obsessive-compulsive patients. *Arch Psychiatr Nurs.* 2014;28(5):345-351.
47. Berger JO. *Statistical Decision Theory and Bayesian Analysis.* Berlin, Germany: Springer Science & Business Media; 2013.
48. Hangya B, Sanders JI, Kepecs A. A mathematical framework for statistical decision confidence. *bioRxiv.* 2015. doi:10.1101/017400.
49. Kepecs A, Sanders JI, Hangya B. A computational framework for relating decision confidence reports across species. In: *Neuroscience 2012.* New Orleans, LA: Society for Neuroscience; 2012. Abstract 527.06.
50. Adams JK, Adams PA. Realism of confidence judgments. *Psychol Rev.* 1961;68:33-45.
51. Nelson TO. A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychol Bull.* 1984;95(1):109-133.
52. Lichtenstein S, Fischhoff B. Do those who know more also know more about how much they know? *Organ Behav Hum Performance.* 1977;20(2):159-183.
53. Griffiths TL, Tenenbaum JB. Optimal predictions in everyday cognition. *Psychol Sci.* 2006;17(9):767-773.
54. Keren G. Calibration and probability judgments: conceptual and methodological issues. *Acta Psychol.* 1991;77(3):217-273.
55. Purdon C, Clark DA. Metacognition and obsessions. *Clin Psychol Psychother.* 1999;6(2):102-110.
56. Hermans D, Engelen U, Grouwels L, Joos E, Lemmens J, Pieters G. Cognitive confidence in obsessive-compulsive disorder: distracting perception, attention and memory. *Behav Res Ther.* 2008;46(1):98-113.
57. Lazarov A, Dar R, Oded Y, Liberman N. Are obsessive-compulsive tendencies related to reliance on external proxies for internal states? Evidence from biofeedback-aided relaxation studies. *Behav Res Ther.* 2010;48(6):516-523.
58. Campbell WK, Goodie AS. Narcissism, confidence, and risk attitude. *J Behav Decis Making.* 2004;17(4):1-15.
59. Bacon E, Izaute M. Metacognition in schizophrenia: processes underlying patients' reflections on their own episodic memory. *Biol Psychiatry.* 2009;66(11):1031-1037.
60. Moritz S, Ramdani N, Klass H, et al. Overconfidence in incorrect perceptual judgments in patients with schizophrenia. *Schizophr Res Cogn.* 2014;1(4):165-170.
61. Moritz S, Woodward TS. The contribution of metamemory deficits to schizophrenia. *J Abnorm Psychol.* 2006;115(1):15-25.
62. Alloy LB, Abramson LY. *Depressive Realism: Four Theoretical Perspectives.* New York: Guilford Press; 1988.
63. Dunning D, Story AL. Depression, realism, and the overconfidence effect: are the sadder wiser when predicting future actions and events? *J Pers Soc Psychol.* 1991;61(4):521-532.
64. Fu T, Koutstaal W, Fu CHY, Poon L, Cleare AJ. Depression, confidence, and decision: evidence against depressive realism. *J Psychopathol Behav Assess.* 2005;27(4):243-252.
65. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry.* 2010;167(7):748-751.
66. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* 2013;11:1.
67. Cuthbert BN. Research domain criteria: toward future psychiatric nosologies. *Dialogues Clin Neurosci.* 2015;17(1):89-97.