Author/s:
D'Acremont, M;Bossaerts, P

Title:
Neural Mechanisms behind Identification of Leptokurtic Noise and Adaptive Behavioral Response

Date:
2016-04-01

Citation:
D'Acremont, M. & Bossaerts, P. (2016). Neural Mechanisms behind Identification of Leptokurtic Noise and Adaptive Behavioral Response. Cerebral Cortex, 26 (4), pp.1818-1830. https://doi.org/10.1093/cercor/bhw013.

Persistent Link:
https://hdl.handle.net/11343/261956

Cerebral Cortex, April 2016;26: 1818–1830

doi: 10.1093/cercor/bhw013
Advance Access Publication Date: 4 February 2016
Original Article

ORIGINAL ARTICLE

Neural Mechanisms Behind Identification of Leptokurtic Noise and Adaptive Behavioral Response

Mathieu d'Acremont[1,2] and Peter Bossaerts[3]

[1]Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA, [2]AIG–Science, New York, NY, USA and [3]Faculty of Business and Economics, The Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Parkville, Victoria, Australia

Address correspondence to Prof. Peter Bossaerts, Faculty of Business and Economics, The University of Melbourne, 198 Berkeley Street, Parkville, VIC 3010, Australia. Email: peter.bossaerts@unimelb.edu.au

Abstract

Large-scale human interaction through, for example, financial markets causes ceaseless random changes in outcome variability, producing frequent and salient outliers that render the outcome distribution more peaked than the Gaussian distribution, and with longer tails. Here, we study how humans cope with this evolutionary novel leptokurtic noise, focusing on the neurobiological mechanisms that allow the brain, 1) to recognize the outliers as noise and 2) to regulate the control necessary for adaptive response. We used functional magnetic resonance imaging, while participants tracked a target whose movements were affected by leptokurtic noise. After initial overreaction and insufficient subsequent correction, participants improved performance significantly. Yet, persistently long reaction times pointed to continued need for vigilance and control. We ran a contrasting treatment where outliers reflected permanent moves of the target, as in traditional mean-shift paradigms. Importantly, outliers were equally frequent and salient. There, control was superior and reaction time was faster. We present a novel reinforcement learning model that fits observed choices better than the Bayes-optimal model. Only anterior insula discriminated between the 2 types of outliers. In both treatments, outliers initially activated an extensive bottom-up attention and belief network, followed by sustained engagement of the fronto-parietal control network.

Key words: anterior insula, fronto-parietal control network, leptokurtic noise, outliers, reinforcement learning

Introduction

Humans have always had to navigate a natural environment replete with extreme events or "outliers." Formally, outliers are observations in the tails of the distribution of expected outcomes (visual or auditory stimuli, rewards, etc.). Usually, outliers signal a change in contingency that requires adaptation. In the laboratory, this situation has been emulated in "contingency-shift paradigms" (Daw et al. 2006; Behrens et al. 2007; Brown and Steyvers 2009; Payzan-LeNestour and Bossaerts 2010; Payzan-LeNestour et al. 2013; McGuire et al. 2014). There, postoutlier adaptation requires extensive learning and exploration. The situation has also been emulated in oddball paradigms (Kim 2014), where the focus is on adequate behavioral control after detection of an outlier. Mathematically, reversal learning tasks (Hampton et al. 2006) fall in the same category. Humans are known to do well in contingency-shift or oddball paradigms, and the neurobiology supporting this capacity has begun to be understood. The crucial role of noradrenergic neurons in the locus coeruleus (LC) in detecting outliers and facilitating subsequent belief and behavioral adjustment has long been acknowledged (Yu and Dayan 2003, 2005; Cohen et al. 2007; Preuschoff et al. 2011; Nassar et al. 2012; Payzan-LeNestour et al. 2013). The outliers engage a large attentional network (AN), including thalamus, posterior parietal cortex (PPC) and anterior cingulate cortex (ACC) (Kim 2014). This network

is not monolithic: Recently, a dissociation has been found between PPC, involved in tracking the surprise reflected in the outliers (i.e., detection of improbable events), and ACC, engaged in subsequent belief updating (i.e., adjustment of probabilities) (O'Reilly et al. 2013). Animal models have provided details as to how LC alters ACC output to facilitate belief adjustment and, hence, behavioral adaptation (Tervo et al. 2014).

Outliers could occur for a very different reason, however. They could be the result of continuous random shifts in observational error variance. In that case, the resulting distribution is unusual: It is far more peaked and displays much heavier tails than the Gaussian distribution (Fig. 1*a*). The distribution is said to be "leptokurtic," a reference to its fourth moment, kurtosis, which is higher than that of the Gaussian distribution. Under sustained variance shifting (technically, variance "mixing"), outliers are frequent and salient. Unlike in contingency-shift paradigms or oddball paradigms, the outliers constitute noise because the shift in variance affects the observation error and not the underlying value of the process. But, importantly, the outliers must not be ignored. They have to be acted upon, in ways that differ often substantially from optimal choice in contingency-shift or oddball paradigms. In a situation emulating the movement of stock market prices during periods of high volatility, we will show—using Bayesian principles—that this type of outlier calls for restraint. This "leptokurtic noise" is rare in the natural environment of humans, where most noise is either Gaussian or in the domain of attraction of the Gaussian distribution (e.g., Poisson, exponential, binomial). Leptokurtic noise has emerged in the modern social sphere of humans, however, where it is the consequence of large-scale interaction. Financial markets constitute the most widely studied case (Mandelbrot 1963; Embrechts et al. 1997). For instance, empirical distributions of daily price changes of equity, a typical financial security, exhibit kurtosis as high as 10 (compared with 3 for the Gaussian distribution); see Figure 1*b*. Leptokurtic noise has emerged through other types of large-scale human interaction, such as internet communication and air traffic (Hsu 1979; Wolpert and Taqqu 2005; Fischer 2010). Leptokurtic noise is to be distinguished from other phenomena that create high kurtosis, such as the ubiquitous "pink noise" (Mandelbrot and Van Ness 1968). The distinguishing feature of leptokurtic noise is not only the peakedness of the distribution (small outcomes are disproportionally frequent relative to medium-sized outcomes), but more importantly its complete lack of memory—its power spectrum is flat; the noise is "white." In sharp contrast, "pink noise," for instance, is characterized by exceptionally long memory.

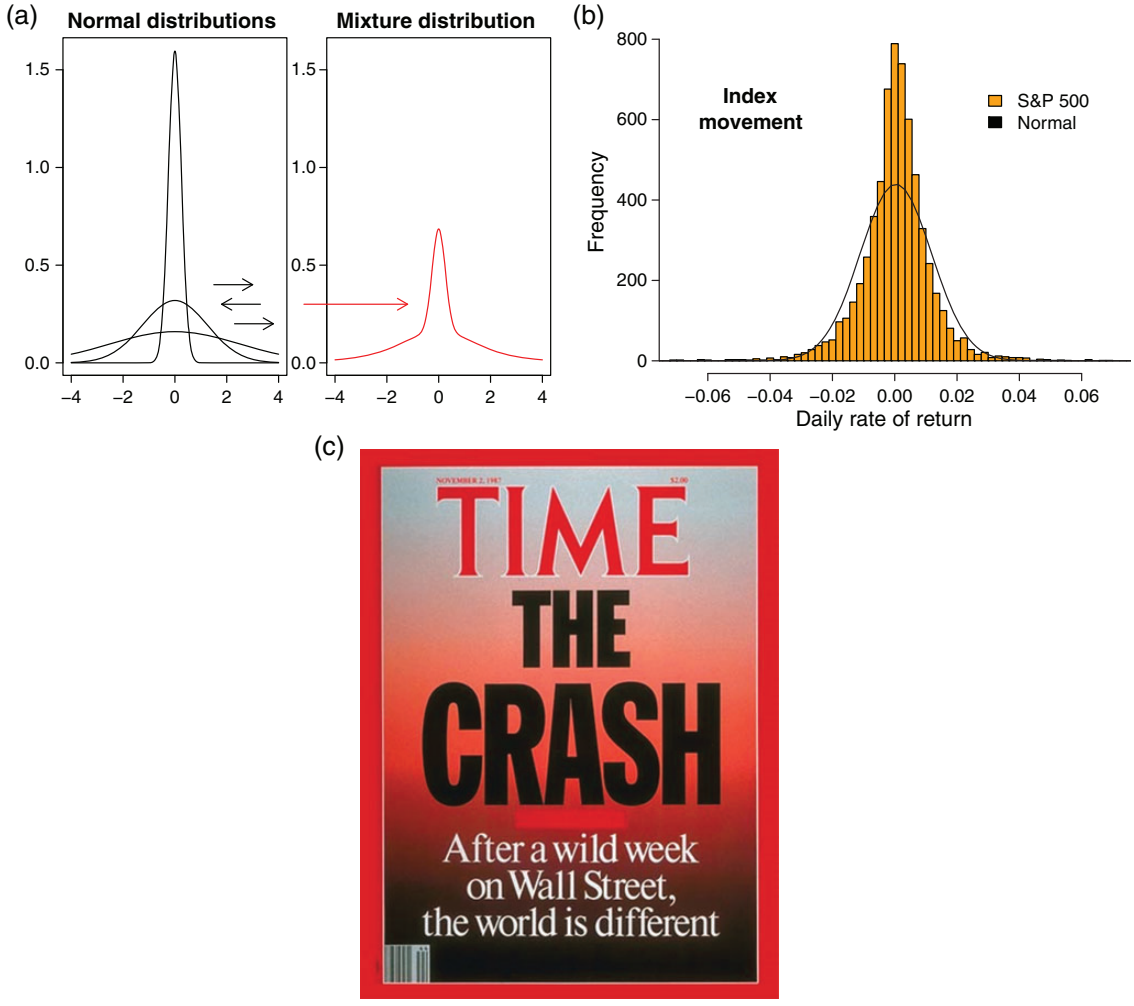

Figure 1. (*a*) Random shifting in ("mixing" of) the variance of a Gaussian distribution (left) generates an outcome distribution that is leptokurtic (right); outliers are frequent and salient. (*b*) Histogram and fitted Gaussian curve, daily rates of return (rates of movement of value from market close in 1 day to the next) of Standard and Poor's (S&P) index of 500 major US common stock, 1 June 1988 to 28 June 2013. (*c*) Cover of Time magazine for the week after the October 1987 stock market crash.

Casual evidence would seem to suggest that humans overreact to leptokurtic noise. The 1987 stock market crash is one example. At the time, it looked arresting and indeed many thought it to be the beginning of a new era; see Figure 1*c*. Nevertheless, the outlier was merely an aberration. Within 1 year, the stock market had more than re-covered. Traders who had pulled out of the market in 1987 thinking that something fundamental had changed, ended up bearing large opportunity costs. That even professional investors at times overreact is a well-documented phenomenon (De Bondt and Thaler 1990).

To date, there does not exist systematic evidence of how humans react to leptokurtic noise, let alone a deeper understanding of its neurobiological foundations. We designed a simple experimental paradigm where subjects had to guide a robot to track as closely as possible a target that moved on a circle. Movements were affected by leptokurtic noise (red dot; Fig. 2*a*). This meant that the target frequently made large swings that reverted in the subsequent trial. Because of the reversals, we refer to this setting as the "Transitory Treatment."

Evolutionary biologists have long conjectured that the human brain was developed to adapt rapidly to novel types of risk, and

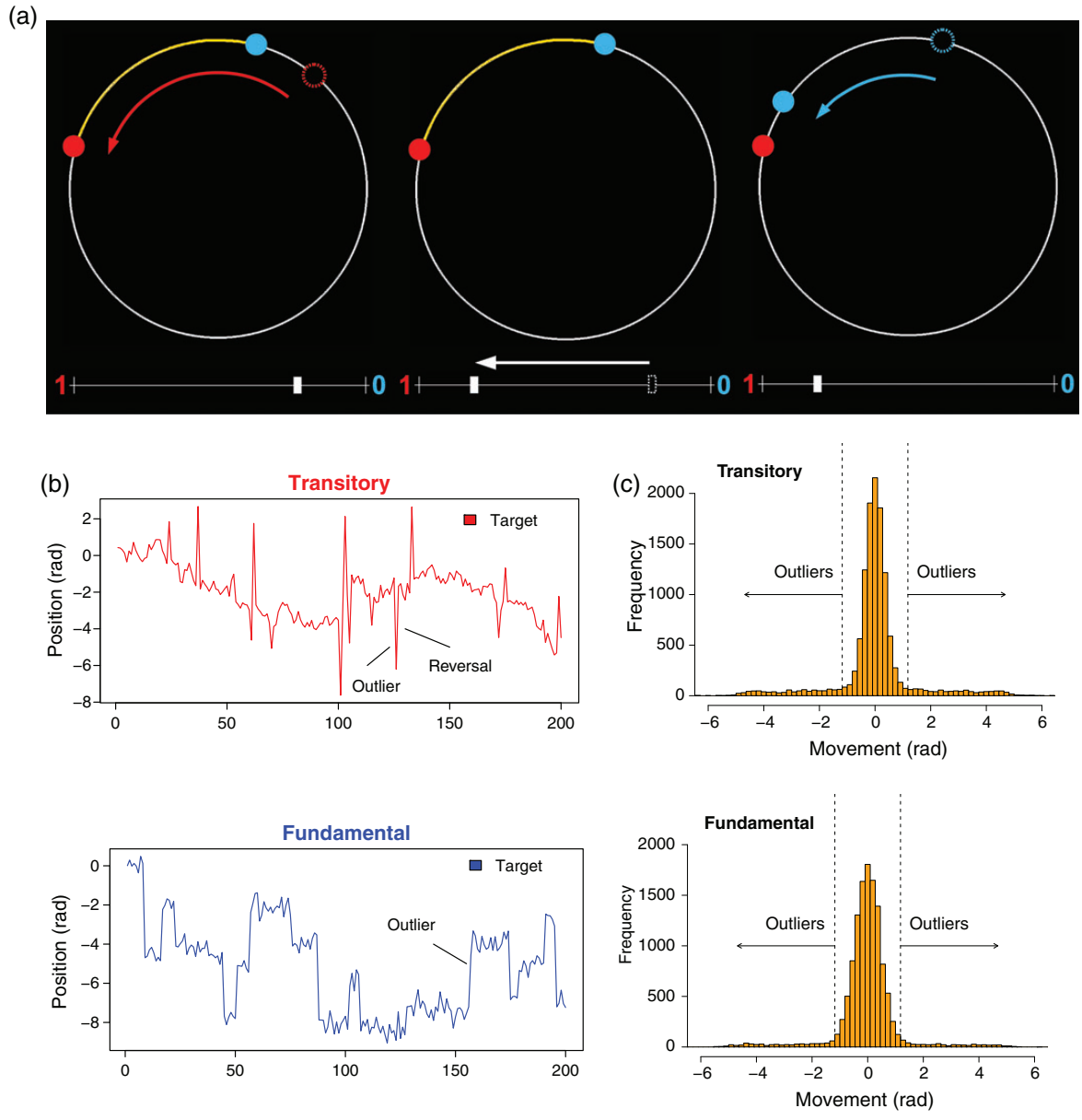

Figure 2. (*a*) Sequence of events in one trial. Left: Following a movement of the target (old position: open red circle; new position: filled red circle; movement direction: red arrow), the arc between the position of the robot (filled blue circle) and the target is indicated in yellow. This is also the "prediction error," that is, the size of the mistake of the participant's guess of the new position of the target. Middle: Participant adjusts position of robot by moving a slider (old position: open white rectangle; new position: filled white rectangle) somewhere between 0 (robot stays at current location) to 1 (robot moves all the way to target position). Here, participant instructs the robot to move about 80% toward the target. The goal is to move the robot as close as possible to the position of the target in the next trial. Right: Robot executes the instruction and moves about 80% toward the target (blue arrow; old position indicated by open blue circle; new position indicated by closed blue circle). (*b*) Sample paths of locations of target in one Run, Transitory Treatment (top), and Fundamental Treatment (bottom). Outliers are salient, showing as sudden large movements. In the Transitory Treatment, outliers revert in the subsequent trial. They are permanent in the Fundamental Treatment. (*c*) Empirical distribution of target movements in the Transitory Treatment (top) and in the Fundamental Treatment (bottom). Movements are expressed in radians; one full circle corresponds to 2π radians, or approximately 6.28 radians.

hence, eventually to cope with them, even if adaptation is not effortless. Anterior insula (AI), and to a lesser extent, the medial wall of the ACC, have long been thought to be crucial for such rapid adaptation, and this capacity has been linked to the presence of an evolutionary recent type of neuron, the von Economo neuron (Allman et al. 2011). To be able to recognize that the uncertainty in one's surroundings is unusual requires appropriate integration of external sensory signals and ensuing bodily reactions (emotions), and fast behavioral adaptation to unfamiliar uncertainty demands continued awareness—2 tasks that evolutionary biologists have attributed to AI (Craig 2009, 2011). The first aim of our study was to challenge the theory that AI facilitates adaptation to highly unfamiliar types of risk, by establishing whether AI activates preferentially in the presence of leptokurtic noise. The emphasis is on "preferentially:" AI is generally engaged whenever outliers or oddballs occur. This is best exemplified by studies showing that AI activation correlates with risk prediction errors, because outliers trigger risk prediction errors—outcomes are larger than expected (Preuschoff et al. 2008; d'Acremont et al. 2009). We set out to determine to what extent, 1) AI manages to discriminate between leptokurtic noise and other types of outliers, even if these are equally salient and frequent, and 2) whether this discrimination is a property observed uniquely in human AI.

In our design, behavioral adaptation to leptokurtic noise was less an issue of belief adjustment, but more one of adaptive control. Specifically, while belief updating played a role immediately after the outlier, in that the decision maker learned about the frequency and size/nature of outliers in the task at hand, subsequent adjustment required the right behavioral response. Importantly, if mistakes were made initially, corrective actions were required subsequently, in order not to stray farther afield. Prior research has focused on outlier detection and belief adjustment (d'Acremont, Fornari, et al. 2013; O'Reilly et al. 2013; McGuire et al. 2014), whereas, in our design, the issue of control came in addition, and subsequent, to outlier detection and belief adjustment. Consequently, a second goal of our study was to determine to what extent, 1) a neural network associated with behavioral control—as opposed to attention—would engage, namely, the fronto-parietal control network (FPCN) (Glaescher et al. 2012; Cole et al. 2013), and 2) whether there is temporal separation between the attentional and control networks, whereby the former would activate first. To better evaluate behavioral responses and corresponding neural activation, we designed a contrasting treatment where outliers did not constitute noise, but instead reflected permanent shifts in the mean, to be referred to as the "Fundamental Treatment." We ensured that the average (unconditional) distribution of movement surprises was identical across the 2 treatments. For this purpose, we forced the permanent shifts of the target to be driven by the same leptokurtic law that generated the noise in the Transitory Treatment. At the same time, the Gaussian driving process of permanent shifts of the target in the Transitory Treatment provided the noise in the Fundamental Treatment. This way, outliers were equally salient and frequent under either treatment; only their nature differed. To wit, in the Transitory Treatment, outliers reflected noise and hence reverted (so the distribution of potential future locations of the target remained unaltered), whereas in the Fundamental Treatment, outliers reflected permanent moves of the target (outliers signaled a shift in the mean of the distribution of potential future target locations). Figure 2*b* displays sample paths of target locations for the 2 treatments. The difference of the sample paths is immediately clear, both qualitatively and quantitatively. Figure 2*c* displays the distributions of the target movement for the 2 treatments. They are clearly leptokurtic. While there were twice as many outliers in the Transitory Treatment because outliers reverted, these reversals could be expected, and hence, half of the outliers did not constitute surprises. So, crucially, the number of "unexpected" outliers was the same; only their effect differed (transitory vs. permanent).

Materials and Methods

Participants

Thirty-one participants took part in the study (12 women and 19 men). They were all students and staff of the California Institute of Technology (Caltech). The median age was 20 years old (min = 18, max = 32). The study took place at the Caltech Brain Imaging Center and was approved by the local institutional review board. The experimenter read the task instructions (see Supplementary Information, Instructions) aloud, checked comprehension, and participants were allowed to practice with one demonstration Run before the functional magnetic resonance imaging (fMRI) scanning.

Task Design

Participants were asked to manipulate a robot in such a way that it came as close as possible to the upcoming position of a target that moved in either direction along a circle (Fig. 2*a*). The target first moved physically around the edge of the circle to its new location. This move took 0.25 s, so that (rare) movements of more than a full circle could be distinguished from those of less than a full circle. Subsequently, participants had a maximum of 2.25 s to move the slider. The slider indicated, between 0 and 1, how far the robot was to move toward the new position of the target, with 0 indicating no move, 1 indicating a complete catch-up with the target, and numbers in between 0 and 1 corresponding to fractional moves. After subject choice, the robot itself moved to its new location, also in 0.25 s. If the target had moved more than a full circle while the robot was at the original position of the target, and the subject subsequently chose to make the robot catch up with the target, then the robot too would move more than a full circle. The total trial duration amounted to 2.75 s. The slider direction changed every 10 trials, adding 2 s to the trial.

Slider direction was changed to balance visual and motor activation across left and right brain hemispheres.

Borrowing from the previous modeling of outliers in computational neuroscience (Yu and Dayan 2003), the target position in trial t (y_t) was determined by a simple hidden state model where the state (x_t) was driven by a state transition shock ξ_t that was independent of the observation error ε_t:

$$y_t = x_t + \varepsilon_t,$$
$$x_{t+1} = x_t + \xi_t.$$

Positions were measured in radians.

Importantly, and unlike in traditional contingency-shift paradigms, the underlying state "changes each trial." Our setting is an extension of the traditional Kalman filter paradigm, where the optimal filter becomes nonlinear (Yu and Dayan 2003; Kitagawa 1987): the learning rate changes with the prediction error in ways we specify below. To appreciate the difference with traditional contingency-shift paradigms, notice that filtering, and hence learning (about the location of the new state), is continuous. The issue our paradigm addresses is whether human subjects understand that, under leptokurtosis, learning intensity (i.e., the learning rate) changes appropriately. In particular, we

introduce 2 Treatments. In one, learning becomes less intense after an outlier, while in the other one, the learning rate is to be increased.

In the so-called Transitory Treatment, the state transition shock ξ_t was Gaussian with mean zero and standard deviation (SD) 0.25 radians while the observation noise ε_t was drawn from a leptokurtic distribution. See Figure 2*b*,*c*, Top. The leptokurtic distribution was obtained as a mixture of 2 zero-mean Gaussian random variables, one with SD 0.25 radians and chosen with probability 0.85, and another one with SD 2 radians and chosen with probability 0.15. The mean and skewness of target moves were both equal to zero; their SD and kurtosis equaled 1.16 and 9, respectively. The leptokurtic noise had a SD and kurtosis of 0.80 and 17, respectively. In the so-called Fundamental Treatment, the distributions of the observation error and state transition shock were switched, keeping the parameters the same: the state transition shock became leptokurtic whereas the observation noise became Gaussian. See Figure 2*b*,*c*, Bottom. The mean and skewness of target moves were both equal to zero; their SD and kurtosis equaled 0.88 and 13, respectively. The state transition shocks had the same SD and kurtosis as the leptokurtic noise in the Transitory Treatment.

In the scanner, participants completed 2 Runs of each of the Treatments. Each Run lasted 200 trials, so participants made choices across 800 trials in total. The Treatment order was counterbalanced between subjects. Participants were informed that target behavior, in particular, the way the target might reverse, would change after the first 2 Runs. The explicit rule governing the target movement in a given run was not revealed and had to be learned.

Participants were paid for accuracy, as follows. In each Run, 3 randomly chosen trials were rewarded with $2. The chance of winning $2 increased with accuracy of prediction of the position of the target in the trial. The chance (in percentage points) was determined as

$$\max\{0, 100 - 0.04(y_t - r_t)^2\},$$

where y_t is the (new) position of the target in trial t, and r_t is the position of the robot going into trial t. Here, position differences between target and robot are measured in degrees rather than radians, to facilitate participant comprehension. The payoff from the rewarded trials was added to a fixed payment of $50. The rewarded trials typically added between $12 and $24 to the fixed payment.

Bayes-Optimal Choice

Optimal adjustment of the position of the robot is equal to the change in belief about the subsequent position of the latent variable x_{t+1}. This latent variable drives permanent changes in the position of the target. The adjustment can be expressed as a proportion of the prediction error, that is, the difference between the new location of the target and the prior location of the robot. Consistent with terminology for Kalman filtering (Kitagawa 1987), we refer to the proportional adjustment as the "learning rate" (also known as the Kalman gain). Unlike in the traditional Kalman filter, however, the learning rate will change with the size of the prediction error. As such, behavioral adjustment was tied to learning rates, which themselves related to prediction errors.

A Bayesian model was developed to determine how the learning rate should be adjusted. To simplify the analysis, we approximated our infinite-range (the target could move an unlimited number of circles), continuous-state (the target could end up anywhere) state-space model by a finite-range, discrete-state Markov transition model. Bayesian posteriors could then readily be obtained using simple matrix multiplications. Details can be found in the Supplementary Information. We then determined the average optimal learning rates and prediction errors per trial category: outlier trials, where the target moved more than 1 SD, nonoutlier trials (the target moved less than 1 SD) and reversal trials (trials following an outlier trial in the Transitory Treatment). (The next section elaborates on the categorization.) Average optimal learning rates and corresponding prediction errors per trial category are indicated with horizontal green line segments in Figure 3*a*,*b*.

The optimal Bayesian learning rate decreases with the size of the prediction error in the Transitory Treatment, and increases in the Fundamental Treatment. See Supplementary Figures 5 and 6, left panels. This is because, in the Transitory Treatment, large moves are most likely leptokurtic noise, and hence, not to be reacted to, whereas in the Fundamental Treatment, large moves are most likely leptokurtic state changes, and hence, need to be followed. Since inference about the cause of a target move is clearest for outliers and for small moves, confidence in the appropriateness of the learning rates is highest for those types of outcomes; confidence in correct adjustment for medium-sized moves is lowest, because the entropy of the posterior belief about the true location from which the target will move next is highest. See Supplementary Figs 5 and 6, right panels.

Our Bayesian model assumes that the decision maker knows the parameters of the generative process. This may be unrealistic. Likewise, we assume that she knows which treatment she is in. The latter is less objectionable: from the Instruction Set (see Supplementary Information, Instructions), participants could readily infer the treatments for Runs 2–4. As to Run 1, because of the salient reversals of outliers in the Transitory Treatment, it takes a Bayesian only a few outliers to determine which treatment one is in (see Supplementary Information).

Model-Free Reinforcement Learning

We developed a learning model that requires neither knowledge of parameters nor treatment, unlike for the Bayesian approach. We modified the model-free reinforcement learning (RL) algorithm in Sutton (1992). The algorithm relies on autocorrelation of prediction errors to adapt the learning rate. It was originally conceived to accommodate contingency shifts. Here, we altered the learning rate adjustment such that, for outliers only, the learning rate is adjusted appropriately relative to a baseline. In the Fundamental Treatment, lack of adjustment of the learning rate would give rise to positive autocorrelation, and the algorithm reacts by increasing the learning rate. In the Transitory Treatment, noise that is not accommodated induces negative autocorrelation (of prediction errors) and, hence, the algorithm decreases the learning rate. The reaction increases with the size of the prediction error. In the Transitory Treatment, this leads to a contrarian strategy, whereby the learning rate is scaled back more the larger the prediction errors. Therefore, we call our RL algorithm the Contrarian RL model.

Defining the error as the distance between the inferred and true target position (estimated and true value of x_{t+1}), the root mean square error of the Contrarian RL model is only marginally worse than that of the optimal Bayesian algorithm, and much better than that of the Sutton RL algorithm, or a simple RL algorithm with fixed learning rate. Importantly, while relative performance of the Contrarian RL Model is worse in the Transitory

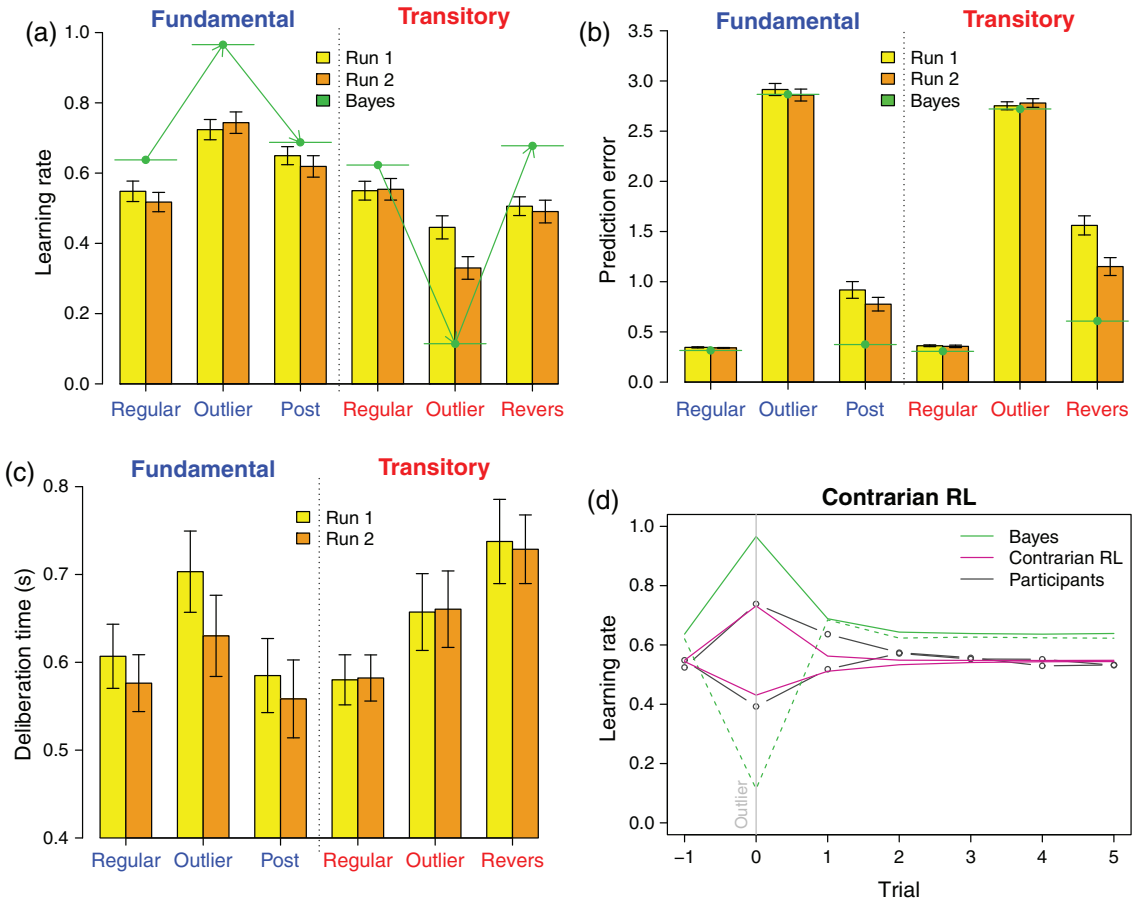

Figure 3. (*a*) Average (±1 standard error [SE]) learning rates (fraction of length of arc spanning pretrial position of robot and new position of target that participant instructs robot to cover; a learning rate of 1 instructs robot to fully catch up with the target). Trials are divided into 3 categories: Nonoutlier (regular) trials; outlier trials; postoutlier trial (Fundamental Treatment; this refers to the trial following an outlier) or reversal trial (Transitory Treatment; this refers to the trial following an outlier, but only if the movement back covers at least the size of an outlier). First Run of 200 trials in yellow; second Run in orange. Green line segments indicate Bayes-optimal learning rates. (*b*) Average (±1 SE) prediction errors (length of arc spanning pretrial position of robot and new position of target), stratified by Treatment, Run, and Trial types. Green line segments show prediction errors from implementing the Bayes-optimal strategy. (*c*) Average (±1 SE) reaction times (in seconds), stratified by Treatment, Run, and Trial types. (*d*) Average learning rates in event time (outlier trial = "0"), stratified by treatment (Transitory: learning rates decrease in outlier trial; Fundamental: learning rates increase upon an outlier). Bayes-optimal model averages: green; Contrarian RL model averages: magenta; participant averages: black.

Treatment, it is much better than that of the Sutton and simple RL algorithms. See Supplementary Figure 8, right panel. To control for the number of free parameters, the bayesian information criterion (BIC) was calculated for the 4 models. As expected, the Bayesian model had the lowest BIC, but the Contrarian RL outperformed both the Sutton and simple RL algorithms. See Reinforcement Learning section in Supplementary Information for details.

Behavioral Analysis

We tracked learning rates, performance, and deliberation times. Learning rates are adjustments of the belief about the true value of the process driving permanent changes in the position of the target. Belief adjustments are revealed by how much participants let the robot catch up with the target. The performance is a function of the prediction error. The prediction error can be measured by the distance between the robot's position and the target's new position. Deliberation times are the length of the interval between target movement and initiation of robot position adjustment.

Trials were divided into separate categories, depending on whether an outlier occurred. Outliers were defined as a target movement greater than 1.18 radian (vertical dashed lines in Fig. 2*c*), which is 1 SD of the distribution of target movements after pooling outcomes in the 2 treatments. The chance of an outlier was about 1 in 10 trials. Regular trials then corresponded to target movements smaller than 1 SD. They form what we shall call the baseline. Outlier trials were trials when an outlier occurred. Postoutlier trials were regular movements (<1 SD) that immediately followed outliers. They were typically found in the Fundamental Treatment. Reversal trials were the signature of the Transitory Treatment. They were large target movements (>1 SD) that followed an outlier and went in the opposite direction.

For each criterion (learning rate, prediction errors, deliberation times), we estimated a mixed linear model, with a factor that identified the type of trial (regular, outlier, postoutlier/reversal). Across-subject variability was accommodated through random-effect regressors for the intercept as well as the trial type factor. This allowed us to test differences between trials using simple *t* statistics; *P* values reported below refer to these *t* statistics. Subject-specific parameters could be obtained by summing the fixed and random effects. More details can be found in the Supplementary Information.

Brain Analysis

Image Acquisition

Blood oxygen level–dependent (BOLD) fMRI acquisitions were performed with a 32-channel head coil on a 3-T Siemens Tim-Trio system. Functional MRI images were acquired with an echo planar imaging gradient echo T_2*-weighted sequence (Flip angle 80°, repetition time = 2000 ms, echo time = 30 ms, 64 × 64 matrix, generalized autocalibrating partially parallel acquisitions acceleration 2, voxel size 3 × 3 × 3 mm, 38 slices, covering the whole brain). Field maps were recorded to correct for head movement. High-resolution morphological data were acquired with a sagittal T_1-weighted 3D magnetization prepared rapid acquisition gradient echo sequence, 176 slices (with voxel size of 1 mm isotropic). This became the structural basis for brain segmentation and surface reconstruction.

Preprocessing

fMRI preprocessing steps were conducted with SPM8 (Wellcome Department of Cognitive Neurology, London, UK), included unwrapping to correct for head movement, normalization to a standard template (Montreal Neurological Institute template, MNI) to minimize interparticipant morphological variability, and resampling to isotropic voxels of 2 × 2 × 2 mm to improve superposition of functional results and morphological acquisitions, and convolution with an isotropic Gaussian kernel (full width at half maximum = 6 mm) to increase signal-to-noise ratio. The signal drift across acquisitions was removed with high-pass filter (only signals with a period <240 s were retained).

Voxel-Based Analysis

A general linear model (GLM) was estimated on the BOLD data using SPM8. Subject was defined as a random factor. The default orthogonalization of regressors in SPM8 was turned off to avoid arbitrary results due to regressor order. Regressors were convoluted with a standard hemodynamic response function to accommodate the typical delay patterns in the fMRI signal. The events defined in the GLM were the target movement (0.25 s), the robot movement (0.25 s), the slider movement (variable duration), and the slider direction switch (2 s). Each event was defined with an onset and a duration variable. Two additional regressors were included to study the early and late effects of outliers. These regressors modulated the effect of the target movement. The early onset regressor takes the value one for outlier trials and the next 2 target movements. Then it reverts to zero (e.g., **1** 1 1 0), unless a new outlier occurs. The late onset regressor is obtained by shifting all the values to the next target movement (e.g., **0** 1 1 1; the boldface value refers to the outlier trial). To control for additional visual effects, the distance the robot moved was included as additional regressor in the GLM. The choice of early/late onset regressors was based on a preliminary analysis, where we had included dummy variables separately for each of the four trials including and following an outlier. The retained regression specification maintains parsimony. Initially, we tested a GLM model with the parametric prediction error as regressor, thus allowing the size of an outlier to modulate the BOLD signal. This produced qualitatively similar results. We only report findings for the dummy variable analysis because the dummy variables allowed the effect of an outlier trial to span over several trials (1 1 1 0) or to be delayed (0 1 1 1).

Altogether, there were 7 regressors per Run in the GLM, and since there were 4 Runs, the GLM used 28 regressors. In addition, we added regressors to capture head movements, as well as Run-specific dummy variables.

ROI Analysis

To avoid circularity or "double dipping" (because the same data were used as in the voxel-based analysis), ROIs for each participant were localized based on the data of all other participants excluding the participant at hand (Kriegeskorte et al. 2009). Average activation in regions of interest (ROIs) was calculated using Marsbar (Brett et al. 2002). The 20 s following an outlier onset were divided in 10 bins of 2 s for the Finite Impulse-Response model. An additional model was defined to estimate the effect of each target movement separately. The beta values obtained with Marsbar served in turn as a dependent variable in mixed linear regressions (R Core Team 2015). The independent variables in these regressions were the early onset regressor (coding for outlier target movements as in the GLM), the treatment (0 = fundamental, 1 = transitory) and their interaction (Outlier × Treatment). Subject was set as a random factor. One mixed linear model was estimated for each ROI. The analysis was repeated by replacing the early onset regressor by the late one. The interaction term allowed us to test if the early and late effect of outliers was significantly different between the Transitory and Fundamental Treatments.

Results

Behavioral Results

In the Fundamental Treatment, the Bayes-optimal strategy is to increase the learning rate as a function of the prediction error. See Figure 3*a*, Left (green line segments). The optimal strategy can best be understood when considering that small moves are likely to be just noise, while large moves reflect permanent changes in the position of the target. Therefore, the robot should be moved closer to the target when the target has just made an outlier move. In the Transitory Treatment, small moves are likely to be permanent position changes, while large moves are most likely reversed subsequently. Therefore, but perhaps counter-intuitively, the robot should not be moved much when the target moves substantially, while small target movements require full robot adjustment. In other words, the learning rate should decrease in the prediction error. See Figure 3*a*, right. In postoutlier trials of the Fundamental Treatment (which, as mentioned before, excludes outlier trials), the learning rate returns close to the baseline. In the Transitory Treatment, the learning rate also returns near the baseline upon reversal (an outlier trial in the other direction).

Consistent with Bayes-optimal choice, in the Fundamental Treatment, the participant learning rate for outliers was higher compared with the baseline ($P < 0.001$) and subsequently decreased significantly ($P = 0.01$, Supplementary Table 1). Still, adjustment in the outlier trials was below the Bayes-optimal policy ($P < 0.001$, Supplementary Table 2). In the Transitory Treatment, participants correctly reduced the learning rate for outlier trials ($P < 0.001$) and then increased it at the time of reversals ($P < 0.001$, Supplementary Table 3). Still, participants overreacted in outlier trials: their adjustment was significantly above the Bayes-optimal level ($P < 0.001$, Supplementary Table 4).

Thus, when controlling the robot, participants distinguished fundamental and transitory outliers and followed the general Bayesian schema (following fundamental changes, resisting transitory changes). However, analysis of runs, prediction errors, and reaction time will show that it is more difficult for participants to adapt to transitory compared with fundamental changes.

Across Runs, we observed an effect of learning during outlier trials in the Transitory Treatment, where the learning rate decreased significantly between the first and second Runs ($P < 0.001$, Supplementary Table 5), bringing participants closer to the

Bayesian optimal solution. No learning effect was detected in the Fundamental Treatment ($P = 0.26$, Supplementary Table 6). This differential learning effect was confirmed by a significant Treatment × Run interaction ($P < 0.001$, Supplementary Table 7).

Prediction errors (the distance between the new target position in a trial and the initial position of the robot) measured performance: the lower, the better. The Bayes-optimal solution generated large prediction errors during outlier trials in both conditions (horizontal green lines, Fig. 3*b*), because outliers could not be predicted. The error decreased in the postoutlier trials, because these involved only regular movements in the Fundamental Treatment, whereas in the Transitory Treatment, perfectly predictable reversals occurred in these trials. In the Fundamental Treatment, participant learning rates mirrored the Bayes-optimal pattern, with significant increases in errors in outlier compared with regular trials ($P < 0.001$), followed by a significant decrease in postoutlier trials ($P < 0.001$, Supplementary Table 8). The increase in outlier trials and decrease in subsequent reversals were also significant in the Transitory Treatment (both $P < 0.001$, Supplementary Table 9). The relative prediction error (computed as the difference between a participant's prediction errors and the Bayesian prediction errors) was larger in the Transitory compared with the Fundamental Treatment for regular trials ($P < 0.001$, Supplementary Table 10) and outlier trials ($P = 0.02$, Supplementary Table 11) alike. It was also larger for reversal trials when compared with postoutlier trials ($P < 0.01$). This suggests that it was more challenging for people to predict the target movement in the Transitory Treatment. Relative to the Bayes-optimal choice, prediction errors in the Transitory Treatment were on average 67% larger than in the Fundamental Treatment (0.786 radians compared with 0.470; Supplementary Table 12).

We now consider prediction errors in absolute terms (not anymore relative to the Bayes-optimal solution). For postoutlier trials, participant errors were smaller in the second Run compared with the first Run ($P = 0.01$, Supplementary Table 13). This effect of training was also observed in reversal trials in the Transitory Treatment ($P < 0.001$, Supplementary Table 14). Importantly, the improvement in performance was more pronounced in the Transitory Treatment, as indicated by a significant Treatment × Run interaction ($P < 0.01$). Quantitatively, training reduced prediction errors in the Transitory Treatment by 27% (1.156 radians compared with 1.585) and by 15% in the Fundamental Treatment (0.777 radians compared with 0.909; Supplementary Table 15).

Deliberation time was defined as the length of time (in seconds) between target movement in a trial and the moment the participant started pulling the slider to adjust the learning rate. Trials where participants did not change their learning rate were not taken into account for the analysis. Compared with regular trials, deliberation time increased for outliers in the Fundamental Treatment ($P < 0.001$) and decreased for postoutlier trials ($P < 0.001$, Fig. 3*c*, Supplementary Table 16). In the Transitory Treatment, deliberation times were significantly higher for outliers compared with regular trials ($P < 0.001$), and further increased for reversal trials ($P < 0.001$, Supplementary Table 17). This suggests that outliers recruited additional cognitive resources, in particular when an outlier reverts. Across Runs, deliberation times decreased in the Fundamental Treatment ($P < 0.01$, Supplementary Table 18), but this was not the case in the Transitory condition ($P = 0.95$, Supplementary Table 19). The latter indicates that with training, behavioral adjustment to fundamental outliers became more automatic, while it remained effortful for transitory outliers.

The Behavioral Results section in Supplementary Information provides detailed analyses of behavior beyond the first postoutlier trial. From the second postoutlier trial on, learning rates, prediction errors, and deliberation times across Transitory and Fundamental Treatments are all much more alike than in the first postoutlier trial. See Supplementary Figure 1. But there are important training effect differences. Specifically, in the Fundamental Treatment, changes in the learning rate upon an outlier settle after about 80 trials, while convergence in the learning rate toward the Bayes-optimal levels lasts several hundred trials in the Transitory Treatment. See Supplementary Figure 2*b*. Individual differences reveal that, for the large majority of participants, the learning rate increased after an outlier in the Fundamental Treatment and decreased in the Transitory Treatment. See Supplementary Figure 2*a*.

The Contrarian RL model has 4 free parameters (a baseline learning rate a, a parameter identifying outliers f, a scaling parameter controlling the effect of covariance of prediction errors s, and a learning parameter updating estimates of variance and covariance of prediction errors θ). We fit it to the participant choices. We computed the error as the difference between the participant choice and the model learning rate. This error was calculated for each of the 12 400 experiment trials (31 participants × 400 trials per treatment). Minimizing the root mean square error, the estimated parameters for the Contrarian RL models were: $a = 0.55$, $f = 0.59$, $s = 1.29$, and $\theta = 0.002$. Figure 3*d* displays the learning rates for the 2 models (Bayesian, Contrarian RL), and the learning rate revealed in participants' choices, per treatment, and across several trials straddling the outlier trials. The Contrarian RL produces learning rates that track those of participants much more faithfully than the Bayesian model. In particular, upon an outlier, both the Contrarian RL model and participants adjusted the learning rate in the correct direction, but this adjustment was smaller in comparison with the Bayes-optimal solution. This under-reaction obtains in both treatments. The smallest BIC was found for the Contrarian RL model (−61 126), followed by the Bayesian model (−55 028), Sutton's algorithm (−43 635), and the simple RL model (−30 851). A formal comparison of model thus confirmed that the Contrarian RL model matched participant choices better than the other 3 models.

Brain Activation Results

Using fMRI, we correlated brain activation with mathematical representations of outlier stimuli (controlling for the usual confounds such as movement or ancillary stimuli).

Absent significant learning effects in brain activation, we merged data from the 2 Runs of each Treatment. We did not find a significant difference in the effect of target movement across Treatments, consistent with our hypothesis that our Fundamental Treatment was a good control for our Transitory Treatment, and hence, that any treatment effect would be picked up by decision-relevant regressors such as type of trial (outlier/nonoutlier). To determine significance, we used a cutoff P level of 0.05, corrected for multiple comparisons at the cluster level [Accepted False Discovery Rate (qFDR) < 0.05] (Chumbley and Friston 2009; Chumbley et al. 2010), unless indicated otherwise. Here, we report activation in brain regions that reacted to outliers.

Exploratory analysis confirmed that some regions reacted at the onset of outliers while others had a delayed response. This analysis also showed that the effect of outliers spanned multiple trials. To formalize the timing effects, the target movement was modulated by 2 regressors in the GLM with which brain activation was analyzed. The regressor corresponding to the early onset equaled 1 in the outlier trial and the following 2 trials, 0 otherwise. The regressor corresponding to the late onset equaled 1 in

the trial following the outlier and the following 2 trials, 0 otherwise.

Areas of significant early activation comprised an extensive bottom-up AN, including thalamus, bilateral PPC (stretching from occipital cortex to the parietal lobes), medial ACC (reaching the supplementary motor area [SMA]), superior part of the precentral gyrus (frontal eye field [FEF]; bilateral), inferior part of the frontal gyrus (bilateral) extending to the (inferior) precentral gyrus, and AI (bilateral), see Figure 4*a*. Detailed imaging results in table format are provided in Supplementary Table 20. Significant late activation upon outliers emerged in bilateral inferior parietal cortex (angular gyrus extending to the supramarginal gyrus) and bilateral middle frontal gyrus (dorsolateral prefrontal cortex; Brodmann areas [BA] 8/9), extending to the frontal pole (BA 10), see Figure 4*b*. This set of regions forms the fronto-parietal network (Supplementary Table 21).

To explore the precise nature of these activations, we conducted a region-of-interest (ROI) analysis, properly adjusting for the fact that we re-visited the same data by localizing the ROI for a given participant using voxel-based activation results for the remaining participants only (leave-one-out procedure). A finite impulse-response model with 10 bins of 2 s was defined to explore the time course of BOLD activity upon outliers. The model was fitted to the mean activity of each ROI separately and the estimated time course was averaged, across the 6 ROIs forming the AN on the one hand (Fig. 4*a*), and the 3 ROIs forming the fronto-parietal network on the other hand (Fig. 4*b*). Results revealed that peak BOLD response to outliers in the fronto-parietal network was delayed by about 4 s (ca. one trial) compared with the AN (Fig. 4*c*).

We then tested whether the early or the late effects of outliers were different between the Fundamental and Transitory Treatments. Results of the GLM contrasts showed no difference for the late onset regressor. Differences were also insignificant when analyzing the average BOLD signal in each of the fronto-parietal ROIs ($P > 0.34$). At a more liberal threshold ($P < 0.001$, uncorrected), a significant difference was found in the bilateral insula for the early onset GLM regressor. See Figure 5 (right) and Supplementary Table 22. For illustrative purpose, the early effect of outliers was computed for the 2 Treatments separately. For the Transitory Treatment, a large activation was observed in the bilateral insula and the medial wall of the ACC (qFDR < 0.05). No voxel was significantly activated in the Fundamental Treatment, see Figure 5 (left). A ROI analysis confirmed that the early effect of outliers was significantly stronger in the bilateral insula ($P < 0.01$; Supplementary Table 23). This significant differentiation already emerged in the outlier trial itself (i.e., before outlier reversal; $P < 0.001$). No significant difference between the 2 Treatments was observed in the remaining 5 ROIs of the AN ($P > 0.26$).

Discussion

We studied human reaction to leptokurtic noise and its neurobiological foundations. Leptokurtic noise is an unusual type of risk

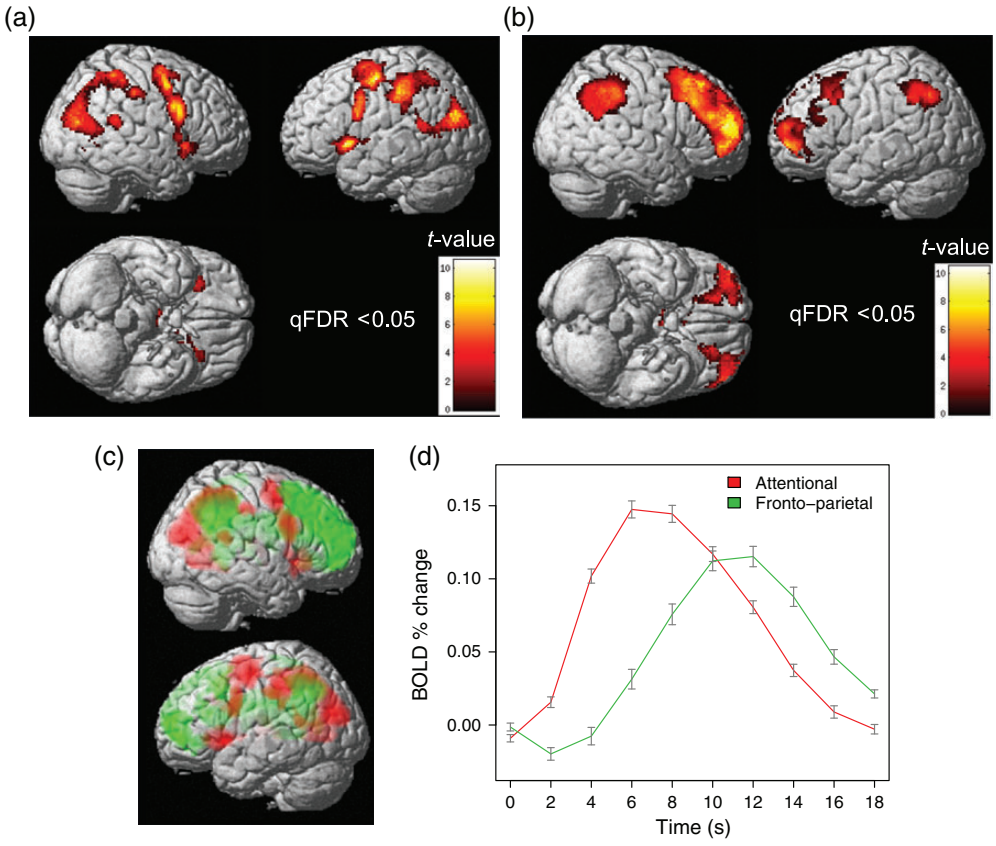

Figure 4. (*a*) Statistical parametric brain maps of significant BOLD activation correlating with outlier trials (early onset), whole-brain corrected, all Runs of both Treatments. Significant activation emerges in a wide network engaged in bottom-up attention to sensory stimuli. (*b*) Statistical parametric brain maps of significant BOLD activation correlating with outlier trials (late onset), whole-brain corrected, all Runs of both Treatments. Significant delayed activation emerges in the fronto-parietal control network. (*c*) Statistical parametric brain maps (left) and time courses (right) of activation in the attentional (red) and fronto-parietal (green) networks correlating with outlier trials, all Runs of both Treatments. Time courses aligned with initiation of target movement (time = 0).

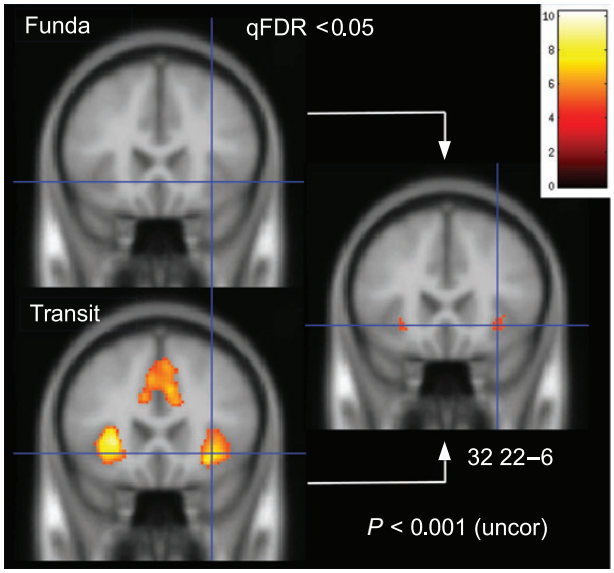

Figure 5. Localization of significant BOLD activation correlating with outlier trials (early onset), separated by treatment. BOLD signals in anterior insula and medial wall of the anterior cingulate cortex were significant (whole-brain corrected) in the transitory treatment (bottom left) while they were not in the Fundamental Treatment (top left). Right: Only anterior insula survived (at $P < 0.001$ uncorrected) direct testing of differential activation across treatments.

that has recently emerged in the human environment as a result of large-scale social interaction through, for example, financial markets. We had 2 goals in mind. First, we wanted to challenge a conjecture in evolutionary neurobiology that the human AI has been shaped, among others, to respond quickly to novel types of uncertainty. We considered leptokurtic noise an excellent example of an unfamiliar kind of risk. Our aim was to determine whether AI played a privileged role in discriminating between, on the one hand, leptokurtic noise, and, on the other hand, outliers that were otherwise equally salient and frequent but that did not constitute noise. AI is known to generally activate in response to outliers, presumably because outliers translate into risk prediction errors, and AI encodes such errors (Preuschoff et al. 2008; d'Acremont et al. 2009). But here we were interested in differential activation, and whether significant differential activation is unique to AI. Second, we wanted to determine whether the FPCN became engaged after outliers that required demanding behavioral adjustment. It is known that outliers activate a broad AN, and that this activation can be separated in an outlier detection component and a belief updating component (O'Reilly et al. 2013). Here, we wanted to investigate neural processes supporting a crucial third stage, namely, behavioral control. We were interested in localization, and in potential temporal separation between (early) attention-related activation and (later) control processes. We targeted FPCN because of its role in tasks requiring behavioral control (Glaescher et al. 2012; Cole et al. 2013).

The behavioral results revealed that participants were indeed unfamiliar with leptokurtic noise: they initially overreacted grossly and performance was poor. Yet consistent with adaptation, behavior improved gradually. Adaptation was not effortless, however, because reaction times remained high, even after several hundred trials, and learning rates took hundreds of trials to converge. This is in contrast to the results we obtained in our version of a paradigm that humans are far more familiar with, namely, the traditional mean-shift paradigm. There, choices were far closer to optimal from the beginning, performance higher, and reaction times shorter, despite the fact that the same leptokurtic law was used to generate outliers. Our results extend prior findings in mean-shift paradigms, demonstrating that these findings continue to obtain even if the random shifts in the mean occur continuously and follow a leptokurtic law. Altogether, our results highlight the difficulty humans have in adjusting their reaction to outliers generated by an increase in error volatility (leptokurtic noise). This difficulty could partly explain why investors and the media sometimes overreact to inconsequential outliers (Fig. 1c).

We found that a Contrarian RL model fit participants' choices better than the Bayes-optimal model. Like an earlier RL model (Sutton 1992), it exploits information in the autocorrelation of outliers, reducing the learning rate below baseline when the autocorrelation is estimated to be negative, and increasing it when the autocorrelation is positive. Negative autocorrelation obtains when the decision maker overreacts, so reduction in the learning rate is called for. Positive autocorrelation obtains when she underreacts, so the learning rate is to be increased above baseline. Unlike the earlier model, however, the Contrarian RL model identifies outliers explicitly and adjusts learning rates only based on observations in outlier trials, and with the learning rate in regular trials as baseline. In contrast with Bayes-optimal choices, the Contrarian RL approach is model-free: it requires knowledge of neither the parameters of the generative model nor identification of which regime one is in (Transitory; Fundamental). It does require one to track autocorrelation of prediction errors. Given how well the Contrarian RL model fits participants' choices, an interesting question for future research is to explore where and how autocorrelation is recorded in the brain.

Our findings elucidate the nature of the neural processes required to properly control behavior after salient and frequent outliers. In prior work, the focus had been on neural processes involved in detecting those outliers, and in updating beliefs. In both Treatments, outliers initially engaged a broad bottom-up AN, including sensory regions such as thalamus and attention-directing regions such as FEF. Within this network, involvement of inferior frontal gyrus (extending to the inferior precentral gyrus) and the PPC could be attributed to the surprise that the outliers generated (Friston 2010; d'Acremont, Schultz, et al. 2013; O'Reilly et al. 2013). Updating of beliefs about the outliers would explain engagement of ACC/SMA, consistent with (O'Reilly et al. 2013). In subsequent trials, activation shifted, tracing the FPCN—inferior parietal cortex and dorsolateral prefrontal cortex extending to the frontal pole (Cole et al. 2013; Glaescher et al. 2012).

Consistent with our conjectures, we thus observed neural dissociation between the attentional/belief updating and control phases of our task. The results highlight how cognition and control are separated neurally, both spatially and temporally. We were able to detect such separation because subjects had to engage in complex behavioral control after an outlier (the robot had to be moved the right distance toward the target). As such, our paradigm had a more prominent control component than, for example, in oddball detection tasks. In traditional contingency-shift paradigms (Behrens et al. 2007; O'Reilly et al. 2013; McGuire et al. 2014), control after an outlier is likewise demanding, but intertwined with extensive belief adjustment (what is the nature of the new contingency?), and belief adjustment sometimes requires exploration (what action would reveal more about the new contingency?), making it difficult to separate belief adjustment and control. In contrast, our task was built on the (nonlinear) Kalman filter, where adjustment of the learning rate is continuous—forecasting the location of the true state should

have taken place in each trial, whether an outlier occurred or not. The distinguishing element was robot control (how far was the robot to be moved?). If anything, forecasting (the location of the state) was simpler in outlier trials than in nonoutlier trials: in the Transitory Treatment, expected future target location (the true state) coincided with target location in the previous trial; in the Fundamental Treatment, expected future target location coincided with the new target location. The entropy (which moves inversely with precision) of the Bayesian posterior beliefs reflects this: the entropy is higher after medium-sized target moves, but lower both after small-sized and large target moves (Supplementary Figs 5 and 6). The intuition is simple: medium-sized target moves could equally likely have been caused by leptokurtic noise or Gaussian disturbances.

Only AI survived a direct contrast between the two Treatments. AI activated more strongly under the Transitory Treatment. Given the unusual nature of leptokurtic noise, our findings support the view among evolutionary biologists that AI has developed to allow humans to rapidly respond in unfamiliar stochastic environments (Allman et al. 2011). This capacity is attributed to the presence of von Economo neurons. Closer inspection revealed that the area within AI where we observed the activation overlapped with the region where von Economo neurons are alleged to reside (Craig 2009). The same subregion of AI has been found to encode risk prediction errors as well (Preuschoff et al. 2008; d'Acremont et al. 2009). The ability of AI to discriminate between leptokurtic noise and other frequent and salient outliers fits well with the notion that AI is involved in building awareness about unusual sensory stimuli from one's environment and emotional reactions to these stimuli (Craig 2009, 2011). As such, AI activation may not simply reflect identification of salient, exceptional events, and may instead reveal an attempt to maintain awareness around outlier trials in the Transitory Treatment, to avoid distraction that would otherwise lower performance significantly. Our hypothesis here is that AI takes control of a person's attention and emotions.

Because outliers were equally frequent and salient across the 2 Treatments, our results suggest that the role of AI is not merely that of bottom-up detection of salient events and subsequent modulation of attention and control (Menon and Uddin 2010). Our results show that AI plays a far more fundamental role. It identifies the nature of salient events, so that responses can be fast and adaptive. AI is thought to engage in tasks that require restraint (Brass and Haggard 2007). To a certain extent, restraint was also needed in our task: under leptokurtic noise, subjects had to keep themselves from following large movements of the target (in contrast, they were to catch up with the target when it made small moves). The role of AI in controlling restraint is controversial though: recent evidence from a stop-signal task suggests that AI is involved only in initial acquisition of restraint, while continued assuring of restraint is attributed to inferior frontal gyrus (Berkman et al. 2014). Moreover, under leptokurtic noise, optimal reaction to outliers was not simply to ignore the outliers (unlike in, e.g., "task-irrelevant" oddball paradigms; Kim 2014), but required careful adjustment of the position of the robot depending on the size of the outlier. Specifically, the slider had to be moved toward the zero point, to reduce the learning rate. Finally, AI is known to signal awareness of upcoming mistakes, perhaps as a result of an increase in noise that could affect performance (Klein et al. 2007; Eichele et al. 2008). But anticipatory activation of this kind is actually located in a more rostral subregion of AI than where we observed our activation (Preuschoff et al. 2008). An interpretation consistent with the localization reported in previous studies would be that the AI activation reflects the higher entropy associated with leptokurtic noise (compare Supplementary Figs 6*b* and 5*b*).

The medial wall of ACC generally activates after outliers in traditional contingency-shift paradigms, and this activation has been associated with belief updating (Behrens et al. 2007; O'Reilly et al. 2013; McGuire et al. 2014). Therefore, it may be surprising that we did not record significant (whole-brain corrected) activation when examining the Fundamental Treatment "in isolation." However, belief updating in our Fundamental Treatment was actually much simpler than in traditional contingency-shift paradigms. As mentioned before, what happened after an outlier did not require elaborate belief adjustment (about the nature of the new contingency). We would offer the following explanation for the strong activation of the medial wall of ACC in the Transitory Treatment in isolation. Our explanation is consistent with the idea that this region is crucial for belief updating, but here belief updating does not concern location of the true state, but identification of the true outcome-generating process. Our subjects exhibited coping problems because they were unfamiliar with a setting where outliers reverted constantly. One can think of humans as favoring a world model where outliers signal fundamental changes and reversals clash with this model. Several results support the idea that fundamental changes are considered "normal," such as the hot hand fallacy (Gilovich et al. 1985), and the general tendency to seek meaningful patterns in random information (apophenia) (Brugger 2001). The belief that outliers should signal fundamental changes could have been exacerbated in our paradigm because our paradigm uses a geometric interface and subjects have been shown to attribute intention to geometric objects that move on a computer screen (Blakemore et al. 2003). In addition, people are particularly prone the hot hand illusion when random series are produced by an intelligent agent, as opposed to a lottery (Ayton and Fischer 2004). Thus, a possible interpretation of ACC activation in the Transitory Treatment (Fig. 5, bottom-left) is that participants' beliefs need substantial revision, and this is what ACC is involved in.

The activation in the medial wall of ACC should also be put in perspective against similar activation in Stroop (Shenhav et al. 2013) and AX-CPT tasks (Carter et al. 1998). ACC activation squares with the idea that this brain region is involved when stimuli are associated with conflicting response. In our Transitory Treatment, a conflict exists between the meaning of the stimulus (the outlier), on the one hand, and subjects' beliefs about how outliers are usually generated (their "world model"). This conflict translated into longer reaction times and larger errors among participants, and goes beyond conflict within the stimulus (Stroop task) or conflict between temporary instruction by the stimulus and usual instruction (AX-CPT task). Conflict had to be resolved by means of adequate control (appropriate moving of the robot), but equally importantly, by means of updates of one's "world model." As such, we could interpret ACC activation as signaling not only conflicting responses, but also of belief updating, in the process merging 2 different opinions about the role of ACC.

We have shown how participants managed to adapt to unfamiliar types of uncertainty. At the same time, substantial effort was needed throughout, reflected in persistently long reaction times. This should caution policies that implicitly force humans to be exposed to leptokurtic noise. The recent tendency to make citizens personally accountable for investments (e.g., through a move from defined-benefit to self-administered, defined-contribution retirement plans) has led to such increased exposure. In fact, risks in financial markets are far more complex than those underlying target movements in our paradigm, because there,

it is never immediately clear whether an outlier constitutes noise or reflects a permanent shift (Mandelbrot 1963; Embrechts et al. 1997). We would advocate the development of new statistical models that adequately identify when outliers constitute noise. Popular financial econometric tools such as GARCH (Bai et al. 2003) do not. Our Contrarian RL model appears to provide a promising direction for future research.

Finally, we conjecture that there may be a link between AI in its role of securing fast, adaptive response to unfamiliar types of uncertainty, on the one hand, and mental disorders that are characterized by lack of adaptation to environmental changes, on the other hand. Activation in AI has indeed been shown to correlate with anxiety disorder and specific phobia, especially social phobia (Etkin and Wager 2007; Damsa et al. 2009). We hypothesize that abnormal function of AI in modulating adaptation to unfamiliar risks may be related to, if not the cause of, some types of anxiety disorder and phobia.

Supplementary material

Supplementary Material can be found at http://www.cercor.oxfordjournals.org/ online.

Authors' Contributions

M.D. helped design the study, did the statistical analysis of behavioral and imaging data, developed the Contrarian Reinforcement Learning Model, helped prepare the manuscript, and wrote the Supplementary Information. P.B. designed the study, helped with data analysis, and wrote the paper.

Funding

This work was supported by the Ronald And Maxine Linde Institute for Economic and Management Sciences at the California Institute of Technology and through US National Science Foundation (grant SES-1061824). Funding to pay the Open Access publication charges for this article was provided by The University of Melbourne.

Notes

Yutaka Kayaba, former PhD and postdoc at Caltech, helped design and run the experiments, and provided preliminary statistical analysis of the behavioral data. The experiments were ran when PB was still at Caltech, and analysis and writing took place during PB's stay at the University of Utah. *Conflict of Interest*: None declared.

References

Allman JM, Tetreault NA, Hakeem AY, Manaye KF, Semendeferi K, Erwin JM, Park S, Goubert V, Hof PR. 2011. The von Economo neurons in the frontoinsular and anterior cingulate cortex. Ann N Y Acad Sci. 1225:59–71.

Ayton P, Fischer I. 2004. The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? Mem Cognit. 32:1369–1378.

Bai X, Russell JR, Tiao GC. 2003. Kurtosis of GARCH and stochastic volatility models with non-normal innovations. J Econom. 114:349–360.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF. 2007. Learning the value of information in an uncertain world. Nat Neurosci. 10:1214–1221.

Berkman ET, Kahn LE, Merchant JS. 2014. Training-induced changes in inhibitory control network activity. J Neurosci. 34:149–157.

Blakemore SJ, Sarfati Y, Bazin N, Decety J. 2003. The detection of intentional contingencies in simple animations in patients with delusions of persecution. Psychol Med. 33: 1433–1441.

Brass M, Haggard P. 2007. To do or not to do: the neural signature of self-control. J Neurosci. 27:9141–9145.

Brett M, Anton JL, Valabregue R, Poline JB. 2002. Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, Sendai, Japan.

Brown SD, Steyvers M. 2009. Insensitivity to future consequences following damage to human prefrontal cortex. Cogn Psych. 58:49–67.

Brugger P. 2001. From haunted brain to haunted science: a cognitive neuroscience view of paranormal and pseudoscientific thought. In: Houran J, Lange R, editors. Hauntings and poltergeists: multidisciplinary perspectives. London: McFarland. p. 195–213.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science. 280:747–749.

Chumbley J, Friston KJ. 2009. False discovery rate revisited: FDR and topological inference using Gaussian random fields. Neuroimage. 44:62–70.

Chumbley J, Worsley K, Flandin G, Friston K. 2010. Topological FDR for neuroimaging. Neuroimage. 49:3057–3064.

Cohen JD, McClure SM, Yu AJ. 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos Trans R Soc, B. 362:933–942.

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. 2013. Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Neurosci. 16:1348–1355.

Craig AD. 2009. How do you feel—now? The anterior insula and human awareness. Nat Rev Neurosci. 10:59–70.

Craig AD. 2011. Significance of the insula for the evolution of human awareness of feelings from the body. Ann N Y Acad Sci. 1225:72–82.

d'Acremont M, Fornari E, Bossaerts P. 2013. Activity in inferior parietal and medial prefrontal cortex signals the accumulation of evidence in a probability learning task. PLoS Comput Biol. 9: e1002895.

d'Acremont M, Lu Z, Li X, Van der Linden M, Bechara A. 2009. Neural correlates of risk prediction error during reinforcement learning in humans. Neuroimage. 47:1929–1939.

d'Acremont M, Schultz W, Bossaerts P. 2013. The human brain encodes event frequencies while forming subjective beliefs. J Neurosci. 33:10887–10897.

Damsa C, Kosel M, Moussally J. 2009. Current status of brain imaging in anxiety disorders. Curr Opin Psych. 2:96–110.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. 2006. Cortical substrates for exploratory decisions in humans. Nature. 441:876–879.

De Bondt WF, Thaler RH. 1990. Do security analysts overreact? Am Econ Rev. 90:52–57.

Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, Cramon DYV, Ullsperger M. 2008. Prediction of human errors by maladaptive changes in event-related brain networks. Proc Natl Acad Sci USA. 105:6173–6178.

Embrechts P, Klueppelber C, Mikosch T. 1997. Modelling extremal events for insurance and finance. Berlin and Heidelberg (Germany): Springer.

Etkin A, Wager TD. 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry. 164:1476–1488.

Fischer M. 2010. Generalized Tukey-type distributions with application to financial and teletraffic data. Stat Pap. 51:41–56.

Friston K. 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci. 11:127–138.

Gilovich T, Vallone R, Tversky A. 1985. The hot hand in basketball: on the misperception of random sequences. Cog Psych. 17:295–314.

Glaescher J, Adolphs R, Damasio H, Bechara A, Rudrauf D, Calamia M, Paul LK, Tranel D. 2012. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proc Natl Acad Sci USA. 109:14681–14686.

Hampton A, Bossaerts P, O'Doherty J. 2006. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J Neurosci. 26:8360–8367.

Hsu DA. 1979. Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis. J Am Stat Assoc. 74:31–40.

Kim H. 2014. Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. Hum Brain Mapp. 35:2265–2284.

Kitagawa G. 1987. Non-Gaussian state-space modeling of nonstationary time series. J Am Stat Assoc. 82:1032–1041.

Klein TA, Endrass T, Kathmann N, Neumann J, Von Cramon DY, Ullsperger M. 2007. Neural correlates of error awareness. Neuroimage. 34:1774–1781.

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. 2009. Circular analysis in systems neuroscience—the dangers of double dipping. Nat Neurosci. 12:535–540.

Mandelbrot B. 1963. The variation of certain speculative prices. J Bus. 36:394–419.

Mandelbrot B, Van Ness J. 1968. Fractional Brownian motions, fractional noises and applications. SIAM Rev. 10:422–437.

McGuire JT, Nassar MR, Gold JI, Kable JW. 2014. Functionally dissociable influences on learning rate in a dynamic environment. Neuron. 84:870–881.

Menon V, Uddin L. 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct. 214:655–667.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI. 2012. Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci. 15:1040–1046.

O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TE, Mars RB, Rushworth MF. 2013. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc Natl Acad Sci USA. 110:3660–3669.

Payzan-LeNestour E, Bossaerts P. 2010. Risk, estimation uncertainty, and unexpected uncertainty: Bayesian learning in unstable settings. PLoS Comput Biol. 6:29–30.

Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP. 2013. The neural representation of unexpected uncertainty during value-based decision making. Neuron. 79:191–201.

Preuschoff K, Marius't Hart B, Einhäuser W. 2011. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. Front Neurosci. 5:115. doi:10.3389/fnins.2011.00115.

Preuschoff K, Quartz S, Bossaerts P. 2008. Human insula activation reflects risk prediction errors as well as risk. J Neurosci. 28:2745–2752.

R Core Team. 2015. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Shenhav A, Botvinick MM, Cohen JD. 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron. 79:217–240.

Sutton RS. 1992. Adapting bias by gradient descent: an incremental version of delta-bar-delta. In: Proceedings of the 10th National Conference on Artificial Intelligence. Cambridge (MA): MIT Press. p. 171–176.

Tervo DG, Proskurin M, Manakov M, Kabra M, Vollmer A, Branson K, Karpova AY. 2014. Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. Cell. 159(1):21–32.

Wolpert RL, Taqqu MS. 2005. Fractional Ornstein–Uhlenbeck Lévy processes and the Telecom process: upstairs and downstairs. Signal Process. 85:1523–1545.

Yu A, Dayan P. 2003. Expected and unexpected uncertainty: ACh and NE in the neocortex. Adv Neur Inf Process Syst. 173–180.

Yu A, Dayan P. 2005. Uncertainty, neuromodulation, and attention. Neuron. 46:681–692.