Hand-washing, particle filter simulations (for Figures 2 and 3)

Particle filters are a set of sampling algorithm used for Bayesian inference for state-space models, which include both observation uncertainty and transition uncertainty (Speekenbrink, 2016). In contrast to the Kalman filter, which provides a analytic Bayesian solution to state space models, particle filters are not limited to Gaussian variables. The reasons for which we chose to use particle filter, rather than a Kalman filter in these simulations are reviewed below.

State-space models describe the evolution of latent states as a first-order Markovian process, where each state depends probabilistically only on the previous state:

|  |  |
| --- | --- |
|  | (S1) |

, where represents a transition uncertainty parameter, which can be variance, or other uncertainty parameters in the case of other distributions. The observation at time t ( depends probabilistically only on the latent state at time t (:

|  |  |
| --- | --- |
|  | (S2) |

, where represents an observation uncertainty parameters, which again can be the variance or other uncertainty parameters.

 A generic particle filter represents estimated latent states with a finite set of particles generating a Monte-Carlo distribution. At each time step, particles are weighted in accordance with Equation S2. This procedure, known as importance sampling (Speekenbrink, 2016) generates an approximate posterior distribution, integrating the prior estimate of the latent state (particle distribution before weighting), and the likelihood distribution (Equation S2). Most particle filters then use a resampling procedure, where particles are resampled in proportion to their weights, resulting in a particle distribution where the relative number (rather than the weights) of the particles represent the relative posterior probability of each particle. This resampling step was shown to be critical to prevent the 'particle degeneration problem'. As a final step, the particle set at time t is propagated in correspondence with Equation S1, to account for transition uncertainty (i.e. process noise).

 We used a particle filter (rather than a Kalman filter) for two reasons. First, although we assumed that the evolution of latent states is Gaussian, we also assumed that the observations are generated from a Gaussian distribution truncated at zero, to represent the idea that any evidence for hands' dirtiness cannot reach negative values (i.e. hands can be perfectly clean at most). Second, to model an affective bias which can be described as the belief that the latent state generating observable information is usually more negative than the observation, we used a skewed normal distribution, with a negative skew. In our case, this corresponds with the belief that one's hands are usually dirtier than they look/feel.

The actual sensory variance, denoted by was set to 50, whereas actual process noise (i.e. transition uncertainty) was set to 0.1. To focus on action-dependent transitions, process noise was set to 0 when no action (i.e. hand-washing) was applied. The values of observation noise and process noise are important only in relation to one each other. Sensory noise was set to a considerably higher value than process noise to represent the assumption that sensory feedback regarding hand dirtiness is in most cases highly non-informative (except for relatively infrequent cases were one can actually see dirt on one's hands). Whereas hand washing changes hand dirtiness, we assumed that it does so in a relatively predictable fashion; that is: we assumed that hand washing is a more informative cue than sensory information regarding hand dirtiness.

Finally, as a simple response model, the probability with which agents were assumed to wash their hands was determined in accordance with the proportion of particles surpassing a specific criterion, which represents a subjective criterion for "how clean does one expect/need one's hands to be". Hand washing was assumed to decrease the actual hand dirtiness by a (arbitrary) factor of 0.8. A multiplicative rather than additive factor was used, to account for the idea that hand washing can never yield perfectly clean hands, and that the effects of hand washing decrease for cleaner hands. The only modeling assumption critical for the results of the simulation is that of representing hand washing as a sequential Bayesian inference, whereas some of the more peripheral assumptions described above could be relaxed with no marked effect on the results of the simulation.

 We now describe the particle filter algorithm used for these simulations. In this algorithm, represents the level of hand dirtiness at time t, estimated by a set of P particles: . At each time t, the weight of each particle is denoted by . corresponds with the probability of washing one's hands at time t. The model included 3 free parameters. First, *b* corresponded with the ratio of the estimated state uncertainty ( to the actual state uncertainty (, and it was set to 1 for simulated control subjects, and to 100 to simulated OCD patients (see Figure 2A vs 2B, respectively, for results). corresponded with the cleanliness criterion, and it was set to 5 in all but the 'perfectionism' simulation, where it was set to 1 (see Figure 3A for results). Finally, corresponds with the skew parameter of the skewed normal distribution, and it was set to 0 (corresponding to a Gaussian distribution) for all but the 'affective bias' simulation (see Figure 3B for results). 1000 particles were used for the simulations. In the algorithm below, asterisks correspond with changes in the real world (rather than in the Bayesian estimate thereof):

1. **\*(Actual state) (initial level of dirtiness)**
2. **(Initialize) sample P particles from**
3. **For t=1….T**
	1. **\*(Action)**
	2. **\*(Actual state)**
	3. **\*(Observation) sample a random observation from**
	4. **(Propagate) with**
	5. **(Weight) Weight particles in accordance with with**
	6. **(Resample) Resample particles in accordance with**

The free energy framework

 The free energy framework has been suggested as a unified theory of cognition and action, and it can be viewed as proposing a specific mathematical framework for conceptualizing and quantifying predictive processing in the brain. Interested readers can found detailed theoretical reviews (Friston, 2010), and mathematical tutorials (Bogacz, 2017; Buckley, Kim, McGregor, & Seth, 2017) elsewhere. Here we briefly describe the very basic mathematical principles relating the free energy framework to the Bayesian brain, and predictive coding.

 In most realistic cases Bayesian inference requires non-analytic approaches. The particle filter presented above is an example of one type of solutions which use sampling techniques to approximate Bayesian inference. Whereas sampling provides asymptotically correct solutions with infinite sample size, finite samples provide an approximation of the exact solution. This weak point of sampling techniques has been suggested to make it appropriate to explain the often non-optimal cognition (Sanborn, 2017).

 Another set of solutions provide tools for Bayesian inference in complex scenarios by approximating the true posterior distribution by a simpler distribution. This class of solutions, called variational Bayes, converts the Bayesian inference problem to an optimization problem, focusing on finding the distribution that best approximates the true posterior. Specifically, assuming the exact Bayesian posterior is:

|  |  |
| --- | --- |
|  | (S3) |

In variational inference, p(x|o) is approximated by another, simpler distribution q(x), with the goal of minimizing the dissimilarity between these two distributions. The most common measure of dissimilarity is the Kullback-Leibler divergence, defined as:

|  |  |
| --- | --- |
|  | (S4) |

Note than when q(x) equals p(x|o), their ratio is equal to 1, and its logarithm, as well as the entire expression is equal to zero.

 Critically, solving the KL divergence is still impossible, because it requires one to know the posterior distribution, p(x|o). However, one can substitute this term by its definition above (Equation S3), which gives:

|  |  |
| --- | --- |
|  | (S5) |

, where the transitions between the second and third lines were based on logarithm product rule; the transition from the third to the fourth line was based on integral rules; and the transition from the fourth to the last line was based on the fact that is a constant and thus can be pulled out the integral, whereas the integral of q(x) is 1, because it is a probability distribution.

 The integral in the last line if Equation S5 is called *free energy* (*F*), such that Equation S5 can be expressed as:

|  |  |
| --- | --- |
|  | (S6) |

Now, recall that our objective here is to minimize the KL divergence. Because is a constant with respect to q(x), instead of minimizing the KL we can simply minimize the free energy. Thus, in variational Bayes, approximate Bayesian inference is done by minimizing free energy. This explains why according to the free energy framework, the (Bayesian) brain acts to minimize free energy.

 Note further that we can rearrange the terms in Equation S6, to:

|  |  |
| --- | --- |
|  | (S7) |

 Because the KL divergence is, by definition, always positive, the negative free energy is also the lower bound for . That means, that:

|  |  |
| --- | --- |
|  | (S8) |

Equation S8 also serves a critical role in the free energy framework, because the term on its right side is called surprisal in information theory, and it quantifies the extent to which observations are novel to the agent (i.e. observations with low probability are most surprising). This means, that by minimizing free energy one affectively minimizes surprisal. Indeed the key theoretical assumption of the free energy framework is that all living organisms strive to minimize surprisal, which ultimately stands for staying in the homeostatic boundaries allowing for their existence (e.g., for a fish, being out of the water is a highly surprising observation).

In terms of cognition, q(x) can be thought of as beliefs regarding the true state of the world x. Minimizing variational free energy thus minimizes the difference between one's beliefs and observed outcomes interpreted in light of a probabilistic generative model with assumptions regarding how observations and latent states covary:

|  |  |
| --- | --- |
|  | (S9) |

In active inference, the selection of policies (or actions) is also assumed to be guided by free energy minimization. However, here agents are assumed to strive to minimize their expected free energy (rather than the current free energy), because this process is assumed to occur during planning. Mirroring Equation S9, expected free energy, given a specific policy can be defined as:

|  |  |
| --- | --- |
|  | (S10) |

Note that the expectation here is with respect to a ‘posterior predictive’ distribution , so that we take expected future observations into account. Assuming that the agent's preferences (prior expectation) regarding outcomes do not depend on the selected policy, the expected free energy becomes:

|  |  |
| --- | --- |
|  | (S11) |

This formula weights the epistemic value, and the extrinsic value of a policy. The second line shows that, if we approximate , we can rewrite the epistemic value as an expected KL-Divergence (or information gain). The third line rewrites epistemic value involving two terms. First, the expected free energy of a policy decreases as the state transitions become more uncertain, as they lead to a greater uncertainty about future outcomes, or a greater entropy of the predictive distribution . Thus, for example, when a certain policy is believed with high certainty to bring in a certain state, is low, signaling that following that policy provides less knowledge, because there is limited uncertainty to resolve. Second, the expected free energy decreases further, as the observations are more informative, that is – as they disambiguate one's prediction regarding the consequences of a policy. The ambiguity associated with a given policy may be expressed as the entropy of the likelihood (first term on line 3 of Equation S11), as this expresses the fidelity with which states map to outcomes. In sum, policies related with states characterized by relatively high state uncertainty and relatively low observational noise have a higher epistemic value. Finally, the extrinsic value of policies is formalized in active inference in accordance with the prior distribution over observations. That is, desired outcomes are defined as observations with high prior probability, such that policies expected to result in desired outcomes have a lower expected free energy.

Generative model for the active inference simulations

The MDP scheme used in our simulations closely resembled the one specified in (Parr & Friston, 2017), with the addition of habit learning, represented as the prior probability over policies (), which is learned over time. Thus, policy selection is a softmax function () of these prior beliefs, and the expected free energy of each policy (see above), with the inverse temperature parameter :

|  |  |
| --- | --- |
|  | (S12) |

As described above, the expected free energy of a policy is a function of beliefs regarding state transitions for that policy, beliefs regarding the state-outcome mapping, prior preferences regarding outcomes, and prior belief regarding the initial state, defined respectively as:

|  |  |
| --- | --- |
|  | (S13) |

 In the current scheme, the outcome and state space are each defined in terms of two factors. That is, the state space is factorized into hidden states defining *position* (e.g., where did the agent 'click'), and hidden states defining the *context* (i.e. the correct cue). Whereas the transition matrix (B) for position states depends on the action *u*, the transition matrix for context states is defined irrespective of action. Specifically, the probability of each context to persist on the next time step is parameterized by the parameter *b*, which is at the focus of the current simulations. It is important to note that *b* here refers to the *estimated* transition uncertainty (i.e. from the agent's perspective). The actual probability for state transitions was 0 (with the exception of one, predefined transition at the reversal trial).

The outcome space is factorized into two factors: exteroceptive outcomes, that are informative about the agent's position, and affective (which could be defined as interoceptive) outcomes which could be a reward, a punishment or a neutral affective outcome. Each hidden state (in terms of both position and context) is mapped to an outcome. The exteroceptive outcome (position) is a deterministic (identity) function of the position state, whereas the affective (interoceptive) outcome is a probabilistic function of the joint position-context state. Here we assume that the probability for a reward, given that the agent has chosen the correct cue (e.g., choosing the left cue in the context that the left cue is correct), is defined by *a*, whereas the probability of reward for each of incorrect cues is 1-*a*.

The D vectors (i.e. initial states) were parameterized with Dirichlet priors. The agent beliefs each trial starts from the initial/baseline position (position states factor). In the first trial, the agent has a relatively non-informative (symmetric Dirichlet, with the concentration parameter ) prior regarding context (context states factor), and this belief is updated on each trial, based on the generative model and observed outcomes. Policies were defined only in terms of position states, because the agent had no direct influence on context states. Correspondingly, habits were learned only for position related policies. Habits were also parameterized in terms of a symmetric Dirichlet prior, where the concentration parameter defines the amount of prior certainty in a given policy (see Figure 4). Finally, the C vector (i.e. preferred outcomes) was set to include preferences for affective outcomes, and no preferences for position outcomes (i.e. the agent prefers to get rewards, but has no preference regarding its position per se).

**Specific generative model for task 1.**

In task 1 there were 4 hidden position states, and 3 hidden context states. The transition matrix for position, for each policy, was defined such that each action (in this task policies were of length-1) leads directly to the respective position. For example, for policy 'L':

Here, the transitions occur from state represented by each column to those represented by rows, with states being: neutral/initial position (N), left position (L), right position (R), and middle position (M). The transition matrix for context, independent of action, was defined as (Here L, R and M refer to context and not to the agent's position) :

The likelihood matrix for position () for each context was a simple identity matrix (implying a deterministic relationship between position states and observed positions). The likelihood matrices for affective outcomes prescribe the probability for a reward given the position and context states. So for example, when the left cue is correct, the likelihood matrix mapping position states (columns) to affective outcomes (rows) is (note that the ‘L’, in the left term, here refers to the context in which the left cue is correct – not to the policy ‘choose left cue’, as in the transition matrix above):

**Specific generative model for task 2.**

In task 2 there were 4 hidden position states, and 2 hidden context states. The transition matrix for position, for each policy, was defined such that each action (in this task policies were of length-2) leads directly to the respective position, with the exception of the utilitarian cue positions (left/right) being *absorbing states*, such that once entered (in a given trial) the agent stays in this position until the end of the policy (see Friston et al., 2016). Thus, for example, for policy 'L':

The transition matrix for context, independent of action, was defined as:

In this task, the exteroceptive likelihood matrix included information regarding position, as well as information signaling context when the agent was at the 'checking' (*C*) position state. So, when the left cue is correct, the exteroceptive likelihood matrix is:

, and when the right cue is correct it becomes:

The likelihood array for affective outcomes describes the probability for a reward given the position and context states. Here the affective outcome of both the initial and the checking positions is neutral (defined as 0 in the C matrix). So for example, when the left cue is correct, the likelihood matrix mapping position states (columns) to affective outcomes (rows) is:

 Note also, that in the second task *a* was set to 1 (only deterministic state-outcome contingencies), because otherwise the outcome of checking behavior should also have been defined as probabilistic, adding a currently unnecessary level of complexity.

**References**

Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, *76*, 198–211. https://doi.org/10.1016/j.jmp.2015.11.003

Buckley, C. L., Kim, C. S., McGregor, S., & Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology*, *81*, 55–79. https://doi.org/10.1016/j.jmp.2017.09.004

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O’Doherty, J., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience and Biobehavioral Reviews*, *68*, 862–879. https://doi.org/10.1016/j.neubiorev.2016.06.022

Parr, T., & Friston, K. J. (2017). Uncertainty, epistemics and active inference. *Journal of The Royal Society Interface*, *14*(136), 20170376. https://doi.org/10.1098/rsif.2017.0376

Sanborn, A. N. (2017). Types of approximation for probabilistic cognition: Sampling and variational. *Brain and Cognition*, *112*, 98–101. https://doi.org/10.1016/j.bandc.2015.06.008

Speekenbrink, M. (2016). A tutorial on particle filters. *Journal of Mathematical Psychology*, *73*, 140–152. https://doi.org/10.1016/j.jmp.2016.05.006