**Supplementary File 3. Priors distributions over model parameters**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Prior distribution** | **Notes** |
| Unisensory Lapse rate () |  | Weakly informative prior preferring lower lapse rates |
| Unisensory Lapse bias () |  | Weakly informative prior preferring unbiased lapses |
| Lapse rate in implicit and explicit causal inference tasks () | ]  where  and | Hierarchical: This prior allows the model to hierarchically infer a lapse rate that is consistent across experiments unless the data suggests otherwise |
| Lapse bias in implicit & explicit causal inference tasks () | ]  where | Hierarchical (see above) |
| Auditory and visual location prior mean () |  | Weakly informative prior preferring no perceptual bias |
| Auditory location sensory noise standard deviation () |  | Chosen to explore the full range of psychometric curves. The parameters were chosen by finding values of for limiting psychometric curves (shallow and steep) for the range of values used in experiments. The limiting values were translated to parameters of the lognormal distribution such that 90% of the mass lied within the range. We verified that the results were insensitive to the exact value of the prior by varying the mean and standard deviation by a factor of 0.5 and 2 and verifying that the parameter estimates were quantitatively similar |
| Auditory and visual location prior standard deviation () | where | While combining the prior information about the cue locations with the likelihood, the weight on the prior/likelihood depends on the ratio of the likelihood and prior variances (refer and in Eq. S2). In such a case, to ensure that the same range of are a-priori likely; independent of the specific value of , the prior over the auditory prior is varied around in logarithmic space |
| Visual location sensory noise standard deviation (one parameter for each reliability r) | where | Same motivation as for the auditory and visual prior (see above) as the when combining the auditory and visual cues, the weights given to either cue depends on the ratio of their relative variances. (refer in Eq. S2). |
| pcommon |  | Since this is the main parameter of interest, we use a flat prior to prevent any biases due to the prior |
| Choice bias (auditory discrimination) |  | Weakly informative prior preferring unbiased responses |
| Choice bias (visual discrimination , multisensory discrimination , implicit causal inference ) | ]  ]  ]  where | Hierarchical: Since all the mentioned tasks are discrimination tasks, this prior allows the model to hierarchically infer a choice that is consistent across experiments unless the data suggests otherwise |
| pcombined(explicit causal inference) |  | Since this is the main parameter of interest, we use a flat prior to avoid biases due to the prior |