Supporting Information

Solway and Botvinick 10.1073/pnas.1505483112

SI Materials and Methods

Participant Selection. Thirty participants were targeted for each experiment. As described in the main text, each experiment consisted of two parts. The first involved rating items and the second making decisions about them. To progress to the second part, participants had to rate a minimum number of items at each rating level (although they were not told about this requirement, discussed below). Recruitment continued until there were 30 valid participants for each experiment. Overall, 31 participants were recruited for experiment 1, and all progressed to the second part. One participant was dismissed with credit before completing the experiment because they did not finish 30 min after the allotted time. Thirty-six participants were recruited for experiment 2, and 30 of them progressed to the second part.

Participants had to rate at least 10 items at each rating level for experiment 1 and 12 for experiment 2. Participants were not told of this specific requirement beforehand but were instructed to give relative ratings by considering each item in the context of all of the other items on the list, and to try to use all of the numbers during the rating period. Participants who did not meet this requirement were dismissed with two hours' worth of compensation. This requirement was enforced to reduce carryover effects between consecutive trials. Because five items were displayed on each trial of the decision phase in experiment 1, and each could be associated with the same rating, having 10 items at each rating level ensured that the same item would not have to appear on two consecutive trials. This same reasoning applies to experiment 2, which had six items per trial.

Item Selection and Experiment Duration. The set of items for each participant was trimmed to have the same number for each rating level. For example, if 90 items were rated 0, 50 were rated 1, 30 were rated 2, 50 were rated 3, and 50 were rated 4, the 30 items rated a 2 were all kept together with a random subset of 30 items from each other rating level. This helped put items associated with different ratings on more equal footing.

Experiment 1. The items on each trial were pseudorandomly chosen such that (*i*) the same item did not appear on consecutive trials and (*ii*) the difference between the value of the best path and the average of the values of the other two paths was between 1 and 5, in increments of 0.5. The latter is the primary measure of trial difficulty described in the main text, and value is defined as before as the sum of the ratings in a single path of the tree. For example, if the rating of the item on the left top was 4, left bottom left 5, left bottom right 3, right top 1, and right bottom left 5, the values were computed to be left top + left bottom left = 9, left top + left bottom right = 7, and right top + right bottom left = 6. The difficulty of the trial in this case was $9 - \text{mean}(6, 7) = 2.5$. Each participant completed 900 trials, 100 at each difficulty level, with four unlimited breaks offered evenly spaced throughout the experiment.

Experiment 2. Items were similarly chosen on each trial, with difficulty measured based on four pairs of items instead of three. Difficulty levels between 1 and 6 were included, rounded to the nearest integer. Seventeen participants completed 660 trials, 110 at each difficulty level, and the remaining 13 participants completed 780 trials, 130 at each difficulty level. Both groups were offered four unlimited breaks evenly spaced throughout the experiment. The number of trials was increased to reflect the amount that fit within a 2-h window. Note that in Fig. 4 difficulty level is rounded to the nearest 10th decimal place rather than to the nearest integer.

Statistical Analysis. Second-stage reaction time effects were analyzed by log-transforming the data, computing the mean for each participant and absolute difference in ratings, and subjecting the results to a repeated measures ANOVA.

Computational Models. We present the modeling details in terms of the second experiment. However, the implementation of the first experiment is exactly the same, with one fewer item (and path) being compared.

Two stages with independent paths (primary model). The primary model treats each path through the tree as an independent competitor. The evidence for each begins at zero and on each iteration of the deliberation process is updated according to the rewards (ratings) along that path:

$$\begin{aligned}
E^{t+1}_{L,L} &= E^{t}_{L,L} + (d_1 \cdot R_L + \epsilon) + (d_1 \cdot R_{L,L} + \epsilon),\\
E^{t+1}_{L,R} &= E^{t}_{L,R} + (d_1 \cdot R_L + \epsilon) + (d_1 \cdot R_{L,R} + \epsilon),\\
E^{t+1}_{R,L} &= E^{t}_{R,L} + (d_1 \cdot R_R + \epsilon) + (d_1 \cdot R_{R,L} + \epsilon),\\
E^{t+1}_{R,R} &= E^{t}_{R,R} + (d_1 \cdot R_R + \epsilon) + (d_1 \cdot R_{R,R} + \epsilon).
\end{aligned} \quad \textbf{[S1]}$$

E refers to the amount of evidence associated with a single path. For example, $E_{L,L}$ is the evidence for the path corresponding to going "left" and then "left" again. R is the reward associated with a particular position, with R_L and R_R referring to the left and right rewards at the first stage, $R_{L,L}$ to the left reward at the second stage after going left at the first stage, and so on. ε is a Gaussian random variable with mean 0 and SD 0.01, and d_1 is a free parameter. A decision is made when the difference between the largest integrator and the next largest integrator ("best vs. next") exceeds threshold (θ_1, a free parameter).

Although the model thus far is able to render a decision for both stages of action (the winning path implies what to do at both stages), it predicts a constant reaction time for the second stage. This contrasts with Figs. 3*E* and 4*E* in the main text, which show an effect of difficulty. To capture this effect, we amend the model to output only the first-stage decision and allow the remaining paths to continue integrating until the difference between them exceeds another threshold. For example, if the right side was selected during the first stage, the second stage continues according to

$$\begin{aligned}
E^{t+1}_{R,L} &= E^{t}_{R,L} + d_2 \cdot R_{R,L} + \epsilon,\\
E^{t+1}_{R,R} &= E^{t}_{R,R} + d_2 \cdot R_{R,R} + \epsilon.
\end{aligned} \quad \textbf{[S2]}$$

d_2 is a free parameter, and a decision is made when the difference between the larger and smaller integrator exceeds θ_2.

Each stage of action also has an additional parameter, T_1 and T_2, respectively, specifying the amount of nondecision time. This includes time spent processing the stimuli and issuing a response.

Two stages with correlated paths. The correlated paths model is exactly the same as above, except the noise term (ε) is correlated for R_L and R_R. That is, on each iteration R_L is sampled once and the sample contributes to both $E_{L,L}$ and $E_{L,R}$, and likewise for R_R.

Two stages with pruning. Each of the two models above was also instantiated with a second decision rule, which says that a first-stage decision can be made when either

$$min(E_{L,L}, E_{L,R}) - max(E_{R,L}, E_{R,R}) \geq \theta_{prune}, \quad \textbf{[S3]}$$

or

$$min(E_{R,L}, E_{R,R}) - max(E_{L,L}, E_{L,R}) \geq \theta_{prune}. \quad \text{[S4]}$$

θ_{prune} is an additional free parameter. A decision is made as soon as either this rule or the max-vs.-next-best rule applies, whichever occurs first.

Two-stage average. Rather than individual paths competing at the first stage, this model posits a single integrator for each first-stage action, with its value incremented according to the average of the (noisy) rewards below:

$$E_L^{t+1} = E_L^t + (d_1 \cdot R_L + \epsilon) + \left[\left(d_1 \cdot R_{L,L} + \epsilon\right) + \left(d_1 \cdot R_{L,R} + \epsilon\right)\right]/2,$$
$$E_R^{t+1} = E_R^t + (d_1 \cdot R_R + \epsilon) + \left[\left(d_1 \cdot R_{R,L} + \epsilon\right) + \left(d_1 \cdot R_{R,R} + \epsilon\right)\right]/2. \quad \text{[S5]}$$

A decision is made when the difference between E_L and E_R exceeds threshold. This simplification comes at a cost. First, to capture the second-stage reaction time effect, these samples also contribute to integrators for second-stage items (which compete after the first stage ends, as in the models above), resulting in an increase in the total number of integrators. Second, the model is suboptimal on trials like the one shown in Fig. S6, which we call "max/mean conflict" trials. On such trials, the overall maximum path is on one side of the tree, whereas the average of the pairs of paths is higher on the other side. For example, in the tree shown in Fig. S6, the best path $(5 + 5 = 10)$ is on the right, but the average of the two paths on the right $[5 + (5 + 1)/2 = 8]$ is less than that on the left $[5 + (4 + 4)/2 = 9]$. The model prefers the side with the maximum average rather than the maximum overall path.

Considering the average best may be a useful heuristic when the decision tree is deep and only one or two steps need to be executed quickly. Although our experiments did not directly address this scenario, and max/mean conflict trials make up a small portion of our data (less than 3% of experiment 1 and less than 2% of experiment 2), we performed an exploratory analysis fitting this and the primary model not only to the data discussed in the main text, but also to first-stage choice accuracy and reaction time as a function of trial type (max/mean conflict vs. no max/mean conflict). The baseline model provided a more parsimonious fit for both experiments (experiment 1, primary model BIC −277, average model BIC −243; experiment 2, primary model BIC −464, average model BIC −415).

Forward greedy search. In forward greedy search, items at each stage compete independently. The top-level items compete first:

$$E_L^{t+1} = E_L^t + d_1 \cdot R_L + \epsilon,$$
$$E_R^{t+1} = E_R^t + d_1 \cdot R_R + \epsilon. \quad \text{[S6]}$$

A decision is made when the difference between the integrators exceeds θ_1. The remaining items compete at the second stage. For example, if the right side was chosen, the items on the right compete as in Eq. **S2**, but both start with zero evidence. Each stage of action also has an additional parameter specifying the amount of nondecision time, as in the models above.

Backward induction. First, items at the second level compete in parallel:

$$E_{L,L}^{t+1} = E_{L,L}^t + d_0 \cdot R_{L,L} + \epsilon,$$
$$E_{L,R}^{t+1} = E_{L,R}^t + d_0 \cdot R_{L,R} + \epsilon,$$
$$E_{R,L}^{t+1} = E_{R,L}^t + d_0 \cdot R_{R,L} + \epsilon,$$
$$E_{R,R}^{t+1} = E_{R,R}^t + d_0 \cdot R_{R,R} + \epsilon. \quad \text{[S7]}$$

A decision is made on the left when one integrator exceeds the other by θ_0, and likewise on the right. This contributes the length of the longer of the two competitions to the first-stage reaction time. The top level items then enter the competition:

$$E_L^{t+1} = E_L^t + d_1 \cdot R_L + \epsilon,$$
$$E_R^{t+1} = E_R^t + d_1 \cdot R_L + \epsilon. \quad \text{[S8]}$$

The integrators for the second-level items remain frozen during this time. A decision is made when the sum of the integrators for the left top-level item and the winning left second-level item exceeds the sum of the integrators for the right top-level item and the winning right second-level item, or vice versa, by θ_1. In other words, when

$$\left|\left[E_L + max(E_{L,L}, E_{L,R})\right] - \left[E_R + max(E_{R,L}, E_{R,R})\right]\right| > \theta_1. \quad \text{[S9]}$$

The model thus far does not predict a reaction time effect at the second stage. To allow the model to capture this effect, the second-stage integrators are unfrozen and continue integrating (with drift d_2, until the difference between them exceeds θ_2). As above, each stage of action has an additional parameter specifying the amount of nondecision time.

We also fit a version of the model where d_1 was constrained to equal d_0 and θ_1 was constrained to equal θ_0.

Backward induction with reset. This version of the backward induction model begins the deliberation process according to Eq. **S7**. However, the losing items are then pruned away, and evidence for the winning items is reset. The remaining branches compete in parallel, with evidence accruing for items at both levels:

$$E_L^{t+1} = E_L^t + (d_1 \cdot R_L + \epsilon) + (d_1 \cdot R_{LW} + \epsilon),$$
$$E_R^{t+1} = E_R^t + (d_1 \cdot R_R + \epsilon) + (d_1 \cdot R_{RW} + \epsilon). \quad \text{[S10]}$$

R_{LW} is the rating associated with the winning item at the second level on the left, and likewise for R_{RW} and the winning item on the right. These two competitions constitute the first-stage deliberation. The second-stage deliberation proceeds as in the above model.

One-stage parallel integration with vigor. The one-stage parallel integration model is similar to the primary model during the first stage and proceeds according to Eq. **S1**. Upon choosing a path, the first-stage reaction time is further incremented by $vigor_1 \cdot (R_{max} - R_{1W})$. $vigor_1$ is a free parameter, R_{max} is the maximum rating (4 in experiment 1 and 5 in experiment 2), and R_{1W} is the rating associated with the top-level item of the winning path. No deliberation takes place during the second stage. The second-stage reaction time is incremented by $vigor_2 \cdot (R_{max} - R_{2W})$, where $vigor_2$ is a free parameter and R_{2W} is the rating associated with the second-level item of the winning path. Each stage has an additional parameter that captures the sum of the time spent looking at the stimuli and the intercept term for the motor response.

One-stage parallel integration with vigor and rating noise. This model is exactly the same as the model above, except that the reward associated with each item is sampled once at the beginning of the trial from a Gaussian distribution centered on the item's rating. For example, reward for the top left item is sampled according to

$$R_L^{'} \sim \mathcal{N}(R_L, rating_{sd}), \quad \text{[S11]}$$

and similarly for the other items. $rating_{sd}$ is a free parameter. The sampled ratings are used in place of the actual ratings in Eq. **S1**.

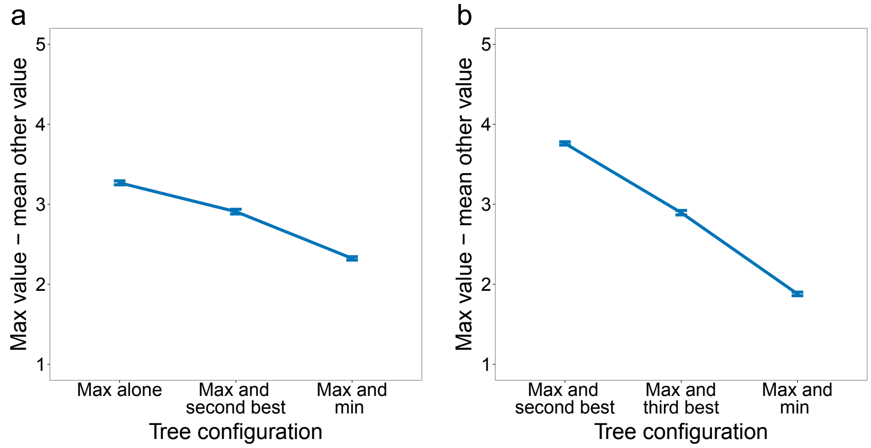

Fig. S1. The difference between the value of the best path and the average of the other paths' values as a function of which paths appear together in the tree. Decision difficulty differs across different tree configurations as a result of the two-stage structure. For example, consider the "Max and min" case vs. the "Max and second best" case. Because paths are correlated at the first stage, the value of the minimum path is on average closer to the maximum in the former compared with the latter. However, the second-best path in the former, which is on the other side, has to be even higher and is even more correlated. This makes the decision harder on average. (*A*) Experiment 1. (*B*) Experiment 2.

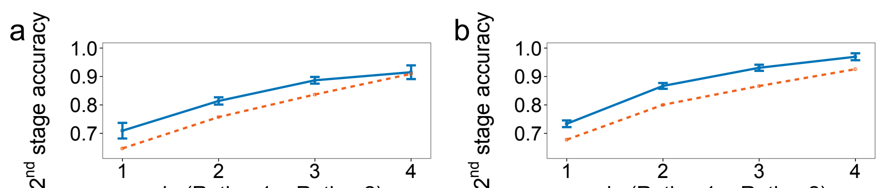

Fig. S2. Simulation of both experiments using the best-fitting parameters of the one-stage parallel integration with vigor model. The model underestimates second-stage choice accuracy due to the lack of deliberation there. (*A*) Experiment 1. (*B*) Experiment 2.

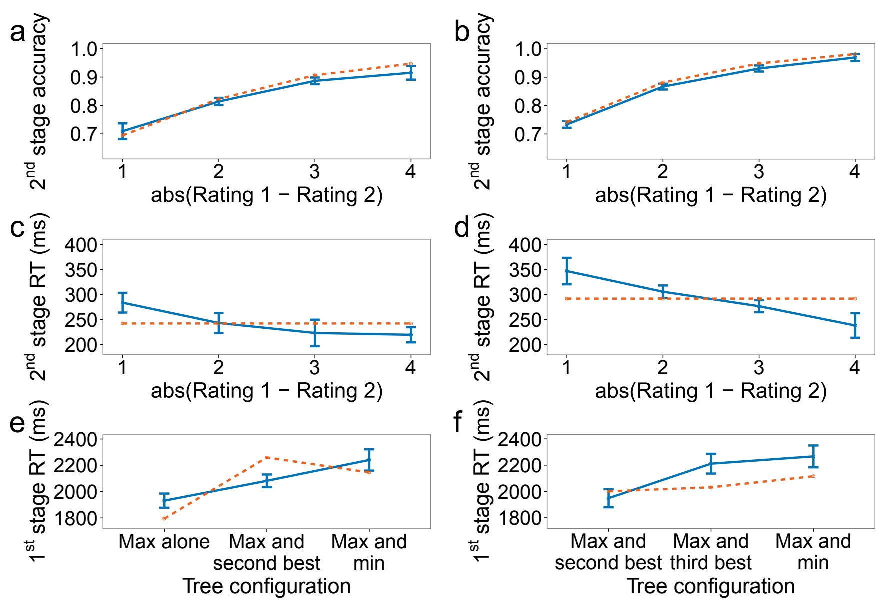

Fig. S3. Simulation of both experiments (*A*, *C*, and E, experiment 1; *B*, *D*, and *F*, experiment 2) using the two-stage model with correlated noise. Because noise is correlated at the top level, the model forces a resolution at the first stage between the bottom-level items on the winning side. However, the level of resolution required to match the first-stage accuracy and reaction time effects overestimates second-stage accuracy without additional deliberation. This in turn predicts a flat second-stage reaction time curve. The reduction in noise at the first stage differentially affects different tree configurations, also resulting in a poor fit to this aspect of the data.

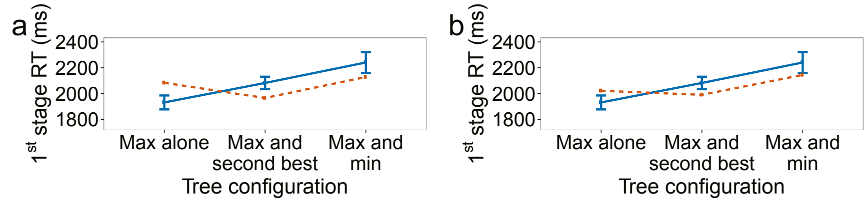

Fig. S4. Simulation of experiment 1 using the two-stage integration model with (*A*) pruning and (*B*) pruning and correlated noise. The pruning mechanism predicts that first-stage decisions should be fastest when the best and second-best paths appear on the same side of the decision tree, contrary to the data.

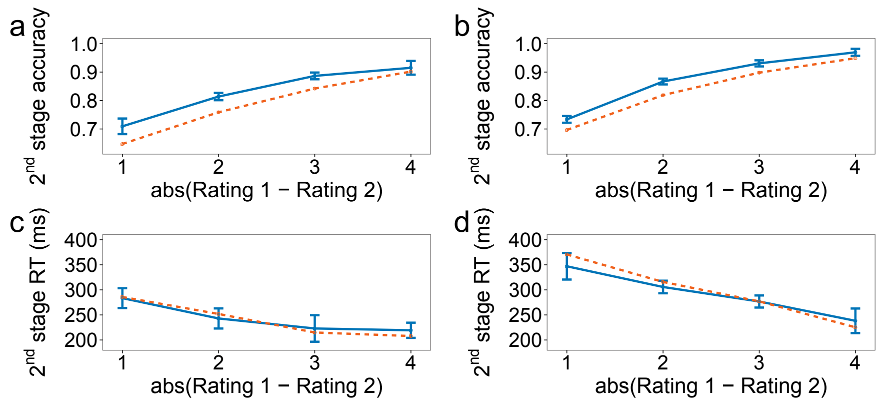

Fig. S5. Simulation of both experiments (*A* and *C*, experiment 1; *B* and *D*, experiment 2) using the primary model with a single drift rate for both stages. The model underestimates second-stage choice accuracy because the drift rate required to fit the first-stage data is much lower than that required to fit the second-stage data.

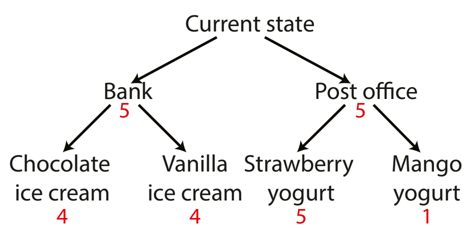

Fig. S6. Example of a max/mean conflict trial. Here the maximum overall path is on the right side (5 + 5 = 10); however, the average value of the two paths on the right [5 + (5 + 1)/2 = 8] is less than the average of the two paths on the left [5 + (4 + 4)/2 = 9].

Table S1. Best-fitting parameters of the winning model

Parameter	Exp. 1	Exp. 2
d_1	0.000107600	0.000122600
θ_1	0.727058500	0.738474100
T_1	394.9164580	188.2272950
d_2	0.002418200	0.001776700
θ_2	0.746330100	0.778294900
T_2	205.2812689	236.5221880

Table S2. Bayesian information criterion values

Model	Exp. 1	Exp. 2	Parameters
Two-stage, independent paths (primary model)	−265	−431	6
Forward greedy	−204	−395	6
Backward search	−222	−399	8
Backward search, same first-stage parameters	−129	−207	6
Backward search with reset	−257	−375	8
Backward search with reset, same first-stage parameters	−247	−415	6
One-stage with vigor	−199	−330	6
One-stage with single vigor	−202	−331	5
One-stage with single vigor and rating noise	−236	−373	6
Two-stage, correlated paths	−226	−360	6
Two-stage, independent paths, with pruning	−262	−418	7
Two-stage, correlated paths, with pruning	−249	−408	7
Two-stage, independent paths, single drift	−207	−351	5