Supplemental Text 1

Simulation analysis: differences in inverse temperature cause different estimates of advice-related learning biases

Method

Data simulation procedure

We simulated choice data from 200 synthetic participants on a simplified version (one instead of three stimulus pairs) of the task used by Decker et al. (2015). The task consisted of 60 choices between stimulus A and stimulus B. Reward probabilities for stimulus A and B were, respectively, .70 and .30. Stimulus B was assumed to be (incorrectly) advised to have a high reward probability.

Choices were simulated using a reinforcement-learning model extended with an ‘advice-bias parameter’, and a softmax decision function (as in Decker et al., 2015; Doll et al, 2009, 2011). Specifically, each time the non-advised stimulus (stimulus A) is chosen, the model updates the expected reward probability of that stimulus, $Q_A$, according to a standard delta-rule learning algorithm:

$$Q_{A,t+1} = Q_{A,t} + \alpha \delta_t$$

where $\delta_t = O_t - Q_{A,t}$ is the prediction error—the difference between the outcome (reward = 1, no reward = 0) and the expected reward probability at trial $t$—and $\alpha$ is the learning rate.

When the advised stimulus (stimulus B) is chosen, the learning algorithm is extended with bias parameter $\varphi$ that allows the impact of advice-consistent (reward) outcomes to be amplified and the impact of advice-inconsistent (no reward) outcomes to be diminished:

$$Q_{B,t+1} = Q_{B,t} + \varphi \alpha \delta_t$$

when $\delta_t$ is positive (i.e., when a reward is received)

$$Q_{B,t+1} = Q_{B,t} + \alpha \delta_t / \varphi$$

when $\delta_t$ is negative (i.e., when no reward is received).

The Q value of the unchosen stimulus is not updated.
The probability of choosing stimulus A on trial $t$ ($P_{A,t}$) was computed using a softmax function:

$$P_{A,t} = \frac{e^{\beta Q_{A,t}}}{e^{\beta Q_{A,t}} + e^{\beta Q_{B,t}}} = \frac{1}{1 + e^{-\beta(Q_{A,t} - Q_{B,t})}}$$

with inverse-temperature parameter $\beta$. If $\beta$ is 0 both options are equally likely to be chosen, irrespective of their expected reward probabilities (high degree of choice randomness or exploration). As the value of $\beta$ increases, the probability that the stimulus with the highest expected reward probability will be chosen increases (less choice randomness, higher degree of exploitation). For each synthetic participant, we simulated trial-specific choices using a Bernoulli process: on each trial $t$, stimulus A and B were chosen with, respectively, probability $P_{A,t}$ and probability $(1-P_{A,t})$.

We set $\alpha$ to 0.10 and $\phi$ to 3 in all simulations, resembling the estimated values from Decker et al. (2015). Because we assumed that stimulus B was incorrectly advised to have a high reward probability, we set the initial Q value of stimulus B ($Q_{B,1}$) to .80. The initial Q value of stimulus A ($Q_{A,1}$), for which no advice was received, was set to .50. Importantly, we set $\beta$ to either 1.25 (low inverse temperature) or 4.55 (high inverse temperature), in half of the simulations each. These values correspond to the estimated $\beta$ values for the adolescents and adults in Decker et al.’s (2015) study. Thus, we simulated two sets of data with different $\beta$, but identical values for the other parameters.

**Parameter recovery procedure**

We then fitted the same model to each synthetic participant’s simulated choice data, using a Bayesian parameter estimation approach, and estimated posterior distributions for $\alpha$, $\phi$, $Q_{A,1}$, $Q_{B,1}$ and $\beta$. We used a Beta(1,1) prior distribution for $\alpha$, $Q_{A,1}$ and $Q_{B,1}$, a uniform distribution ranging from 0.5 to 5 as prior for $\phi$, and a uniform distribution ranging from 0.1 to 10 as prior for $\beta$. We used the nonparametric Mann–Whitney test to compare the recovered parameter estimates (posterior medians) between the datasets that were simulated with low and high inverse temperature.
Parameter recovery results

Supplemental Figure 1 shows the average recovered parameters, as a function of the simulated (true) value of inverse-temperature parameter $\beta$. As expected the recovered value of $\beta$ was higher for the datasets simulated with a high than a low $\beta$ ($W = 9540, p < .001$; Supplemental Figure 1A). Recovered learning rate, $\alpha$, was overestimated for both sets of data, but higher for the datasets simulated with a low than a high $\beta$ ($W = 2684, p < .001$; Supplemental Figure 1B). Importantly, the recovered advice-bias parameter $\phi$ was significantly higher for the datasets simulated with a high $\beta$ ($W = 7887, p < .001$; Supplemental Figure 1C). In addition, the initial Q value for the advised option was higher, whereas the initial Q value for the nonadvised option was lower, for the datasets simulated with a high $\beta$ ($W = 8500, p < .001$ and $W = 2192, p < .001$. respectively; Supplemental Figure 1D). Thus, the effects of advice on both learning rate and initial estimates were overestimated when $\beta$ was high, as compared to when $\beta$ was low. This suggests that if groups of participants differ in choice randomness, or exploration, this can result in different estimates of advice-related biases, even when these biases do not actually differ between the groups.
Supplemental Text 2: Additional analyses on gambler’s fallacy-like strategies

A gambler’s fallacy-like strategy is more prevalent in adolescents than adults

41% of the adolescents and 9% of the adults used negative learning rates on more than 30% of the trials, suggestive of a gambler’s fallacy-like strategy. A non-preregistered and thus exploratory, analysis, showed that adolescents were more likely to show this behavior than adults ($\chi^2(1, n = 208) = 24.1, p < .001$). Within each age group, the prevalence of a gambler’s fallacy-like strategy was independent of the set-size version ($\chi^2(1, n = 120) = 1.8, p = .19$ and $\chi^2(1, n = 89) = .14, p = .71$, for the adolescents and adults, respectively).

Gambler’s fallacy-like behavior did not change over the course of the task

Our exclusion criterion was the use of negative learning rates on more than 30% of the trials, regardless of when these trials occurred. However, it is possible that participants switched from a gambler’s fallacy to a more optimal (learning) strategy over the course of the task, or vice versa. To examine this idea, we divided the task into five successive bins (each bin included two 16-trial blocks in the set size 1 condition, and one 32-trial block in the set-size 2 condition), and tested for a linear effect of task bin on the proportion of trials with negative learning rates. We also modeled effects of age group and set size, and all bin x age group x set size interactions. We conducted this analysis twice: once for the participants who were excluded according to our preregistered exclusion criterion (negative learning rates on more than 30% of the trials), and once for the included participants (negative learning rates on less than 30% of the trials).

For the included participants, there was no main effect of task bin ($p = .17$), and no interactions of bin with age group or set size either (all $p$’s > .52). Thus, gambler’s fallacy like behavior in the included participants did not change over the course of the task (and was low overall; Supplemental Figure 2, blue lines). For the excluded participants, there was no
main effect of task bin either \( p = .90 \); Supplemental Figure 2, pink lines). However, the proportion of negative learning rates showed a relatively stronger increase over task bins in the adults than in the adolescents (bin x age group interaction, \( t(53) = 3.5, p < .001 \), in particular when working-memory load was low (bin x age group x set size interaction, \( t(53) = 3.9, p < .001 \)). Note that, given the small number of adults in this analysis (5 and 3 adults in the set size 1 and 2 versions, respectively), these results should be interpreted with caution.

Together, these analyses suggest that the use of negative learning rates did not systematically increase or decrease over the course of the task, and was particularly stable in the participants who were included in the analyses.

**A gambler’s fallacy-like strategy is associated with reduced estimation accuracy**

Absolute estimation error (averaged across all trials and blocks) was larger for the participants who showed gambler’s fallacy-like behavior than for the participants who did not, and this was the case for the adults (21.7 vs. 12.0, \( t(7.6) = 3.4, p = .01 \)) as well as the adolescents (22.9 vs. 14.3, \( t(9.5) = 4.2, p = .002 \)), confirming that the use of a gambler’s fallacy strategy was maladaptive in our task.

**Effects of including participants with a gambler’s fallacy-like strategy on age-related differences in estimation accuracy**

We repeated the two analyses on estimation error without excluding any participants who showed gambler’s fallacy-like behavior. For the analysis on the correct- vs. no-advice conditions, this resulted in a significant main effect of age group (\( t(204) = 2.5, p = .01 \)—reflecting that estimation error was overall higher in the adolescents than adults—which was absent when participants with gambler’s fallacy-like behavior were excluded (\( p = .17 \)). For the analysis on the incorrect- vs. no-advice conditions, including participants with gambler’s fallacy-like behavior revealed age group x trial-linear (\( p = .048 \) and group x trial-linear x set
size ($p = .04$) interactions, reflecting that the adults showed a faster decrease in estimation error over trials, especially when working-memory load was low. These effects were absent when participants with gambler’s fallacy-like behavior were excluded. Together, these results suggest that including participants with gambler’s fallacy-like behavior resulted in stronger age effects on overall task performance. This can be explained by the fact that many more adolescents than adults used a maladaptive gambler’s fallacy-like strategy; hence including these participants caused a stronger decrease in average estimation accuracy in the adolescent group. However, the inclusion of participants with gambler’s fallacy-like behavior did not result in additional age group x advice, or age group x advice x set size, interactions in either analysis. This suggests that our exclusion criterion did not obscure age-related differences in advice effects.

**Supplemental Text 3**

A comparison of estimation accuracy in adolescents, 18-21-year-olds and 22-31-year-olds

To test whether the younger and older adults within our adult group were affected differently by the advice and set-size manipulations, we conducted additional non-preregistered hence exploratory analyses on estimation error comparing (i) the younger (18-21 years old, N=43) and older (22-31 years old, N=37) adults to each other, and (ii) each of these two adult groups to the adolescent group (Supplemental Figure 3). Mirroring the main analyses on estimation error, we separately analyzed the correct-advice vs. no-advice conditions (if win probability was .25 or .75) and the incorrect-advice vs. no-advice conditions (if win probability was .50). Besides age group, we also modeled effects of trial, advice, working-memory load, and all interactions, as in the main analyses.
The age-related effects found in the two main analyses on estimation error were not significant when comparing the younger and older adults (all $p$’s $>.10$). In the analysis on the correct- vs. no-advice conditions, the age-related effects found in the main analysis (age group x trial-linear and age group x set size x advice interactions) were present when comparing the adolescents to the older adults ($p = .02$ and $.003$, respectively), but were absent when comparing the adolescents to the younger adults ($p = .17$ and $.08$, respectively). This suggests that estimation performance continues to improve to some extent during early adulthood, in particular when correct advice is not available and working-memory load is high.

Finally, the age-related effects found in the main analysis for the incorrect- vs. no-advice conditions (main effect of age group and the age group x advice x trial-linear interaction) were present when comparing the adolescents to either the younger or older adults ($p$’s $>.02$), with one exception: the main effect of age group was not significant when comparing the adolescents to the older adults ($p = .29$). These results suggest that the susceptibility to incorrect advice decreases between adolescence and early adulthood, but does not decrease further after the age of 18.
Supplemental Text 4

Advice suppresses learning rate during initial trials

In this exploratory analysis, we tested for effects of advice (correct and incorrect advice combined) vs. no-advice on learning rate. We also modeled the effects of trial (linear and quadratic), age group and set size, and all interactions.

This analysis indicated that learning rate was overall higher in the adults than the adolescents ($t(147) = 2.0, p = .047$), and higher in the set-size 1 than set-size 2 version ($t(147) = 2.7, p = .007$; Supplemental Figure 5). In addition, learning rate decreased over trials, in a nonlinear way (linear trial effect, $t(147) = 4.2, p < .001$, quadratic trial effect, $t(147) = 2.0, p = .048$). Overall learning rate did not differ between the advice and no-advice conditions (main effect advice, $t(147) = 1.5, p = .14$). Importantly, however, there was an advice x trial-linear interaction ($t(147) = 4.1, p < .001$) reflecting a stronger reduction in learning rate over time in the absence of advice. Supplemental Figure 5 shows that this was due to higher learning rates in the absence of advice during the first (approximately five) trials. This effect did not differ between the adolescents and adults (advice x trial-linear x age group interaction, $t(147) = 1.4, p = .15$). However, this effect was stronger in the set size 1 than set size 2 version (advice x trial-linear x set size interaction, $t(147) = 2.3, p = .02$). Thus, advice suppressed learning rates, especially during early trials, but this effect did not differ between the two age groups.

We also examined whether the suppressive effect of advice on learning rate may have been driven by the correct-advice condition, in which initial estimates were more accurate and hence there was ‘less to learn’. To this end, we repeated the above analysis, but excluded the correct-advice conditions (i.e., we compared the no-advice vs. the incorrect advice conditions). All effects reported above were significant in this analysis as well. Moreover,
overall learning rate was higher in the no-advice than the incorrect-advice conditions (main effect of advice, $t(147) = 2.4, p = .02$), especially in the set size 1 version (set size x advice interaction, $t(147) = 2.4, p = .02$). These results were corroborated by our computational-modeling analyses (Supplementary Text 6, Supplementary Figure 7B). This indicates that the suppressive effect of advice on learning rate was not a byproduct of the more accurate initial estimates in the correct-advice condition.

Bayesian learning models assume that an increase in estimation certainty (i.e., a more precise prior) results in a decreased learning rate. Thus, in a Bayesian framework, the advice-related suppression of learning rate during early trials may be explained by an increase in estimation certainty. Analysis of participants’ certainty ratings, which will be reported in Supplemental Text 5, can be used to test this idea.
Supplemental Text 5

**Self-reported certainty increases over time and is higher in the presence of advice, in adolescents as well as adults**

Below we report our two preregistered analyses on self-reported estimation certainty. The first analysis contrasts the correct-advice vs. no advice conditions (with identical win probabilities), and the second analysis the incorrect-advice vs. no-advice conditions (with identical win probabilities), mirroring the analyses on estimation error reported in the main text. Supplemental Figure 6 shows how estimation certainty developed over trials in each age group and set-size version, for each combination of actual and advised win probability.

**Certainty in the correct-advice vs. no-advice conditions (win probability is .25 or .75)**

Participants’ self-reported certainty about their estimates for stimuli with win probability .25 and .75 increased over trials in a sublinear way (linear and quadratic effect of trial, $t(147) = 23.5, p < .001$ and $t(147) = 4.3, p < .001$, respectively). Adults showed a stronger linear increase in certainty over trials than adolescents, in particular in the set size 2 version (age group x trial-linear x set size interaction, $t(147) = 4.6, p < .001$). There were age group x trial-linear and age group x trial-quadratic interactions ($t(147) = 2.2, p = .03$ and $t(147) = 2.2, p = .03$, respectively), as well, reflecting that adolescents reported a stronger sublinear increase in certainty over trials. Finally, the age group x trial-quadratic interaction was stronger in the absence of advice (age group x trial-quadratic x advice interaction, $t(147) = 2.5, p = .01$), and especially so in the set-size 1 condition (age group x set size x advice x trial-quadratic interaction, $t(147) = 2.0, p = .04$).

In addition, certainty was overall higher in the correct-advice than the no-advice condition (main effect of advice, $t(147) = 3.0, p = .003$). The advice effect interacted with the quadratic effect of trial ($t(147) = 2.9, p = .004$), reflecting that the advice effect on certainty
was stronger during early and late than during intermediate trials (there was a nonsignificant trend for an advice x trial-linear interaction as well, suggesting that the advice effect decreased over trials, \(t(147) = 1.8, p = .07\)).

**Certainty in the incorrect-advice vs. no-advice conditions (win probability is .50)**

Participants’ self-reported certainty about their estimates for stimuli with win probability .50 increased linearly over trials (\(t(147) = 9.0, p < .001\)), in both age groups. In addition, certainty was overall higher in the incorrect-advice than no-advice condition (\(t(147) = 2.5, p = .01\)), and this advice effect decreased over trials in a non-linear way (advice x trial-linear interaction, \(t(147) = 3.1, p = .002\), and advice x trial-quadratic interaction, \(t(147) = 3.3, p < .001\)). There were no effects of age group or set size (all \(p\’s > .10\), except for an age group x advice x trial-quadratic interaction (\(t(147) = 2.0, p = .047\)). Thus, incorrect advice increased certainty especially during early trials, in both age groups.

Taken together, the analyses on certainty imply that both correct and incorrect advice increased self-reported certainty (relative to no advice), especially in the beginning of learning. For the stimuli with win probability .25 and .75, the adolescents showed a smaller increase in certainty over time than the adults, especially when working memory load was high. There were no other prominent age effects on certainty ratings.
**Supplemental Text 6A**

**No relationship between self-reported certainty and learning rate**

In this analysis, we explored whether participants used their current (un)certainty to regulate their learning rate, as would be predicted by Bayesian models of learning. We tested this using multilevel mediation (Supplemental Figure 7). Because advice suppressed learning rate and increased certainty specifically during early trials, we only included the first five trials in this analysis. We included age group and set size, and their interaction, as second-level moderators in the mediation model. In addition, we controlled for trial by including its linear and quadratic effects as covariates.

This analysis confirmed that advice, as compared to no advice, suppressed learning rate during the first five trials (path $c$, $p < .001$). In addition, advice increased certainty during the first five trials (path $a$, $p < .001$). However, certainty was not predictive of learning rate, when controlled for advice (path $b$, $p = .38$), and certainty did not mediate the effect of advice on learning rate (path $a*b$, $p = .72$). Thus, controlling for certainty did not change the effect of advice on learning rate (path $c'$, $p < .001$). Age group did not moderate any of the paths in the mediation model ($p$’s > .26), suggesting that the relationships between advice, certainty and learning rate did not differ between the adolescents and adults. Set size moderated the total effect of advice on learning rate (path $c$, $p = .048$), consistent with the analysis reported in Supplemental Text 4. The other paths were not moderated by set size.

We also repeated this analysis using all trials (instead of only the first five). This did not change the significance of the effects reported above, except that the advice effect on learning rate (path $c$) was no longer moderated by set size ($p = .07$).

Taken together, these results suggest that although advice affected both certainty and learning rate, these two effects were unrelated to each other. Thus, neither adults nor...
adolescents used their current (un)certainty to regulate their learning rate. These results are discussed in Supplemental Text 6B.

**Supplemental Text 6B**

**Discussion of the absence of a relationship between certainty and learning rate**

In a Bayesian framework, the current level of (un)certainty determines the learning rate on each trial, such that more certain expectations are updated less in response to new evidence. Inconsistent with this idea, our analysis reported in Supplemental Text 6A suggests that the participants in our estimation task did not use their current (un)certainty to regulate their learning rate. This seems at odds with our own recent finding that, in another estimation task, self-reported certainty *did* predict learning rate in both adults and adolescents (Jepma et al., 2020). What might explain this discrepancy? In our previous study, participants observed numeric outcomes sampled from a normal distribution, and estimated the mean of this distribution. This is an important difference with the current study, in which binary outcomes were sampled from a Bernoulli distribution, and participants estimated outcome probabilities. Participants in our previous study used much higher initial learning rates (close to 1) and showed a steeper decrease in learning rate over trials than participants in the current study, suggesting that people use different learning strategies when estimating numerical magnitudes and outcome probabilities. When estimating outcome probabilities, high initial learning rates—which would be consistent with normative, Bayesian learning—result in estimates close to either 0 or 1, and it is possible that people tend to avoid such extreme estimates. One way to avoid extreme probability estimates is to use a more gradual (non-Bayesian) updating strategy, e.g., adjusting one’s estimate only slightly following each new outcome, regardless of the number of previous outcomes. The use of this type of strategy
would diminish the relationship between learning rate and certainty, and may explain the lower and more stable learning rates in the current study.
Supplemental Text 7A

Computational modeling: Methods

We modeled the learning process using beta-binomial models. These models represent the win probability for a given stimulus as a beta distribution that is updated following each new outcome, naturally tracking both the estimated win probability as well as the uncertainty around this estimate. The family of sequentially updated beta-binomial models has been successfully used in recent probabilistic-learning studies, and has been shown to explain various types of learning data better than classical reinforcement-learning models that represent estimates as point values (de Boer et al., 2017; Lamba et al., 2020; Stankevicius et al., 2014; Tzovara et al., 2018; Wise et al., 2019).

According to the beta-binomial model, participants assume that the sequence of win and no-win outcomes sampled from stimulus s is the result of a Bernoulli process with a constant win probability. The win probability is represented as a beta distribution which is updated following each new observed outcome. Parameters $\alpha_s$ and $\beta_s$ of the beta distribution represent evidence for, respectively, win outcomes (coded as 1) and no-win outcomes (coded as 0). Thus, the model assumes that, on each trial $t$, participants update (increase) $\alpha_s$ after a win outcome, and $\beta_s$ after a no-win outcome, with the respective update rates determined by parameters $\tau^+$ and $\tau^-$:

\[
\alpha_{s,t+1} = \alpha_{s,t} + \tau^+ outcome_{s,t} \tag{1}
\]

\[
\beta_{s,t+1} = \beta_{s,t} + \tau^- (1 - outcome_{s,t}) \tag{2}
\]

This results in a beta distribution that reflects the expected win probability; it is biased towards 1 and 0 if, respectively, win and no-win outcomes occur most frequently, and becomes more precise as more outcomes are observed (i.e., as $\alpha_s + \beta_s$ increases).
We assume that participants’ estimated win probability reflects the mean of the beta distribution on the current trial, $\mu_{s,t}$:

$$\mu_{s,t} = \frac{\alpha_{s,t}}{\alpha_{s,t} + \beta_{s,t}}$$

(3)

To examine the effects of advice on initial expectations and expectation updating, we compared five versions of this model, which are described below. We applied all models to participants’ estimation data for the stimuli with win probability .50. Participants learned about these stimuli in the absence of advice and with false advice in two opposite directions (advised win probabilities were too low and too high), making these stimuli well-suited to examine advice effects on participants’ initial estimates and update rates.

**Model 1: Normative model**

In our simplest model, update rates $\tau^+$ and $\tau^-$ as well as the initial values of the beta distribution ($\alpha_{s,1}$ and $\beta_{s,1}$) are all fixed to 1. This model thus assumes an uninformative (uniform) prior on win probability. It reflects an ideal Bayesian learner that is not affected by advice. It has no free parameters.

In the next four models, we considered the possibility that participants’ initial win-probability estimates (priors) were biased towards a specific value, and that participants used different update rates for win- and no-win outcomes. To this end, $\alpha_{s,1}$ and $\beta_{s,1}$ as well as $\tau^+$ and $\tau^-$ were estimated as free parameters. To prevent a tradeoff between these parameters we constrained $\alpha_1$ and $\beta_1$ to sum to 2, effectively estimating the ratio rather than the absolute values of $\alpha_1$ and $\beta_1$. Thus, values of $\alpha_1$ below and above 1 reflect initial win-probability estimates below and above 0.5, respectively. The four models differ in the assumed effects of advice on the initial estimates and update rates, as described below.

**Model 2: No advice effects**
This model allows for biased initial estimates and asymmetric updating, but assumes that these effects do not vary between advice conditions. It has three free parameters: $\alpha_{s,1}$, $\tau^+$ and $\tau^-$ (note that $\beta_{s,1}$ is fixed to $2 - \alpha_{s,1}$).

**Model 3: Advice affects priors**

This model assumes that advice affects participants’ initial estimates (i.e., prior beliefs), but not their update rates. Therefore, this model separately estimates $\alpha_{s,1}$ for the no-advice, too-low advice, and too-high advice conditions, resulting in three initial-value parameters ($\alpha_{s,1}^{no}$, $\alpha_{s,1}^{low}$, $\alpha_{s,1}^{high}$). As in the previous model, $\tau^+$ and $\tau^-$ were estimated as well, but were equivalent for all advice conditions. Thus, this model has five free parameters.

**Model 4: Advice affects updating**

This model assumes that advice affects participants’ update rates, but not their initial estimates. Therefore, this model separately estimates $\tau^+$ and $\tau^-$ for the no-advice, too-low advice, and too-high advice conditions, resulting in six update-rate parameters ($\tau_{s,1}^{+no}$, $\tau_{s,1}^{-no}$, $\tau_{s,1}^{+low}$, $\tau_{s,1}^{-low}$, $\tau_{s,1}^{+high}$ and $\tau_{s,1}^{-high}$). As in Model 2, $\alpha_{s,1}$ was estimated as well, but was equivalent for all advice conditions. Thus, this model has seven free parameters.

**Model 5: Advice affects priors and updating**

This model assumes that advice affects participants’ initial estimates as well as their update rates. Therefore, $\alpha_{s,1}$, $\tau^+$ and $\tau^-$ are all estimated separately for the no-advice, too-low advice, and too-high advice conditions. This results in three initial-value parameters and six update-rate parameters; hence nine free parameters in total.

**Parameter estimation**

We applied all models to participants’ trial-to-trial sequences of estimated win probabilities for the stimuli with win probability .50, using a hierarchical Bayesian approach.
This approach assumes every participant has a different set of model parameters, which are drawn from group-level prior distributions (Gelman, 2014). The parameters governing the group-level prior distributions (hyperparameters) are also assigned prior distributions (hyperpriors). We estimated separate group-level distributions for the adults and adolescents, and for the set size 1 and set size 2 task versions. As we are primarily interested in average-participant behaviour within each age group and task version, and in potential group differences, our primary variables of interest are the hyperparameters governing the means of the group-level distributions.

**Prior distributions**

Group-level distributions for $\alpha_{s,1}$, $\tau^+$ and $\tau^-$ in Model 2-5 were assumed to be beta distributions. Beta distributions are typically defined by two shape parameters, but we reparameterized these in terms of a group-level mean and group-level precision (Smithson & Verkuilen, 2006). The group-level means were assigned a uniform hyperprior on the interval [0,1] and the logarithms of the group-level precisions were assigned a uniform hyperprior on the interval $[\log(2), \log(600)]$ (Steingroever & Wagenmakers, 2014). Parameters for individual participants were drawn from the resulting group-level beta distributions. At the individual level, we transformed the range of the $\alpha_{s,1}$ parameter from [0, 1], the range of the beta distribution, to [0, 2]. Individual-level $\beta_{s,1}$ parameter was then fixed to $2 - \alpha_{s,1}$. We transformed the range of the individual-level $\tau^+$ and $\tau^-$ parameters from [0, 1] to [0, 10].

For model fitting, we assumed that participants’ win-probability estimates were normally distributed around the models’ point predictions ($\mu_{s,t}$, see equation 3). The variance of this normal distribution (error variance) was assumed to have a half-Cauchy group-level distribution. The scale parameter governing the half-Cauchy distribution was assigned a uniform hyperprior on the interval [0,1000]. All priors and hyperpriors were chosen for their uninformative nature.
**MCMC sampling**

We inferred posterior distributions for all model parameters using Markov chain Monte Carlo (MCMC) sampling, as implemented in JAGS (Plummer, 2003) via the R2jags package (Su & Yajima, 2015). We ran 3 independent MCMC chains with different starting values per model parameter, and collected 40,000 posterior samples per chain. We discarded the first 20,000 iterations of each chain as burn-in. In addition, we only used every 5th iteration to remove autocorrelation. Consequently, we obtained 12,000 representative samples per parameter per model. All chains showed convergence (R-hat < 1.1) (Gelman & Rubin, 1992).

**Model comparison**

We compared the performance of the different models using the deviance information criterion (DIC) (Spiegelhalter et al., 2002), separately for the two age groups and set size versions. The DIC is a hierarchical modeling generalization of the AIC which is easily computed in hierarchical Bayesian model-selection problems using Markov chain Monte Carlo (MCMC) sampling. It provides an index of the goodness of fit of a model, penalized by its effective number of parameters. Models with smaller DIC are better supported by the data.

We computed ΔDIC values by subtracting each model’s DIC from the DIC of the worst-fitting model (Model 1, the normative model); hence higher ΔDIC values indicate a better fit. To assess whether the ΔDIC values of the different models were significantly different, we computed their 95% confidence intervals using a bootstrap procedure. Specifically, we randomly sampled—*with replacement*—as many participants from the data as there were in the actual dataset, separately for each age group and set size version. We repeated this process 30 times, resulting in 30 bootstrap samples per age group and set size version. We then applied each model to each bootstrap sample (all models were thus applied...
to the same 30 bootstrap samples), and computed the ΔDIC values for each bootstrap sample. We then computed the 95% confidence intervals of the ΔDIC values. We assume a significant difference between two models if the 95% confidence intervals of their ΔDIC values do not overlap.

Supplemental Text 7B

Computational modeling: Results

Model comparison

Model 5 (advice affects priors and updating) had the best fit (i.e., the highest ΔDIC), followed by, respectively, Model 3 (advice affects priors only), Model 4 (advice affects updating only), Model 2 (no advice effects), and Model 1 (normative model). This was the case for both age groups and set-size versions. However, for the set size 1 versions of both age groups, the difference between Model 5 and Model 3 was not significant (their bootstrap 95% confidence intervals partially overlapped; Supplemental Figure 8). This suggests that when working-memory load was low, the advice definitely affected participants’ priors (initial estimates) but not necessarily their update rates. When working-memory load was high (set size 2), the difference between Model 5 and Model 3 was significant (their bootstrap 95% confidence intervals did not overlap), suggesting that advice affected both priors and update rates. Supplemental Figure 9 illustrates the fit of the Model 5 per age group and set size version.

Parameter estimates

We next examined the parameter estimates of the best fitting model (Model 5). We estimated the parameters’ group-level distributions separately for each age group and set-size version, and denote the hyperparameters governing the means of these group-level
distributions with overbars. In addition, we indicate the advice condition in a superscript (no, low, and high refer to no advice, too-high (.61-70) advice, and too-low (.31-.40) advice, respectively; e.g., $\alpha_{1}^{\text{high}}$ refers to the group-level mean of $\alpha_{1}$ for the stimuli paired with too-high advice).

Supplemental Figure 10A shows the initial win-probability estimates computed from the posterior medians of $\alpha_{1}^{\text{no}}$, $\alpha_{1}^{\text{high}}$, and $\alpha_{1}^{\text{low}}$, and Supplemental Figure 11 shows the full posterior distributions. Consistent with the behavioral estimation data (Figure 3, main text), these initial estimates lay around .50 in the absence of advice, and were biased upwards and downwards when, respectively, higher and lower advice was received. To examine the significance of these advice effects, we computed the percentage of MCMC samples for which $\alpha_{1}^{\text{high}}$ was higher than $\alpha_{1}^{\text{no}}$, and for which $\alpha_{1}^{\text{low}}$ was lower than $\alpha_{1}^{\text{no}}$, in each age group and set-size version. We found that $\alpha_{1}^{\text{high}}$ was higher than $\alpha_{1}^{\text{no}}$ for 92% of the MCMC samples in the adolescents who completed the set-size 1 version, and for 100% of the MCMC samples in the other groups. $\alpha_{1}^{\text{low}}$ was lower than $\alpha_{1}^{\text{no}}$ for more than 99% of the MCMC samples in all groups expect for the adolescents who completed the set-size 2 task version (in this group, $\alpha_{1}^{\text{low}} < \alpha_{1}^{\text{no}}$ for 69% of MCMC samples, but $\alpha_{1}^{\text{low}} < \alpha_{1}^{\text{high}}$ for 100% of the MCMC samples). Thus, participants’ initial win-probability estimates were adjusted in the direction of the advice, with the exception that the adolescents who completed the set size 2 task version did not differentiate between the too-low advice and no-advice conditions.

Supplemental Figure 10B shows the posterior medians of update-rate parameters $\tau^{+}$ and $\tau^{-}$ for each advice condition, and Supplemental Figure 12 shows the full posterior distributions. For the adults who completed the set-size 1 version, both too-high and too-low advice resulted in reduced update rates compared to the no-advice condition ($\tau^{+,\text{high}} < \tau^{+,\text{no}}$, ...
\[
\tau^{+,low} < \tau^{+,no}, \tau^{-,high} < \tau^{-,no}, \tau^{-,low} < \tau^{-,no}, \text{ for } 97, 99, 98 \text{ and } 100\% \text{ of the MCMC samples, respectively.} \]

A comparison of the too-high and too-low advice conditions suggests that the upward updating of estimated win probability following win outcomes did not differ reliably between the two advice conditions (\(\tau^{+,high} > \tau^{+,low}\) for 80\% of MCMC samples), whereas downward updating following no-win outcomes was stronger in the too-high advice condition (\(\tau^{-,high} > \tau^{-,low}\) for 95.3\% of MCMC samples). The adolescents who completed the set-size 1 version used lower update rates in both advice conditions compared to the no-advice condition as well (\(\tau^{+,high} < \tau^{+,no}, \tau^{+,low} < \tau^{+,no}, \tau^{-,high} < \tau^{-,no} \) and \(\tau^{-,low} < \tau^{-,no}\) for 96, 92, 99 and 99\% of the MCMC samples, respectively). A comparison of the too-high and too-low advice conditions revealed no reliable differences between these conditions
\(\tau^{+,high} > \tau^{+,low}\) and \(\tau^{-,high} > \tau^{-,low}\) for 72 and 53\% of the MCMC samples, respectively).

For the adults who completed the set-size 2 version, we found no evidence that update rates differed between the advice and no-advice conditions (\(\tau^{+,high} < \tau^{+,no}, \tau^{+,low} < \tau^{+,no}, \tau^{-,high} < \tau^{-,no} \) and \(\tau^{-,low} < \tau^{-,no}\) for 60, 78, 69 and 56\% of the MCMC samples, respectively), nor between the too-high and too-low advice conditions (\(\tau^{+,high} > \tau^{+,low}\) and \(\tau^{-,high} > \tau^{-,low}\) for 79 and 33\% of the MCMC samples, respectively). The same was true for the adolescents who completed the set-size 2 version: update rates did not differ reliably between the advice and no-advice conditions (\(\tau^{+,high} < \tau^{+,no}, \tau^{+,low} < \tau^{+,no}, \tau^{-,high} < \tau^{-,no} \) and \(\tau^{-,low} < \tau^{-,no}\) for 41 61, 61 and 68\% of the MCMC samples, respectively), nor between the too-high and too-low advice conditions (\(\tau^{+,high} > \tau^{+,low}\) and \(\tau^{-,high} > \tau^{-,low}\) for 68 and 51\% of the MCMC samples, respectively).

Taken together, these results suggest that when working-memory load is low, both adults and adolescents updated their expected win probability in response to new outcomes more in the absence than in the presence of advice, regardless of whether the advice was too
low or too high. This corroborates the results reported in Supplemental Text 4. In addition, when working-memory load was low, adults showed stronger downward updating when they had received too-high, as compared to too-low, advice.

**Supplementary Text 7C: Model recovery analysis**

**Procedure**

To test how well our different models can be distinguished, we conducted a model-recovery analysis. To this end, we simulated 50 datasets with each of our five models. Each simulated dataset consisted of a group of 35 synthetic participants who performed three 16-trial blocks of our estimation task. The actual win probability was .50 for all blocks (outcome order was pseudo-randomized in the same way as in our real task), and the blocks were paired with too-low, too-high and no advice (one block each), mirroring the real task blocks used in our modeling analyses. For each simulation, the hyperparameters governing the model’s group-level distributions were sampled randomly from uniform distributions whose range was matched to the range of values obtained from our fits to the real data. Specifically, the minimum and maximum values of the uniform distributions were set to the minimum and maximum values of the hyperparameters’ posterior medians for the two age groups and set size versions. The individual-level parameters used to simulate the data were then sampled from the resulting group-level distributions.

To examine to what extent the data-generating models could be recovered, we applied all models to each simulated dataset and determined the best fitting model (i.e., the model with the lowest DIC) for each simulated dataset, using the same model-fitting procedure as used for the real data.

**Results**
We found that each (100%) of the 50 datasets simulated by a given model was best fit by that same model, as opposed to one of the other four models. Thus, our procedure could perfectly distinguish the different advice effects captured by our different models, suggesting that we can be confident about the model-comparison results.
Supplemental Figure 1. Mean recovered parameter values (posterior medians) for fits of the model to datasets simulated with low and high $\beta$. Horizontal lines indicate the simulated (true) parameter values. Error bars are standard errors.
Supplemental Figure 2. Mean proportion of trials with negative learning rates in five successive task bins, as a function of age group and set size, and separately for the participants who were included and excluded according to our gambler’s fallacy criterion (exclusion criterion was an average proportion of negative learning rates above .30, across all trials). Error bars indicate standard errors.
**Supplemental Figure 3.** Mean estimated win probability for the 22-31-year-old adults, 18-21-year-old adults, and adolescents, as a function of trial, actual and advised win probability, and set size for **A** the correct- and no-advice conditions with win probability .25 and .75, and **B** The incorrect- and no-advice conditions with win probability .50. Error bars indicate standard errors.
Supplemental Figure 4. Trial-to-trial learning rates per age group and set-size version.

Learning rates in each group are plotted separately for the no-advice (A), correct-advice (B) and incorrect-advice (C) stimuli.
**Supplemental Figure 5.** Learning rate as a function of advice (present vs. absent), trial, age group and set size.
**Supplemental Figure 6.** Mean certainty rating as a function of trial, separately for the two age groups and set-size versions, plotted separately for **A** the correct-advice, incorrect-advice and no-advice conditions, **B** the correct-advice and no-advice conditions with win probability .25 and .75, and **C** the too-high advice, too-low advice, and no-advice conditions with win probability .50. Error bars indicate standard errors.
Supplemental Figure 7. Mediation model and results.

Notes. + and – indicate positive and negative relationships, respectively; *** p < .001, n.s. not significant.
Supplemental Figure 8. Model comparison. We obtained ΔDIC values by subtracting each model’s DIC from the DIC of the worst-fitting model (Model 1). Higher ΔDIC values indicate a better fit. The winning model (Model 5, advice affects priors and updating) is indicated in gray. Error bars indicate 95% confidence intervals, computed using a bootstrap procedure (30 bootstrap samples).
**Supplemental Figure 9.** Model fit. The average estimates of the participants (data; straight lines) and the average predicted estimates of the winning model (Model 5; dotted lines) per trial, age group and condition. To obtain the model’s predictions, we simulated estimation data for each participant using the posterior medians of the individual-level parameters, and then plotted the average predictions. Note that we only applied models to the conditions with win probability .50.
**Supplemental Figure 10.** Initial estimates and update rates derived from Model 5. **A.** Initial win-probability estimates, computed from the posterior medians of $\alpha_1$, per advice condition, age group and set-size version. Adul = adults, Adol = adolescents. **B.** Posterior medians of $\tau^+$ and $\tau^-$ (update rates) per age group and set-size version.
Supplemental Figure 11. Posterior distributions of $\alpha$ per advice condition, age group and set size version. Because we constrained $\alpha$ and $\beta$ to sum to 2, the initial win-probability is $\alpha/2$ (see Equation 3). Thus, values of $\alpha$ above and below 1 correspond to initial win-probabilities above and below .5, respectively.

Supplemental Figure 12. Posterior distributions of update-rate parameters $\tau^+$ (thick lines) and $\tau^-$ (thin lines) per advice condition, age group and set size version.
**Supplemental Table 1.** Statistics for all fixed effects in the analysis on absolute estimation error for stimuli paired with no vs. correct advice (actual win probability is .25 and .75)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Coefficient</th>
<th>STD</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.66</td>
<td>1.19</td>
<td>13.17</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age group (adults &gt; adolescents)</td>
<td>-1.99</td>
<td>1.44</td>
<td>-1.38</td>
<td>0.17</td>
</tr>
<tr>
<td>Set size (2 &gt; 1)</td>
<td>4.05</td>
<td>1.70</td>
<td>2.39</td>
<td>0.02</td>
</tr>
<tr>
<td>Advice (correct &gt; no)</td>
<td>-5.18</td>
<td>0.82</td>
<td>-6.32</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Trial-L</td>
<td>-637.65</td>
<td>38.8</td>
<td>-16.43</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Trial-Q</td>
<td>272.61</td>
<td>47.6</td>
<td>5.72</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age group x set size</td>
<td>-1.10</td>
<td>2.11</td>
<td>-0.52</td>
<td>0.60</td>
</tr>
<tr>
<td>Age group x advice</td>
<td>-0.23</td>
<td>1.14</td>
<td>-0.21</td>
<td>0.84</td>
</tr>
<tr>
<td>Set size x advice</td>
<td>-5.20</td>
<td>1.14</td>
<td>-4.55</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age group x trial-L</td>
<td>-124.38</td>
<td>53.4</td>
<td>-2.33</td>
<td>0.02</td>
</tr>
<tr>
<td>Set size x trial-L</td>
<td>131.11</td>
<td>52.2</td>
<td>2.51</td>
<td>0.01</td>
</tr>
<tr>
<td>Advice x trial-L</td>
<td>355.49</td>
<td>83.2</td>
<td>4.27</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age group x trial-Q</td>
<td>58.17</td>
<td>66.0</td>
<td>0.88</td>
<td>0.38</td>
</tr>
<tr>
<td>Set size x trial-Q</td>
<td>-37.28</td>
<td>64.6</td>
<td>-0.58</td>
<td>0.56</td>
</tr>
<tr>
<td>Advice x trial-Q</td>
<td>-208.54</td>
<td>70.2</td>
<td>-2.97</td>
<td>0.003</td>
</tr>
<tr>
<td>Age group x set size x advice</td>
<td>4.53</td>
<td>1.57</td>
<td>2.88</td>
<td>0.004</td>
</tr>
<tr>
<td>Age group x set size x trial-L</td>
<td>28.36</td>
<td>71.5</td>
<td>0.40</td>
<td>0.69</td>
</tr>
<tr>
<td>Age group x advice x trial-L</td>
<td>52.76</td>
<td>112.8</td>
<td>0.47</td>
<td>0.64</td>
</tr>
<tr>
<td>Set size x advice x trial-L</td>
<td>4.33</td>
<td>118.5</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>Age group x set size x trial-Q</td>
<td>-141.27</td>
<td>88.8</td>
<td>-1.59</td>
<td>0.11</td>
</tr>
<tr>
<td>Age group x advice x trial-Q</td>
<td>9.58</td>
<td>97.6</td>
<td>0.10</td>
<td>0.92</td>
</tr>
</tbody>
</table>
Set size x advice x trial-Q  
Age group x set size x advice x trial-L  
Age group x set size x advice x trial-Q

Notes. STE = standard error, L = linear, Q = quadratic. We used R’s poly function to transform our raw ‘trial’ variable (which contained values from 1 to 16) into orthogonal linear and quadratic trial regressors. This transformation rescaled the trial values to very small numbers (between -.01 and .01), which explains the large coefficients for the trial regressors.
### Supplemental Table 2

Statistics for all fixed effects in the analysis on absolute estimation error for stimuli paired with no vs. incorrect advice (actual win probability is .50)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Coefficient</th>
<th>STE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>10.6</td>
<td>1.09</td>
<td>9.71</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age group (adults &gt; adolescents)</td>
<td>-2.99</td>
<td>1.38</td>
<td>-2.17</td>
<td>0.03</td>
</tr>
<tr>
<td>Set size (2 &gt; 1)</td>
<td>2.93</td>
<td>1.88</td>
<td>1.56</td>
<td>0.12</td>
</tr>
<tr>
<td>Advice (no &gt; incorrect)</td>
<td>-0.88</td>
<td>1.00</td>
<td>-0.87</td>
<td>0.38</td>
</tr>
<tr>
<td>Trial-L</td>
<td>-20.2</td>
<td>62.9</td>
<td>-0.32</td>
<td>0.75</td>
</tr>
<tr>
<td>Trial-Q</td>
<td>72.3</td>
<td>45.6</td>
<td>1.59</td>
<td>0.11</td>
</tr>
<tr>
<td>Age group x set size</td>
<td>-0.52</td>
<td>2.30</td>
<td>-0.22</td>
<td>0.82</td>
</tr>
<tr>
<td>Age group x advice</td>
<td>0.20</td>
<td>1.32</td>
<td>0.15</td>
<td>0.88</td>
</tr>
<tr>
<td>Set size x advice</td>
<td>-0.28</td>
<td>2.19</td>
<td>-0.13</td>
<td>0.90</td>
</tr>
<tr>
<td>Age group x trial-L</td>
<td>-122.1</td>
<td>81.5</td>
<td>-1.50</td>
<td>0.13</td>
</tr>
<tr>
<td>Set size x trial-L</td>
<td>150.4</td>
<td>84.5</td>
<td>1.78</td>
<td>0.08</td>
</tr>
<tr>
<td>Advice x trial-L</td>
<td>-79.8</td>
<td>70.3</td>
<td>-1.13</td>
<td>0.26</td>
</tr>
<tr>
<td>Age group x trial-Q</td>
<td>-13.4</td>
<td>70.2</td>
<td>-0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Set size x trial-Q</td>
<td>-103.8</td>
<td>65.6</td>
<td>-1.58</td>
<td>0.11</td>
</tr>
<tr>
<td>Advice x trial-Q</td>
<td>-74.7</td>
<td>57.20</td>
<td>-1.31</td>
<td>0.19</td>
</tr>
<tr>
<td>Age group x set size x advice</td>
<td>0.21</td>
<td>2.6</td>
<td>0.08</td>
<td>0.94</td>
</tr>
<tr>
<td>Age group x set size x trial-L</td>
<td>130.2</td>
<td>116.9</td>
<td>1.11</td>
<td>0.27</td>
</tr>
<tr>
<td>Age group x advice x trial-L</td>
<td>349.9</td>
<td>97.6</td>
<td>3.59</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Set size x advice x trial-L</td>
<td>37.2</td>
<td>116.0</td>
<td>0.32</td>
<td>0.75</td>
</tr>
<tr>
<td>Age group x set size x trial-Q</td>
<td>-65.8</td>
<td>96.4</td>
<td>-0.68</td>
<td>0.50</td>
</tr>
<tr>
<td>Age group x advice x trial-Q</td>
<td>-15.0</td>
<td>79.1</td>
<td>-0.19</td>
<td>0.85</td>
</tr>
<tr>
<td>Factor</td>
<td>Mean 1</td>
<td>Mean 2</td>
<td>SD</td>
<td>p</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>Set size x advice x trial-Q</td>
<td>12.8</td>
<td>91.1</td>
<td>0.14</td>
<td>0.89</td>
</tr>
<tr>
<td>Age group x set size x advice x trial-L</td>
<td>-139.6</td>
<td>157.8</td>
<td>-0.88</td>
<td>0.38</td>
</tr>
<tr>
<td>Age group x set size x advice x trial-Q</td>
<td>1.80</td>
<td>125.65</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Supplemental References

https://doi.org/10.7554/eLife.26424


https://doi.org/10.1177/0956797620910993


https://doi.org/10.1037/1082-989X.11.1.54

https://doi.org/10.1371/journal.pcbi.1003605

https://doi.org/10.1017/S0140525X13000861

https://doi.org/https://doi.org/http://cran.r-project.org/package=R2jags

https://doi.org/10.1371/journal.pcbi.1006243

https://doi.org/10.1371/journal.pcbi.1007341