PNAS

Supplementary Information for:

Aging is associated with maladaptive episodic memory-guided social decision-making

Karolina M. Lempert[1§], Michael S. Cohen[1], Kameron A. MacNear[2], Frances M. Reckers[1], Laura Zaneski[1], David A. Wolk[3], Joseph W. Kable[1*]

[1]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104
[2]Department of Psychology, University of Illinois – Urbana-Champaign, Champaign, IL 61820
[3]Hospital of the University of Pennsylvania, Philadelphia, PA 19104

[§]Current affiliation for Karolina Lempert is: Gordon F. Derner School of Psychology, Adelphi University, Garden City, NY 11530.

***Corresponding author:**
Joseph W. Kable
kable@psych.upenn.edu

This PDF file includes:

SI Materials and Methods
Supplementary text
Figure S1
Table S1

SI Materials and Methods

Stimuli. Prior to data collection for the main experiment, we invited participants (*n* = 106; mean age = 24.56; SD = 7.67; range: 18-58; gender: 61 F, 44 M, 1 not reported; race and ethnicity: 45 White non-Hispanic, 1 White Hispanic, 30 Asian non-Hispanic, 19 Black non-Hispanic, 1 Black Hispanic, 7 mixed race, 1 Other Race Hispanic, 2 not reported), to the lab to complete a survey of social decisions and to have their photograph taken. Participants had three photographs taken – one in which they made a happy face, one in which they made an angry face, and one in which they made a neutral face. Only the neutral photographs were used in the current study. Participants were paid $5 for this session, which took approximately 15 minutes. They were told that they were being invited to be part of a research database, and that their decisions and photographs may be used in future studies. Participants provided their contact information, since there was a possibility that their choices would be played out for real at a later date. One of the decisions participants were asked to make in the survey was a standard Dictator Game decision, in which they were asked to imagine that they were given $10 to split between themselves and an anonymous other person. They were then asked to indicate their split from a multiple-choice menu, in which the options ranged from keeping $0 (and sharing $10) to keeping $10 (and sharing $0) in one-dollar increments.

Once the database was collected, we selected 64 neutral photographs to use as stimuli in the main experiment. 32 of these were presented in the Reward, Decision, and Memory tasks (see details in *Methods*). 16 additional faces were presented as novel foils in the Decision task only, and another 16 were presented as novel foils in the

Memory Task only. In order to be selected as one of the 32 stimuli in the Reward task, the photographs had to be of people who selected either a “keep $10 / share $0” split or a “keep $5 / share $5” split in the Dictator Game, since half the stimuli in the Reward task were meant to be high-value and half were meant to be low-value. In addition, the stimuli were selected such that there was a balance with respect to race and gender between the high-value and low-value stimuli, since we did not want any effects of race and/or gender on memory or decision-making to be confounded with effects of stimulus value on memory or decision-making. Therefore, the final stimulus set contained four Asian female, four Asian male, four Black female, four Black male, eight White female, and eight White male faces, evenly split between the high-value (“keep $5 / share $5”) and low-value (“keep $10 / share $0”) categories. The individuals shown in the final stimulus set were all young adults (mean age = 25.38; SD = 5.08; range: 18-36). Since we were unable to include stimuli from all age groups (due to the sample in the database collection study being composed primarily of young adults), we decided to keep the stimulus set homogeneous with respect to age.

The 32 novel stimuli had a similar race, gender, and age distribution, but note that participants never learned about the decisions of the novel stimuli, so for that set, we did include photographs of people who proposed to share amounts other than $5 or $0. Photographs were cropped such that only faces were shown.

We brought in an independent sample of participants into the lab to view all of the potential stimuli and to make judgments about them (n = 20; mean age = 22.115; SD = 2.35; gender: 16 F, 4 M; race and ethnicity: 9 White non-Hispanic, 7 Asian non-Hispanic, 4 Black non-Hispanic). The participants made ratings (scale from 0-10) of the photograph's perceived attractiveness, trustworthiness, dominance, competence, and warmth. These attributes were chosen based on previous research showing that certain facial features are associated with these qualities, and that these qualities have been linked to social decisions. Participants also were asked to imagine what the person pictured would share in a Dictator Game with an anonymous stranger (from $0-$10; we termed this variable "perceived generosity"). For each stimulus, we took the average rating for each attribute (including perceived generosity) and ran a one-way ANOVA to test if any of these ratings could differentiate stimuli who shared $5 versus those who shared $0. We found no significant effects (attractiveness: $F_{(1, 27.7)} = 0.96$; $p = 0.336$; trustworthiness: $F_{(1, 28.4)} = 0.07$; $p = 0.787$; dominance: $F_{(1, 24.2)} = 1.23$; $p = 0.278$; competence: $F_{(1, 28.1)} = 0.18$; $p = 0.678$; warmth: $F_{(1, 25.5)} = 0.44$; $p = 0.514$; perceived generosity: $F_{(1, 29.1)} = 0.01$; $p = 0.928$). The average ratings for attractiveness, trustworthiness, dominance, competence, warmth, and perceived generosity from this independent set of subjects were used in other analyses, described below.

The houses were color houses selected from a stimulus set provided by the Epstein lab at the University of Pennsylvania. Houses were randomly paired with $5 and $0 values.

Remote data collection. Participants who completed the task remotely underwent the same procedure with a few small changes. The Reward, Distractor, Decision and Memory tasks were presented in Qualtrics instead of E-Prime. Participants initially met with the experimenter over the BlueJeans platform, and shared their screen after opening the Qualtrics link that contained the task. The experimenter could then guide the participant through the instructions just as they would in the lab. While the participant did the task, the experimenter turned their camera and sound off in order to minimize distractions. Although the timing of the task was the same in Qualtrics as it was in E-Prime, participants made their choices in Qualtrics by clicking on an option with a mouse rather than using a keyboard. Remote participants were paid the same fees and bonuses, but these were paid via debit cards mailed to each participant on the day that the study was completed, rather than paid in cash.

Exclusion criteria. Participants were excluded if, in the Decision phase of the experiment, they either chose the stimulus or chose the schematic image more than 95% of the time (*n* = 10 subjects). They were also excluded if, in the Memory phase, they reported that a stimulus that they had seen gave $5 or $0 more than 95% of the time (*n* = 3 additional subjects). For some analyses, not all participants had values for every category (e.g., some participants had no “guess” trials or no “confident” trials). Rather than exclude them from analyses altogether, they are left out of only those analyses for which they had missing values. The number of subjects in each analysis is specified in the text where relevant.

Post-hoc analyses

Effects of face stimulus age, race, and gender on choice. The thirty-two face stimuli learned about in the study varied with respect to age (although all were young: age range = 18-36), gender (50% female, 50% male), and race (50% White, 25% Asian, 25% Black). We did an exploratory analysis to examine if any of these demographic variables impacted choice, above and beyond the normative effects of value and memory. To examine the effects of stimulus age, we ran a mixed-effects logistic regression, with age as a continuous independent variable and choice (1 = chose stimulus; 0 = chose schematic) as the dependent variable. For the gender analysis, gender was coded as a two-level (Male/Female) factor variable with Male as the reference level. For the race analysis, race was coded as a three-level factor variable (White, Black, Asian), with white race as the reference level. In all three regressions, value (0 = $0 stimulus; 1 = $5 stimulus), associative memory (0 = incorrect associative memory; 1 = correct associative memory), and the value*associative memory interaction term were entered as covariates of no interest. We allowed intercepts and slopes (for the independent variable of interest) to vary by subject. All participants across all age groups were included in these three mixed-effects regressions. In addition to examining the fixed effects of these face attributes, we were also interested in how these effects might vary with participant age. Therefore, we extracted the random-effects coefficients on stimulus age, race, and gender for each subject and ran a Pearson correlation between these coefficients and participant age.

Contributors to social avoidance decisions. We found that there are several reasons why older adults may make maladaptive memory-based decisions in the social domain: item and associative memory deficits, a response bias in associative memory, a tendency to approach social partners regardless of retrieved value, and an over-reliance on perceptions of trustworthiness. We ran a post-hoc multiple regression analysis to investigate which of these factors was most predictive of failures to avoid unfair others.

We chose to focus on these low-value social trials, because, as described in the main text and illustrated in Fig. 4, older adults were actually unimpaired at choosing to approach others who were previously generous. Therefore, older adults' decision-making deficits in the social domain were specific to their inability to avoid unfair others. The dependent variable in our model was, for each subject, the proportion of all previously-seen low-value social stimuli that were successfully avoided. Our model contained the following predictors. First, there were two normative influences on choice: social item memory d' and social associative memory d'. Our third predictor was social associative memory response bias (higher values indicate a greater tendency to label previously-seen stimuli as being low-value). Our fourth predictor was the difference in average "perceived generosity" between the low-value stimuli that were chosen compared to those that were not chosen (higher values suggest that this face perception variable was being considered in the decision process; note that we used this measure instead of the random-effects coefficient described in the text, in order to avoid the circularity inherent in predicting choice using a coefficient derived from a model that was used to predicted the same choices). Finally, our fifth predictor was the proportion of times that a stimulus confidently remembered as being low-value was chosen. Note that this variable is not collinear with the other predictors, since it is conditionalized on participants believing that the stimulus was worth $0; only the relative number of such trials is affected by associative memory accuracy and response bias.

Supplementary Text

Recruitment and testing methods do not significantly impact performance. About half of the participants in each group performed a remote version of the task (*n* = 38 in young group; *n* = 35 in middle-aged group; *n* = 33 in older group). When examining general item and associative memory performance, we found no differences between participants who performed the task remotely and those who performed it in the lab (overall item memory d': t_{208} = 0.69; *p* = 0.491; overall proportion correct out of item hits; t_{208} = 0.96; *p* = 0.336). There were also no age differences between these groups (t_{208} = 0.02; *p* = 0.983). Therefore, we decided to collapse across both of these testing modalities in our analyses. We also found no differences in memory performance between the subset of older adults who were recruited from the Penn Alzheimer's Disease Research Center cohort (*n* = 46) and those who were recruited via Facebook (*n* = 18; item memory: t_{62} = 0.97; *p* = 0.334; associative memory: t_{62} = 1.59; *p* = 0.116).

Stimulus ratings differ between faces and houses but not by age. We ensured that participants were attending to the outcomes in the Reward phase by examining their ratings of how they felt after learning about lottery and dictator game outcomes (1 = "good"; 2 = "neutral", 3 = "bad"). Participants felt better after rewarded outcomes compared to unrewarded outcomes, and this difference was more pronounced in the social domain. A repeated-measures ANCOVA with stimulus type (face vs. house) and reward status ($0 vs. $5) as factors, and age as a covariate, revealed that, while there was a significant effect of reward status on ratings, $F_{(1,206)} = 368.51$; $p < 0.001$; η^2_p = 0.641; *n* = 208, with participants feeling more positively about being rewarded $5 compared to \$0, this was qualified by a stimulus type x reward status interaction, $F_{(1,206)} = 6.81$; $p = 0.010$; $\eta^2_p = 0.032$. This interaction was driven both by \$0 faces being rated as more negative than \$0 houses ($t_{206} = 5.37$; $p_{tukey} < 0.001$; estimated marginal means: \$0 face, M = 2.70; SE = 0.03; \$0 house, M = 2.52; SE = 0.03) and by \$5 faces being rated as more positive than \$5 houses ($t_{206} = -5.89$; $p_{tukey} < 0.001$; estimated marginal means: \$5 face, M = 1.13; SE = 0.02; \$5 house, M = 1.29; SE = 0.03). Age did not significantly modulate either of these effects (age x reward status: $F_{(1,206)} = 3.12$; $p = 0.079$; $\eta^2_p = 0.015$; age x reward status x stimulus type: $F_{(1,206)} = 0.59$; $p = 0.442$; $\eta^2_p = 0.003$). Therefore, there is no evidence that older adults were paying less attention to reward outcomes in the Reward phase or that their self-reported reactions to monetary outcomes were different from those of younger adults.

Item and associative memory performance were above chance in all groups. In all three age groups, and for both faces and houses, item memory (d') was above chance (young adults: houses, M = 2.27; SD = 0.85; $t_{75} = 23.4$; $p < 0.001$; faces, M = 2.88; SD = 0.77; $t_{75} = 32.6$; $p < 0.001$; middle-aged adults: houses, M = 2.19; SD = 0.77; $t_{69} = 23.7$; $p < 0.001$; faces, M = 2.66; SD = 0.73; $t_{69} = 32.6$; $p < 0.001$; older adults: houses, M = 1.67; SD = 0.79; $t_{63} = 16.9$; $p < 0.001$; faces, M = 1.95; SD = 0.97; $t_{63} = 16.1$; $p < 0.001$).

Young adults had a liberal response bias for faces (mean criterion = -0.11; SD = 0.38; $t_{75} = -2.53$; $p = 0.013$), being more likely to label stimuli as "old," rather than "new," but this was not the case for houses (mean criterion = 0.08; SD = 0.49; $t_{75} = 1.44$; $p = 0.153$). Middle-aged adults showed a liberal response bias for both faces (mean criterion = -0.41; SD = 0.47; $t_{69} = -7.29$; $p < 0.001$) and houses (mean criterion = -0.36;

SD = 0.48; t_{69} = -6.19; $p < 0.001$). Older adults were also biased toward labeling stimuli as "old" rather than "new," in both social (mean criterion = -0.60; SD = 0.78; t_{63} = -6.14; $p < 0.001$) and non-social domains (mean criterion = -0.35; SD = 0.66; t_{63} = -4.25; $p <$ 0.001).

All three age groups were also above chance at associative memory. In other words, they were better than chance at accurately classifying stimuli as rewarded ($5) or unrewarded ($0) in both the social and non-social domain (young adults: associative memory d' for houses, M = 0.59; SD = 0.87; t_{75} = 5.88; $p < 0.001$; faces, M = 0.52; SD = 0.67; t_{75} = 6.75; $p < 0.001$; middle-aged adults: associative memory d' for houses, M = 0.43; SD = 0.77; t_{69} = 4.75; $p < 0.001$; faces, M = 0.49; SD = 0.63; t_{69} = 6.55; $p < 0.001$; older adults: associative memory d' for houses, M = 0.22; SD = 0.64; t_{63} = 2.71; p = 0.009; faces, M = 0.14; SD = 0.49; t_{63} = 2.26; p = 0.028).

As stated in the main text, young adults had a response bias toward saying that faces shared $0 (mean associative memory criterion = 0.15; SD = 0.49; t_{75} = 2.72; p = 0.008), a bias that did not extend to houses (mean criterion = -0.05; SD = 0.50; t_{75} = -0.84; p = 0.401). In contrast, older adults had a response bias toward saying that faces shared $5 (mean associative memory criterion = -0.22; SD = 0.65; t_{63} = -2.76; p = 0.007). This effect in older adults was also specific to the social domain (mean associative memory criterion for houses = -0.10; SD = 0.63; t_{63} = -1.33; p = 0.189). Middle-aged adults did not show a significant response bias for either faces (mean associative memory criterion = 0.05; SD = 0.57; t_{69} = 0.73; p = 0.466) or houses (mean associative memory criterion = -0.12; SD = 0.54; t_{69} = -1.85; p = 0.068).

Aging impairs decision-making about low-value stimuli. As described in the main text, we ran a repeated-measures ANCOVA with memory level (Confident incorrect associative memory / Item memory only / Confident correct associative memory) and value ($0 / $5) as within-subjects factors, and choice (proportion of times that the “old” stimulus was chosen over the schematic image) as the dependent variable. Age was entered as a covariate. Aging impacted the extent to which the values of the stimuli, as well as memory for the values of the stimuli, affected decision-making: there was a value x age interaction on choice ($F_{(1,163)} = 14.57$; $p < 0.001$; $\eta^2_p = 0.082$) and a value x memory level x age interaction on choice ($F_{(2,326)} = 10.88$; $p < 0.001$; $\eta^2_p = 0.063$). The memory level x age interaction was not significant ($F_{(2,326)} = 0.39$; $p = 0.678$; $\eta^2_p = 0.002$), and the main effect of age was not significant ($F_{(1,163)} = 3.30$; $p = 0.071$; $\eta^2_p = 0.020$).

To understand the nature of the value x age interaction, we ran a post-hoc Pearson correlation between age and (1) the proportion of times a $5 stimulus was chosen, and (2) the proportion of times that a $0 stimulus was chosen. Age was not associated with choice of the $5 stimulus ($r = 0.001$; $p = 0.992$; $n = 210$), but it was associated with choice of the $0 stimulus ($r = 0.38$; $p < 0.001$; $n = 210$), such that older adults were more likely to choose to interact with low-value stimuli. In other words, older adults’ maladaptive decision-making was driven by their tendency to approach low-value stimuli, rather than their avoidance of high-value stimuli.

To interpret the value x memory level x age interaction, we ran a series of post-hoc Pearson correlations, relating age to choices about stimuli in each of the six relevant bins (i.e., confident correct – high and low value, confident incorrect – high and low value, and item memory only – high and low value). After Bonferroni correction for multiple comparisons, we found that age was significantly associated with an increased tendency to select low-value stimuli, both when participants had confident correct associative memory ($r = 0.37$; $p < 0.001$; $n = 199$), and when associative memory responses were guesses ($r = 0.20$; $p = 0.006$; $n = 197$). When deciding about low-value stimuli for which participants had confident incorrect associative memory (i.e., when they thought the $0 stimuli were actually worth $5), there was no relationship between age and choice ($r = 0.04$; $p = 0.566$; $n = 198$). Age was also not associated with the tendency to select high-value ($5) stimuli, whether associative memory was confident and incorrect ($r = 0.11$; $p = 0.140$; $n = 199$), participants were guessing about associative memory ($r = 0.02$; $p = 0.773$; $n = 190$), or associative memory was confident and correct ($r = -0.17$; $p = 0.014$; $n = 202$). Therefore, older adults' impairment in memory-guided decision-making was driven by a maladaptive tendency to approach low-value ($0) stimuli even when they accurately remembered that a stimulus was low-value.

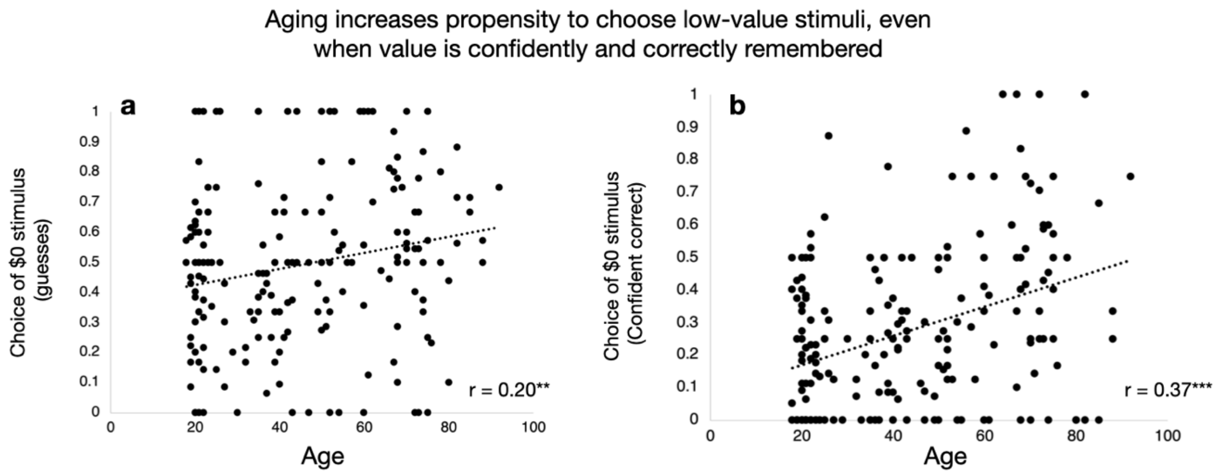

Fig. S1. Age affected the ability to use associative memory to make adaptive decisions, with older adults being less likely to avoid low-value stimuli that they either had no associative memory for (a) or that they had confident and correct associative memory for (b). **p<0.01, ***p<0.001.

Additional ANCOVA results (Social vs. non-social decision-making). When considering the whole sample, participants did not choose faces more than houses (t_{174} = 0.63; p_{tukey} = 0.532). They also did not avoid \$0 faces more than \$0 houses (t_{174} = 12.12; p_{tukey} = 0.151), and they did not approach \$5 faces more than \$5 houses (t_{174} = 1.60; p_{tukey} = 0.382). There was a main effect of stimulus type in the model ($F_{(1,174)}$ = 8.59; p = 0.004; η^2_p = 0.047), but this was qualified by a stimulus type x age interaction ($F_{(1,174)}$ = 8.69; p = 0.004; η^2_p = 0.048), described in the main text. There was also a significant remembered value x age interaction ($F_{(1,174)}$ = 23.71; p < 0.001; η^2_p = 0.120).

Post-hoc correlations with age (Social vs. non-social decision-making). As stated in the main text, there was a significant correlation between age and the tendency to choose faces remembered as being unfair (r = 0.40; p < 0.001; n = 199), but this did not extend to choices about low-value houses (r = 0.04; p = 0.558; n = 187). Not only is this effect of aging on choice specific to the social domain, but it is also specific to the low-value stimuli. The association between age and the tendency to choose stimuli remembered as high-value was negative, and only borderline significant in the non-social domain (r = -0.18; p = 0.012; n = 193) and trending toward significance in the social domain (r = -0.13; p = 0.067; n = 201). Thus, it was not the case that older adults were more likely to select unfair others just because they were more likely to select the stimulus over the schematic image in general.

Associations between perceived generosity and other ratings. The “perceived generosity” rating was correlated with the other ratings, such that faces perceived as more competent (ρ = 0.49; $p = 0.004$), warm (ρ = 0.89; $p < 0.001$), and trustworthy (ρ = 0.88; $p < 0.001$) were perceived as having shared more money in the Dictator game. Perceived dominance (ρ = -0.18; $p = 0.328$) and attractiveness (ρ = 0.28; $p = 0.124$) were not associated with perceived generosity.

Effects of facial appearance on choice are independent of associative memory deficits. If people rely on facial appearance more in their choices because they are compensating for their memory decline, then we might expect to find a significant correlation between associative memory accuracy (d') and the random-effects slopes on these facial attributes, even after controlling for age. This was not the case: perceived generosity partial $r = -0.04$; $p = 0.569$; trustworthiness partial $r = -0.05$; $p = 0.470$; warmth partial $r = -0.06$; $p = 0.419$; competence partial $r = -0.06$; $p = 0.395$; attractiveness partial $r = -0.08$; $p = 0.262$.

However, there was a significant negative relationship between associative memory response bias and the reliance on facial features in choice, even after controlling for age (perceived generosity partial $r = -0.29$; $p < 0.001$; trustworthiness partial $r = -0.28$; $p < 0.001$; warmth partial $r = -0.26$; $p < 0.001$; competence partial $r = -0.26$; $p < 0.001$; attractiveness partial $r = -0.16$; $p = 0.021$), such that those individuals who were more biased toward remembering that partners were generous were also those who engaged more with trustworthy-looking partners. Thus, after accounting for their memory deficits, older adults were not only more likely to remember that people were generous, but they were also more likely to engage with people who *appeared* generous, despite the fact that these appearances were unrelated to their actual behavior.

Effects of face stimulus age, gender, and race on choice. As stated in the main text, there were significant effects of stimulus age ($\beta = -0.020$; $p = 0.001$), gender ($\beta = 0.448$; $p < 0.001$), and race on choice. With respect to race, Black partners were more likely to be chosen relative to White partners ($\beta = 0.331$; $p < 0.001$), while there was no difference between Asian and White partners ($\beta = 0.061$; $p = 0.487$). A follow-up regression with Asian race as the reference level revealed that Black partners were also more likely to be selected compared to Asian partners ($\beta = 0.265$; $p = 0.003$). **Table S1** summarizes these effects and also shows the extent to which they were correlated with the age of the participant. While the effects of participant age were small, they are consistent with older adults relying on irrelevant facial features more in their decisions: they were even more likely to choose female faces and Black faces. For the stimulus age effects, however, the positive sign of the correlation coefficient suggests that older adults were *less* likely than young adults to choose younger-looking faces. Since all of the faces shown were of young adults, however, thus restricting the stimulus age range tested, this finding should be interpreted with caution.

Notably, there were no effects of stimulus race ($F_{(2,29)} = 1.14$; $p = 0.333$) or age ($r = 0.03$; $p = 0.879$; n = 32) on perceived generosity, but female faces were perceived as more generous than male faces ($t_{30} = 2.47$; $p = 0.019$). A follow-up regression that included perceived generosity as an additional covariate found that female stimuli were still more likely to be selected than male stimuli even after accounting for variance related to perceived generosity ($\beta = 0.231$; $p = 0.002$). This suggests that participants are more likely to re-engage with female social partners, even controlling for females' greater perceived generosity.

Table S1. Effects of stimulus age, gender, and race on interaction choices

Attribute of stimulus	Main effect on choice of stimulus (β)	95% CI	Correlation between participant age and subject-specific random slope
Age (range: 18-36)	-0.020**	-0.031 – -0.008	0.189**
Gender (Female relative to Male)	0.448***	0.310—0.586	0.148*
Race (Black relative to White)	0.331***	0.171—0.491	0.154*
Race (Asian relative to White)	0.061	-0.111—0.234	0.171*
Race (Black relative to Asian)[1]	0.265**	0.093—0.438	0.142*

***p<0.001, **p<0.01, *p<0.05;

[1]These values are from a separate regression in which Asian race was coded as the reference level.

Contributors to social avoidance decisions. We ran a multiple linear regression examining the effects of both normative (item and associative memory accuracy) and non-normative (social memory and decision biases) influences on decisions to avoid social partners who were previously unfair. The two strongest predictors of social avoidance behavior were social associative memory response bias (standardized β = 0.41; $p < 0.001$) and the tendency to approach others remembered as being unfair (standardized β = -0.43; $p < 0.001$). In other words, both the tendency toward remembering others as being fair and the inclination toward approaching partners that were remembered as being unfair had large and approximately equal impacts on the successful avoidance of unfair others. The next strongest predictor was associative memory accuracy (standardized β = 0.20; $p < 0.001$). After controlling for those variables, neither item memory accuracy (standardized β = 0.07; $p = 0.177$) nor the influence of perceived generosity (β = -0.07; $p = 0.151$) had a significant effect on avoidance choices. These data tentatively suggest that interventions targeting social motivations are likely to be the most effective at improving older adults' decisions to avoid social partners who were previously unfair.

Control experiment: Assessing memory prior to decision-making does not alter the effects of associative memory on choice. In our main experiment, the Decision task was given prior to the Memory task so that each subject saw each stimulus only once prior to deciding whether or not to interact with that stimulus in the Decision phase. However, this ordering introduces the possibility that participants would give memory responses that were in line with their previous decisions, even if they did not have intact associative memory. If that were the case, then the effects of associative memory on decision-making might be unduly amplified by a bias in memory reporting.

To address this potential issue, we ran a control experiment online via Amazon's mechanical Turk. Seventy-nine adults aged 55+ (mean age = 61.14; SD = 6.79; 62 F, 17 M; race: 68 White non-Hispanic, 3 White Hispanic, 6 Black non-Hispanic, 2 Asian non-Hispanic) performed the study, but six were excluded (see *Exclusion Criteria* in *SI Appendix*), yielding 73 total participants. In this control experiment, the Decision phase was split into two blocks (48 trials each) – one that came prior to the Memory task and one that followed the Memory task. This allowed us to ensure that the effects of associative memory on decision-making were similar whether the decisions were made before or after the memory test. We performed the same ANOVA described in the main text, with memory level (Confident incorrect associative memory / Item memory only / Confident correct associative memory) and value ($5 / $0) as factors predicting choice, but we added in a third within-subjects factor: Decision block (Before Memory Test / After Memory Test). As expected, there was a significant main effect of Value on choice ($F_{(1,19)}$ = 45.45; p < 0.001; η^2_p = 0.705), with people choosing high-value stimuli more than low-value stimuli overall. There was also a significant Memory level x Value interaction ($F_{(2,38)}$ = 38.20; p < 0.001; η^2_p = 0.668), such that associative memory for the item and its value was necessary for approaching high-value stimuli and avoiding low-value stimuli. Critically, however, the results were not at all modulated by *when* decisions about the stimuli were made. There was no significant three-way (Memory level x Value x Decision block) interaction ($F_{(2,38)}$ = 1.98; p = 0.152; η^2_p = 0.094) on choice; nor were there any two-way interactions with Decision block (Decision block x Memory level: $F_{(2,38)}$ = 1.48; p = 0.240; η^2_p = 0.072; Decision block x Value: $F_{(2,38)}$ = 0.22; p = 0.648; η^2_p = 0.011), or a main effect of Decision block ($F_{(1,19)}$ = 0.88; p = 0.359; η^2_p = 0.044) on choice.

We therefore concluded that the finding that associative memory is necessary for adaptive decision-making is not just driven by participants responding to Memory questions in a way that was consistent with their behavior in the Decision task. These results should be interpreted with caution, however, as only twenty participants had data in all twelve bins for this analysis.