The influence of episodic memory decline on value-based choice

Fedor Levin a, Susann Fiedler a and Bernd Weber b

aMax Planck Institute for Research on Collective Goods, Bonn, Germany; bInstitute of Experimental Epileptology and Cognition Research, University Hospital Bonn and Center for Economics and Neuroscience, University of Bonn, Bonn, Germany

ABSTRACT
Recent studies suggest the involvement of episodic memory in value-based decisions as a source of information about subjective values of choice options. We therefore tested the link between age-related memory decline and inconsistencies in value-based decisions in 30 cognitively healthy older adults. Within the pre-registered experiment, the inconsistencies were measured in two ways: i) the consistency between stated preferences and revealed choices; ii) the amount of intransitivities in choice triplets, revealed in a forced paired choice task including all possible pairings of 20 food products. Although no significant association of memory functions to number of intransitive triplets was observed, participants with lower memory scores were more likely to choose the item for which they stated a lower preference. The results suggest a higher noise in the underlying preference signal in participants with lower memory. We discuss the results in the context of the unique needs of elderly consumers.

ARTICLE HISTORY
Received 8 February 2018
Accepted 2 August 2018

KEYWORDS
Episodic memory; aging; value-based choice; transitivity; decision-making

Introduction
The global population is aging rapidly. According to the UN, by 2050 the proportion of people older than 60 is expected to double compared to 2015 (United Nations, 2015). Altered decision-making abilities of older adults (OAs) in domains such as finance (Agarwal, Driscoll, Gabaix, & Laibson, 2009; Hershey, Austin, & Gutierrez, 2015; Li et al., 2015), healthcare (Morrow & Chin, 2015), or consumer choices (Carpenter & Yoon, 2015) may have a negative impact on their wellbeing. Thus, investigating the relation of decline in cognition to decision-making abilities is considered highly relevant (Carpenter & Yoon, 2011; Strough, Löckenhoff, & Hess, 2015). A number of recent empirical studies (Barron, Dolan, & Behrens, 2013; Bornstein & Daw, 2013; Gluth, Sommer, Rieskamp, & Büchel, 2015; Wimmer, Braun, Daw, & Shohamy, 2014; Wimmer & Shohamy, 2012) and reviews (Palombo, Keane, & Verfaellie, 2015; Shadlen & Shohamy, 2016) suggested an important role of episodic memory in value-based learning and value-based choices. These findings are relevant to research on decision-making in aging, as episodic memory starts decreasing at the age of approximately 60–65 with...
large individual differences (Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012). However, the precise impact of this age-related decline on value-based choices is unclear. In the current study we address the relationship between episodic memory function and value-based decision-making of OAs using a food choice task.

In value-based decisions, an individual chooses between the available options according to their subjective values, which can be influenced by previous experiences (Rangel, Camerer, & Montague, 2008). Our research is in line with the framework of value-based decision-making proposed by Rangel et al. (2008), which assumes several computational steps within a decision: (1) representation of the choice as well as internal and external states, (2) valuation of the alternatives, (3) comparison of their values, (4) evaluation of the outcome of the choice and, finally, (5) learning process aimed at improving future behavior. Valuation of alternatives and the learning process potentially depend on memory, since they involve memorizing and, consequently, retrieval of information related to the properties of choice alternatives.

Our approach to testing the link between memory and value-based choice was additionally informed by the Query Theory (QT) framework (Weber et al., 2007). The QT assumes that preferences of the decision-maker are not immediately known to the decision-maker or are not sufficiently precise. Hence, decision-makers construct preferences during the choice process by retrieving memories related to alternatives under consideration. As a result, preference-based decisions rely on memory and follow the same dynamics and biases as memory processes, such as, for example, output interference (Weber et al., 2007). Previous empirical studies tested whether the QT framework could account for behavioral effects observed in a set of value construction decisions such as the endowment effect (the difference in willingness to pay between owners and buyers, Johnson, Häubl, & Keinan, 2007) or asymmetry in delay discounting (Weber et al., 2007). Results within the domain of delay discounting suggest, for example, that individuals show more patient behavior under conditions when they were thinking about more distant outcomes first. Building on this previous work we adopted the general assumptions of the QT framework about the processes of preference construction and their dependence on memory.

Recent research provided findings supporting this role of memory in value-based choice. Wimmer and Shohamy (2012) demonstrated how the hippocampus, a key structure involved in episodic memory, contributes to the spread of reward value across related memories. They observed a behavioral bias toward non-rewarded stimuli that had been presented together with rewarded stimuli earlier. This bias was associated with hippocampal activation and functional connectivity between the hippocampus and the striatum. These findings provided support for the role of the hippocampal memory system in value-based choice. Further evidence was offered by research showing altered feedback-driven learning and disadvantageous decisions in patients with medial temporal lobe (MTL) damage suffering from amnesia (Foerde, Race, Verfaellie, & Shohamy, 2013; Gupta et al., 2009). A recent study assessed an effect of the episodic future imagination on delay discounting in participants with subjective cognitive decline (SCD) who reported a decrease in episodic memory performance (Hu et al., 2017). The results showed increased delay discounting and a decreased effect of episodic future imagination in the SCD group as compared to a cognitively healthy control group. Shadlen and Shohamy (2016) reviewed the evidence for the role of episodic memory and the hippocampus in decision-making. They suggested that assignment of values to options under consideration in value-based
decisions involves retrieval of value-relevant memories and experiences in a way that is similar to the process of evidence accumulation described by sequential sampling models (SSMs). SSMs have previously been applied to perceptual and preference-based decisions (Krajbich, Lu, Camerer, & Rangel, 2012; Philiastides & Ratcliff, 2013; Polania, Moisa, Opitz, Grueschow, & Ruff, 2015; Ratcliff, Thapar, & McKoon, 2011), and describe them as an evidence accumulation process for the choice options. A decision is made when the amount of evidence reaches a specific threshold. The account by Shadlen and Shohamy (2016) thus offers an interesting perspective on the role of memory processes in choice and allows to make predictions about regularities of value-based choice. To summarize, recent studies have provided considerable evidence for neural mechanisms linking memory processes to value-based choice. A decrease in memory performance could lead to decreased choice quality.

Building upon the existing research which links episodic memory to value-based choice, we hypothesized that heterogeneity in episodic memory performance would be related to differences in retrieval processes of value-related information during choice. Lower episodic memory would thus lead to less reliable value construction. In turn, this would decrease efficiency of decision-making, leading to less accurate and slower value-based choices. There are different approaches to measuring accuracy of value-based choices, which vary in how they define optimal choice. We combined several approaches in order to comprehensively describe the link between memory and quality of choice. In the next section we outline these approaches to assessing accuracy, as well as specific hypotheses and our analysis plan. One approach to measuring precision of value-based choices is the assessment of choice accuracy based on stated personal preferences. Some previous studies collected independent liking ratings of items (Milosavljevic, Malmaud, & Huth, 2010) or bids on a Becker–Degroot–Marshak (BDM) auction (Krajbich et al., 2012). This allowed to classify choices as correct or incorrect based on whether an item with a higher stated preference is picked in a forced choice task. High differences between stated preferences of alternatives represented choices with low difficulty, whereas low differences between stated preferences represented choices with high difficulty. This approach allows testing factors that affect accuracy on trial level, while also taking into account choice difficulty and other relevant variables that differ between trials. Additionally, it is possible to evaluate the overall accuracy of a decision-maker by calculating the rate of errors in choices. We formulated three hypotheses with regard to this measure. We hypothesized that episodic memory performance would be related to accuracy. Specifically, we predicted that participants with lower memory would be more likely to choose inconsistently with their independently stated liking ratings, and would commit more errors—decisions where a food item with a lower stated preference is selected—than participants with higher episodic memory (Hypothesis 1a). Furthermore, we tested whether inconsistent choices (errors) were more likely when differences between ratings were low rather than high (Hypothesis 1b). For example, choices between products that are very similar in their values and are thus hard to discriminate would more often result in errors. We planned this test as confirmatory analysis because it allowed to link the errors that we observed to the amount of value-relevant information in support of presented options. We additionally tested whether an interaction effect between memory and rating differences predicted the probability of making an error (Hypothesis 1c). We expected that
better episodic memory would not provide a strong advantage when choice options are close to identical in their subjective value, because there would be relatively little value-relevant information available for retrieval to help discriminate between the options. Additionally, people with lower episodic memory performance would have a smaller improvement of accuracy in easier trials, which for them would result in a weaker relationship between difficulty and accuracy, than for people with higher memory performance. Our prediction is partially in line with findings from a study by Polanía et al. (2015), who used transcranial alternating current stimulation (tACS) to influence synchronization between the frontopolar and the fronto-parietal areas of the cortex. The stimulation produced a larger decline in accuracy of value-based choices in trials with highest value differences. The authors interpreted these value differences as corresponding to intermediate levels of evidence. We also adopted another approach to assessing accuracy using transitivity of value-based choices (Fellows & Farah, 2007; Lee, Amir, & Ariely, 2009). For example, after choosing option A over option B and B over C, subjects were expected to choose A over C. Selecting C in favor of A would thus constitute an intransitive choice. Since a triplet of choices is a minimal structure that can be intransitive, it is possible to calculate intransitive triplets among choices between all possible combinations of items and use them to construct a measure of choice accuracy. The advantage of the measure of intransitivities is that it does not require any assumption about the link between liking ratings and the respective choices. Testing the relationship between memory performance and the measure of intransitivity would allow us to explore whether the effect of memory on accuracy persists with the measure based only on choices and independent of liking ratings. We hypothesized that participants with lower episodic memory would make more internally inconsistent choices and display a higher rate of intransitivities (Hypothesis 2). One disadvantage of the intransitivity measure is that it is not straightforward for use on the level of trials. Therefore, in order to complement the analysis of intransitivities, we also assessed a test–retest measure of reliability (Brown & Peterson, 2009) as a part of exploratory analysis. Presenting participants with the same sets of choices twice allowed us to test whether memory performance predicted a reversal of the initial choice. A participant with lower choice accuracy would be more likely to choose differently when confronted with a same pair of products a second time. We also tested whether the overall frequency of choice reversals by participants was predicted by memory performance. We expected that lower memory performance would be associated with lower test–retest reliability, both on the level of trials and the level of individuals.

With regard to the speed of information processing, we expected that differences in retrieval of value-related information would also have an effect on reaction time. We hypothesized that increased average reaction times would be linked to reduced episodic memory confirming slower processing in OAs with reduced memory (Hypothesis 3a). Thus, a participant with lower episodic memory performance and a lower efficiency of memory retrieval processes would take longer to make choices. Similarly to Hypotheses 1b and 1c, we tested whether reaction times depended on rating differences with slower decisions between similarly valued options (Hypothesis 3b) consistently with the previous studies on value-based choice (Krajbich, Armel, & Rangel, 2010; Milosavljevic et al., 2010). As an example, a decision in a trial with high difficulty would take longer than in an easy trial. We also tested whether an interaction between memory and rating
differences indicated a weaker relationship between rating differences and reaction times for people with reduced memory (Hypothesis 3c). The rationale for this prediction was similar to the rationale for the Hypothesis 1c—the advantage of higher memory performance for the speed of choices would be most apparent in the trials with a relatively lower difficulty, as compared to trials with the highest difficulty.

Apart from our measures of interest, several other individual difference factors could influence processes of decision-making. As has been shown by previous literature, executive function is related to aspects of decision-making (Del Missier, Mäntylä, & Nilsson, 2015) and memory (Bouazzaoui et al., 2013). We therefore included the measure of the executive function as control for all analyses. Additionally, the literature has described an age-related decline of psychomotor functions (Lim, et al., 2012). We added the psychomotor function measured with a simple reaction time task as a control. Previous studies using eye-tracking suggested that stimuli difference measures (position of the chosen food item, trial number) could potentially influence attention to choice options (Orquin & Mueller Loose, 2013). Therefore, in multilevel trial analyses we controlled for these stimuli difference measures.

Method

The study was approved by the ethics committee of the University of Bonn. All participants signed the informed consent before taking part in the study. The sample size was determined a priori based on practical considerations such as anticipated recruitment rate as well as time constraints. Sample size was pre-registered together with all hypotheses and analysis plan before the start of data collection at the Open Science Framework (https://osf.io/dpfn7/register/565fb3678c5e4a66b5582f67). The complete instructions (https://osf.io/y695g/), data (https://osf.io/h6jat/) and analysis script (https://osf.io/rqtu9/) are also available at the Open Science Framework.

Participants

Thirty-seven OAs participated in the study. We recruited participants from the general population via local community organizations, with the help of flyers, email-based advertisements and by word of mouth. Based on pre-specified selection criteria, six participants were excluded from analyses due to having a BDI (Beck Depression Inventory—II; Kühner, Bürger, Keller, & Hautzinger, 2007) score greater than 14 and thus meeting criteria for depression. This exclusion criterion was necessary since depression is associated with cognitive impairments (Rock, Roiser, Riedel, & Blackwell, 2014). Additionally, one participant reported intake of a central nervous system-active medication and was excluded. The final sample consisted of 30 participants (63% female). Ages ranged from 65 to 88 years ($M = 74.6$, $SD = 6.1$). All participants had normal or corrected-to-normal vision.
Material

ISLT

To measure episodic memory performance, we used the delayed *International Shopping List Test* (ISLT) score. We administered the German version of the ISLT using the computerized test battery Cogstate Research™. ISLT is a verbal learning task with three learning trials and a delayed recall trial (Lim, Pietrzak, Snyder, Darby, & Maruff, 2012; Lim et al., 2009). The delayed ISLT was measured as a number of words remembered in the delayed recall 20 min after the last learning trial.

MoCA

We additionally administered the *Montreal Cognitive Assessment* cognitive screening test (MoCA) to rule out dementia or mild cognitive impairment (MCI). Only 43% of our sample have reached a cut-off threshold recommended in the original validation study by Nasreddine et al. (2005). This could be explained by a relatively low specificity of MoCA when used with a standard threshold. The original study (Nasreddine et al., 2005) reported specificity of 87%. However, in later studies using the MoCA (Rossetti, Lacritz, Cullum, & Weiner, 2011; Tiffin-Richards et al., 2014), specificity was lower—only 38% of participants reached the standard threshold in the study by Rossetti et al. (2011; note that participants were not screened for cognitive impairment) and about 57% of healthy participants reached the standard threshold in the study by Tiffin-Richards et al. (2014). Consequently, in order to control for possible differences related to the overall cognitive performance, we added the MoCA score as a control variable to our analyses while still including participants who scored below the threshold.

Additional cognitive assessment

To control for possible alternative mechanisms underlying the variability in reaction times, we additionally measured the psychomotor function using the Detection task (DET) from the Cogstate Research™ software (Maruff et al., 2013).

In order to control for executive function, participants completed the *Trail Making Test* (TMT; Bowie & Harvey, 2006). TMT consists of two parts—TMT A and TMT B. This data allowed to calculate two measures of executive function—the TMT ratio score defined as a ratio between TMT B and TMT A, and the TMT difference score defined as a difference between TMT B and TMT A.

Behavioral task

In order to record food choices we developed a computer-based task consisting of four parts (see Figure 1). First, participants rated 50 food stimuli according to how much they liked them on a discrete scale from 0 to 10. Based on these ratings, 20 of the 50 items were selected to ensure that a set of products used in the paired choice task included items with various liking ratings, and without the experiment taking too long. The
algorithm ordered the list of 50 products by liking ratings and then picked every second product ignoring the 10 lowest-rated products.

In the food choice task, participants then chose their preferred products from a randomized sequence of all possible pairs of the selected 20 products, which amounts to 190 pairs. In order to assess test–retest reliability we presented the same choice pairs at a later point of the experiment again. Choice pairs had a new presentation order and reversed display sides in the retest. Participants chose products by pressing buttons on a keyboard corresponding to the left or the right product. There were no enforced time limits on trials, and participants were instructed to press a button as soon as they made a decision. Participants were informed that at the end of the session they would receive a product of their choice from one randomly picked trial. Since any trial could be selected, it provided an incentive for participants to treat each choice as equally important.

Figure 1. Panel A depicts the procedure of administering the tests. Panel B demonstrates the behavioral task. In the first part participants rated 50 food stimuli on a scale from 0 to 10. Next, they made choices between all possible pairings of 20 food stimuli selected based on their ratings ($N = 190$). In a number comparison task participants were then asked to select the higher number of two. Lastly, they were presented again with the choice pairs from the second part of the experiment, but in a re-randomized order. Brand names are covered on this figure, but they were present in the task. Panel C demonstrates a trial from the paired choice task. Reaction times for paired choices were not limited. On this figure the fixation cross is presented larger, than it was in the task, for demonstration purposes.
In order to test participants’ ability to make comparisons without the need for retrieving preferences, we administered a simple paired values choice task between the test and the retest of the food choice task. It included overall 10 prices of the featured food products (with two decimal places, range 0.79–2.59) presented in all possible pairings, which amounted to 45 trials. Participants were instructed to select the highest number in each trial. The set of numbers was identical for each participant. The behavioral paradigm was presented using in-house software of the Department of NeuroCognition/Imaging, Life & Brain based on Python.

**Overall procedure**

The study took overall between 70 and 120 min and participants received a fixed payment of 25 Euro for their participation. To incentivize their choices, we additionally allowed them to receive one product of their choice from one randomly selected choice trial of the experiment. During the study, participants were shown an assortment of food before the beginning of the session to demonstrate that the products were available in the lab. They read instructions describing the procedure and after that MoCA was administered. Participants completed the initial list learning part of the ISLT with three repetitions of the list. The retrieval part of the ISLT was conducted after a 20 min delay. During the delay, participants completed a set of other tasks. Within the Cogstate Research™ they completed the DET. Then participants answered questions about their age, level of education, whether they had neurological or psychiatric disorder, and whether they were taking central nervous system-active medications. Next, they completed the TMT, the BDI, and the delayed part of the ISLT. Following these tests, participants performed the behavioral task consisting of the liking rating task, the food choice task, number comparison task and the retest part of the food choice task (Figure 1). At the end of the session participants received their payment and a product of their choice from one randomly selected food choice trial.

**Data preparation**

Both parts of the paired choice task—test and retest—were used in the analyses. We used data from the paired choice task to examine triplets of choices and calculate percentages of intransitive triplets for each subject. According to the pre-registered criteria, reaction time outliers and other nonviable trials were excluded.

**Reaction times**

As pre-registered, we excluded trials with reaction times shorter than 300 ms and reaction time outliers (± 3 SD) to remove trials in which participants most likely did not pay attention. This criterion applied to 1.8% of all trials.

**Error measure**

We identified trials in which participants picked a product with a lower rating. For the error measure analysis, in addition to excluding trials based on reaction time criteria, we excluded trials in which products had the same ratings, since errors could not be
identified for those trials. Overall, 16.7% of trials met this criterion. To account for different numbers of trials excluded for each participant, we calculated the error measure as a percentage of error trials among all included trials. Higher values on this measure represented higher error rates and less accurate decision-making.

**Missing data**

One participant’s responses in the rating task had likely been caused by accidental presses on a key not intended for use in the task, and were then registered as a series of entries unrelated to the liking rating scale with implausibly short reaction times. This affected 8 out of 20 preselected products and subsequently 65.3% of this participant’s trials. Hence, the participant was excluded from all error measure analyses and mixed effects analyses involving rating differences. This participant was included in the analyses of intransitive triplets, average reaction times and percentages of choices changed in retest since these measures did not depend on product ratings. For another participant, the DET measure in the testing program output was missing most likely due to incorrect execution of the task, and was excluded from all analysis including DET.

**Intransitivity**

We calculated percentages of intransitive triplets among all triplets for each participant. For example, a triplet containing trials with choices between products A, B and C was

![Figure 2](image-url)

**Figure 2.** Linear regression models with the percentage of errors in the food choice task and percentage of errors in the number comparison as dependent variables and delayed ISLT as an independent variable. Fitted linear regression line for the percentage of errors in the food choice task: $r = -.25$, $p = .199$. Fitted linear regression line for the percentage of errors in the number comparison task: $r = -.03$, $p = .863$. 

Intransitivity We calculated percentages of intransitive triplets among all triplets for each participant. For example, a triplet containing trials with choices between products A, B and C was
labeled intransitive if product A was selected over product B, product B was selected over product C, but product C was selected over product A. Each sequence of 190 trials produced 1140 possible triplets and percentages of intransitive triplets were averaged across the two sequences of choices for each participant.

**Data transformation**

Due to the skewed distribution of reaction time data, we log-transformed it for all analyses. In all regression models we centered independent variables around the mean before including their respective interaction in the analysis. Additionally, for the logistic regression model predicting errors in trials, we mean centered the two control variables TMT difference and presentation order, as well as rescaled them dividing by 100 in order to ensure the convergence of the model. For the linear regression model predicting reaction times, we mean centered and rescaled the presentation order variable.

**Results**

Regression analyses included models specified on the level of individuals as well as mixed effects models specified on the level of trials, which included random intercepts for participants in order to account for individual differences. Reported linear repeated measurement mixed effects regression models were fit by the restricted maximum likelihood (REML) and associated t-tests used Satterthwaite approximations for degrees of freedom. Descriptive statistics for demographic variables and collected measures are presented in Table A1. Coefficients of correlations between cognitive and behavioral task measures are presented in Table A2. Additionally, frequencies of MoCA scores are presented in Figure A2 and frequencies of delayed ISLT scores are presented in Figure A3.

**Choice analysis**

We examined the relationship between episodic memory and the percentage of deviations of choices from previously stated preference ratings (percentage of errors) using a linear regression model on an individual level with the percentage of errors as the dependent variable, delayed ISLT as the independent variable and controlling for MoCA scores. The delayed ISLT was significantly negatively associated with percentages of errors confirming that better memory performance was associated with a lower rate of errors (Table 1).

Surprisingly, MoCA scores included for control were related to proportions of errors and longer average reaction times; however, the effect was absent in multilevel models. We tested the relationship between trial difficulty, episodic memory and errors in a repeated measurement logistic regression that predicted error in a trial by differences in product ratings, delayed ISLT and their respective interaction (Table 2). Confirming Hypothesis 1b, the analysis showed a significant effect of rating differences on probability to make an error, meaning that as trials become more difficult, participants are more likely to make errors (Figure A1). To test whether lower memory would correspond to a weaker relationship between the delayed ISLT and rating differences signifying stronger effects of ISLT on easier trials (Hypothesis 1c), we included the interaction effect
of delayed ISLT and rating difference. The results showed a significant interaction effect implying a stronger relationship between the delayed ISLT and rating differences for individuals with lower memory. The finding indicated a larger effect of memory on the probability to make errors in more difficult trials and contradicted our prediction made with regard to Hypothesis 1c.

Comparison between these results and the performance in the number comparison task indicated that errors in the food choice task were not due to deficiencies in simple comparisons, but due to differences specific to value-based decision-making (Figure 2). Nine of the participants made one error each and one participant made three errors out of 45 trials. The rest of the participants did not make mistakes.

Next, we assessed the relationship between episodic memory and intransitivity using a linear regression model on an individual level with intransitivity as the dependent variable, delayed ISLT as the independent variable and MoCA score as a control variable. Contrary to Hypothesis 2, our results did not show a significant relationship between episodic memory and percentage of intransitive triplets (Table 1).

### Reaction time analysis

We next tested the relationship between episodic memory and reaction times using a linear regression model on an individual level, with the average reaction times as the dependent variable, delayed ISLT as the independent variable and MoCA scores as a control variable (Table 1). The analysis showed a significant inverse relationship between delayed ISLT scores and average reaction times in accordance with our Hypothesis 3a. We also tested the relationship between trial difficulty, episodic memory and reaction time on each trial, using a repeated measurement linear regression model predicting transformed trial reaction time by differences in ratings of products, delayed ISLT and their respective interaction (Table 2). The analysis revealed a significant effect of difficulties on reaction times—higher difficulties were associated with longer reaction time—confirming Hypothesis 3b. Additionally, the results showed a significant interaction effect of delayed ISLT and reaction times. The positive interaction effect indicated that lower delayed ISLT scores corresponded to a stronger relationship between difficulty and reaction times, which does not support Hypothesis 3c.

### Table 1. Linear regression analyses predicting percentages of errors, percentages of intransitive triplets and average reaction times from delayed ISLT.

|                          | Percentage of errors                       | Average reaction time                      |
|--------------------------|--------------------------------------------|--------------------------------------------|
|                          | Simple model | With controls | Percentage of intransivities | Simple model | With controls |
| **Delayed ISLT**         | 1.51* (−2.26) | −1.51* (−2.12) | 0.03 (−0.15) | −178.54** (−3.62) | −177.4*** (−4.48) |
| **MoCA score**           | 1.24* (2.7)  | 1.25* (2.11)  | 0.03 (0.2)   | 72.63* (2.12)    | 49.93 (1.48)      |
| **TMT difference**       | 0.01 (0.03)  | −1.92 (−1.36) | 4721.32*** (6.46) |
| **Detection task speed** |                     |                         |               |                   |
| **Constant**             | −1.24 (−0.11) | −1.55 (−0.09) | 1.88 (0.57)  | 1408.27 (1.74)   | −9902.4*** (−3.75) |

Note. Unstandardized estimates are presented with t-statistics in parentheses. *p < .05, **p < .01, ***p < .001. ISLT = International Shopping List Test; TMT = Trail Making Test; MoCA = Montreal Cognitive Assessment. Percentage of errors was calculated as a percentage of trials in which a product with lower rating had been chosen among all trials. Percentage of intransitivities was calculated as a percentage of intransitive choice triplets among all choice triplets.
Controls

In order to control for additional inter-individual differences which might be crucial for value-based decision-making and our experimental task in particular, we additionally collected individual difference measures of psychomotor (DET) and executive function (TMT ratio and TMT difference), as well as stimuli difference measures (position of the chosen food item, trial number). The reaction time in the Detection task (Psychomotor function) was significantly related to both average reaction times and reaction times for each trial, but did not affect the predictive power of delayed ISLT or the rating difference (see Tables 1 and 2).

To explore the links between age, cognitive measures as well as decision-making, additional correlation analyses were run (see Table A2). Notably, older age was associated with lower episodic memory performance ($r(28) = -0.52, p = .003$), lower executive function as measured by TMT difference ($r(28) = 0.42, p = .022$), a higher percentage of intransitivities ($r(28) = 0.38, p = .041$) and a higher rate of choice changes in retest ($r(28) = 0.38, p = .036$).

Table 2. Logistic and linear random intercept regression analyses predicting errors and reaction times from differences in ratings of products, delayed ISLT and their respective interaction.

| Estimate                    | Errora                      | Log-transformed reaction timeb |
|-----------------------------|-----------------------------|--------------------------------|
|                             | Simple model | With controls | Simple model | With controls |
| Delayed ISLT                | 0.04 (0.76) | 0.04 (0.71)   | -0.07** (−3.07) | -0.07*** (−4.18) |
| Rating difference           | -0.51*** (−25.79) | -0.51*** (−25.51) | -0.03*** (−18.40) | -0.03*** (−19.11) |
| Rating difference × delayed ISLT | 0.06*** (6.05) | 0.06*** (6.11) | 0.01* (2.39) | 0.01* (2.51) |
| MoCA score                  | 0.02 (0.66) | 0.01 (0.2) | 0.02 (1.52) | 0.01 (0.66) |
| Detection task speed        | 2.05 (1.53) | 2.05 (1.53) | 2.21*** (5.17) | 2.21*** (5.17) |
| TMT difference              | -0.11 (−0.58) | 0.01 (−1.89) | 0.01 (−1.89) | 0.01 (−1.89) |
| Right product selected      | 0.03 (0.41) | -0.02** (−3.33) | -0.02** (−3.33) | -0.02** (−3.33) |
| Presentation order          | 0.06* (2.09) | -0.1*** (−30.26) | -0.1*** (−30.26) | -0.1*** (−30.26) |
| Constant                    | -1.98*** (−18.51) | -2*** (−18.07) | 6.8*** (17.71) | 6.8*** (17.71) |
| Observations                | 9009 | 8680 | 10,819 | 10,445 |

Note. Unstandardized estimates are presented with z statistics in parentheses for the models predicting errors and with t-statistics in parentheses for the models predicting reaction time. * $p < .05$, ** $p < .01$, *** $p < .001$.
ISLT = International Shopping List Test; TMT = Trail Making Test; MoCA = Montreal Cognitive Assessment.

aRating differences equal to zero were excluded. Rating differences and delayed ISLT were mean centered, reaction times were log-transformed due to skewness of the distribution. MoCA scores and Detection task speed were mean centered, TMT difference and presentation order were mean centered and rescaled by dividing by 100 to ensure convergence of the model. Errors were defined as trials in which participants had picked a product with a lower rating.

bPresentation order was mean centered and rescaled by dividing by 100 to ensure convergence of the model.

Analyses predicting changes of choice in retest

In order to explore whether episodic memory performance also influenced the probability of a choice reversal between the test and retest measure of the behavioral task, two regression models were specified (Table A4). One model used the individual measure of test–retest reliability, which was calculated as a percentage of choices that differed between the first and the second runs of the paired choice task. This analysis did not show an effect of episodic memory. Utilizing the repeated measurement structure of the data, we specified a mixed effects logistic regression model predicting on a trial level whether the choice was changed in the retest part of the task. In line with the analysis predicting errors on a trial level, the results of the repeated measurement
random intercept logistic regression showed no relationship between delayed ISLT and choice reversals, but identified as predictors task difficulty ($B = -0.27, z = -10.47, p < .001$) and its interaction with delayed ISLT ($B = 0.04, z = 2.68, p = .007$; see Table A4). Meaning, that participants were more likely to reverse choices made in trials with higher difficulty. The effect of the interaction implies a stronger negative relationship between memory and probability to change the choice in the retest in harder trials, as opposed to easier trial.

**Discussion**

Recent advances describing the role of episodic memory in value-based choices as a source of preference-related information have the potential for improving the understanding of decision-making of OAs. Since episodic memory declines with age, it has been hypothesized to lead to less accurate value-based decision-making of OAs. The current study assessed the relationship between episodic memory performance and value-based decisions of community-dwelling OAs. A simple binary food choice task was used to test whether episodic memory performance predicted accuracy and speed of decisions made by OAs. Results provided partial support of this assumption, in that OAs with lower memory performance exhibited more inconsistencies between their stated preferences and their actual choices. Even though our results showed a link between episodic memory and the percentages of errors, we found no support for the hypothesized inverse relationship between episodic memory and choice consistency as measured by the percentage of intransitive triplets of choices.

We confirmed an association between lower memory performance and increased average reaction times, indicating slower memory retrieval. This finding makes it unlikely that higher rates of errors are explained by a speed-accuracy trade-off. However, there is still a possibility that a speed-accuracy trade-off somewhat improved accuracy with even lower reaction times. Similarly, in a study by Lighthall, Huettel, and Cabeza (2014) OAs took more time to make memory-dependent choices than younger adults, which allowed them to reach similar accuracy.

We also confirmed an inverse relationship between trial difficulties and the performance measures of speed and accuracy. Choices between products closer in stated subjective preference were slower and more likely to result in a choice of the previously less preferred product. These results replicate consistent findings from previous literature on economic decision-making (Milosavljevic et al., 2010; Oud et al., 2016; Polanía et al., 2015). This effect of difficulty indicates that observed errors and reaction times relate to differences in subjective valuations of alternatives, and thus likely reflect properties of value construction processes.

We predicted that a relationship between episodic memory and accuracy as well as speed would differ with trial difficulty. This relationship would reflect an interaction between evidence availability, as determined by trial difficulty, and potential restrictions imposed by memory function. We hypothesized that, in addition to lower overall accuracy and speed, people with lower memory performance would display relatively smaller improvements in easier trials compared to people with higher memory. However, present results suggested a different effect and stronger memory-related differences in accuracy and speed in harder trials, rather than in easy trials. We also observed effects of trial difficulty
and an interaction between memory performance and trial difficulties in the model predicting reversal of choice in the retest part of the paired choice task.

One potential limitation of the current study was that valuations of products and, consequently, difficulties of trials, differed to some extent from person to person, even though they were based on the pre-selection based on ratings. This has likely decreased the precision of measures on the individual level; however, the repeated measurement models accounted for ratings of products and interactions with memory. An additional limitation might be the slightly unbalanced composition (63% female) of the participant sample. Previous research has shown that female participants perform better on episodic memory tasks, and this advantage is retained in older age (Herlitz, Nilsson, & Bäckman, 1997; Kramer, Yaffe, Lengenfelder, & Delis, 2003). The effect depends on the type of the task, with a maximal advantage observed for tasks involving verbal material such as word lists (for a review, see Herlitz & Rehnman, 2008). This protective effect of gender on episodic memory performance is potentially relevant to the interpretation of the current study since the measure of verbal memory, the delayed ISLT, could have slightly overestimated the episodic memory of female participants and this could have affected results from multiple regression models. However, this reported effect of gender on memory is small (Herlitz & Rehnman, 2008) and we refrained from testing gender effects. One potential direction of future research could address a question about whether male OAs are more vulnerable to the effect of memory on decision-making due to lower average memory performance. Another limitation was that we measured episodic memory performance at a given time point rather than longitudinally. An inverse relationship between episodic memory and age was consistent with the previous literature though (Bouazzaoui et al., 2013; Nyberg et al., 2012), and indirectly supported an assumption that memory heterogeneity in our sample was related to age. Further research—including, for example, neuroimaging methods—would be needed to confirm contributions of specific neural structures and their activity to properties of choice behavior. Another potential future research direction would be to allow participants to learn properties of novel choice options in a controlled fashion (for example, similarly to the procedure used in the study by Lighthall et al., 2014). Such an approach with a greater control over option values and trial difficulty would allow to test how memory performance and forgetting influence retrieval of value-relevant information. Assessment of the learning and forgetting of new information, for example, introduced in a form of a dietary advice, could contribute to identification of factors underlying dietary adherence and thus link the findings about the role of memory in choice to practical health-related outcomes.

**Conclusion**

The present work provides novel findings on the association between episodic memory and performance on a value-based task. They are broadly consistent with previous experimental studies and reviews that proposed involvement of memory and related neural structures in value-based decisions (Barron et al., 2013; Bornstein & Daw, 2013; Gluth et al., 2015; Shadlen & Shohamy, 2016; Wimmer et al., 2014; Wimmer & Shohamy, 2012). Previous literature has described decision-making deficiencies in learning (Foerde et al., 2013) and advantageous complex decision-making (Gupta et al., 2009) in individuals with amnesia.
due to hippocampal or MTL damage. Overall, our results suggest that the quality of
decision-making is linked to variations in memory performance in the absence of overt
amnestic symptoms. We employed a food choice task, but similar regularities can be
potentially expected in other value-based decisions relying on common mechanisms
(Rangel et al., 2008). The results have implications for research on decision-making in
aging and, more specifically, for mechanisms of value-based decisions as well as decision
quality in various areas. Within the domain of food choice, one relevant topic is adherence
to dietary guidelines, which has been linked to health outcomes in OAs (Jankovic et al.,
2014) including cognitive health (Gopinath, Russell, Kifley, Flood, & Mitchell, 2015). Further
research is warranted in order to assess the impact of episodic memory on consumer
satisfaction and potential personal losses due to less efficient decision-making.

Notes
1. For the first seven participants, MoCA was administered after the retrieval part of the ISLT and
controlling for the order of administration in the analysis did not indicate an effect of the
order change.
2. For this model, we excluded the participant for whom the proportion of errors was likely
affected by a high number of trials excluded due to missing ratings. For completeness we
performed additional analysis including this participant and regression model yielded similar
estimates (Table A3).
3. DV indicated if participants changed their choice between test and retest even though the
same pair of two items was presented.

Acknowledgement
The authors thank Dr Peter Trautner for programming the behavioral task and Minou Ghaffari for
helpful comments.

Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Fedor Levin http://orcid.org/0000-0002-0518-1715

References
Agarwal, S., Driscoll, J. C., Gabaix, X., & Laibson, D. (2009). The age of reason: Financial decisions
over the life cycle and implications for regulation. Brookings Papers on Economic Activity, 2009
(2), 51–117. doi:10.1353/eca.0.0067
Barron, H. C., Dolan, R. J., & Behrens, T. E. J. (2013). Online evaluation of novel choices by
simultaneous representation of multiple memories. Nature Neuroscience, 16(10), 1492–1498.
doi:10.1038/nn.3515
Bornstein, A. M., & Daw, N. D. (2013). Cortical and hippocampal correlates of deliberation during
model-based decisions for rewards in humans. PLoS Computational Biology, 9(12), e1003387.
doi:10.1371/journal.pcbi.1003387
Bouazzaoui, B., Fay, S., Taconnat, L., Angel, L., Vanneste, S., & Isingrini, M. (2013). Differential involvement of knowledge representation and executive control in episodic memory performance in young and older adults. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 67*(2), 100–107. doi:10.1037/a0028517

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the trail making Test. *Nature Protocols, 1*(5), 2277–2281. doi:10.1038/nprot.2006.390

Brown, T. C., & Peterson, G. L. (2009). An enquiry into the method of paired comparison: Reliability, scaling, and thurstone’s law of comparative judgment. *Geneneral Technical Report RMRS-GTR-216WWW*. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station (January), 1–104.

Carpenter, S. M., & Yoon, C. (2011). Aging and consumer decision making. *Annals of the New York Academy of Sciences, 1235*(1), 1–12. doi:10.1111/j.1749-6632.2011.06390.x

Carpenter, S. M., & Yoon, C. (2015). Aging and consumer decision making. In: *Aging and decision making* (pp. 351–370). Elsevier. doi:10.1016/B978-0-12-417148-0.00017-0

Del Missier, F., Mäntylä, T., & Nilsson, L.-G. (2015). Aging, Memory, and Decision Making. In Aging and Decision Making: Empirical and Applied Perspectives, eds T. M. Hess, C. E. Loeckenhoff, and J.-N. Strough (Vol. 1235, pp. 127–148). Elsevier Academic Press. doi: 10.1016/B978-0-12-417148-0.00007-8

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? *Cerebral Cortex, 17*(11), 2669–2674. doi:10.1093/cercor/bhl176

Foerde, K., Race, E., Verfaellie, M., & Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: Evidence from amnesia. *The Journal of Neuroscience, 33*(13), 5698–5704. doi:10.1523/JNEUROSCI.5217-12.2013

Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory. *Neuron, 86*(4), 1078–1090. doi:10.1016/j.neuron.2015.04.023

Gopinath, B., Russell, J., Kifley, A., Flood, V. M., & Mitchell, P. (2015). Adherence to dietary guidelines and successful aging over 10 years. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 71*(3), 349–355. doi:10.1093/gerona/glv189

Gupta, R., Duff, M. C., Denburg, N. L., Cohen, N. J., Bechara, A., & Tranel, D. (2009). Declarative memory is critical for sustained advantageous complex decision-making. *Neuropsychologia, 47*(7), 1686–1693. doi:10.1016/j.neuropsychologia.2009.02.007

Herlitz, A., Nilsson, L.-G., & Bäckman, L. (1997). Gender differences in episodic memory. *Memory & Cognition, 25*(6), 801–811. doi:10.3758/BF03211324

Herlitz, A., & Rehnman, J. (2008). Sex differences in episodic memory. *Current Directions in Psychological Science, 17*(1), 52–56. doi:10.1111/j.1467-8721.2008.00547.x

Hershey, D. A., Austin, J. T., & Gutierrez, H. C. (2015). Financial decision making across the adult life span. In: *Aging and decision making* (pp. 329–349). Elsevier. doi:10.1016/B978-0-12-417148-0.00016-9

Hu, X., Uhle, F., Fliessbach, K., Wagner, M., Han, Y., Weber, B., & Jessen, F. (2017). Reduced future-oriented decision making in individuals with subjective cognitive decline: A functional MRI study. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 6*, 222–231. doi:10.1016/j.dadm.2017.02.005

Jankovic, N., Geelen, A., Streppel, M. T., De Groot, L. C. P. G. M., Orfano, P., Van Den Hooven, E. H., & Fesken, E. J. (2014). Adherence to a healthy diet according to the world health organization guidelines and all-cause mortality in elderly adults from Europe and the United States. *American Journal of Epidemiology, 180*(10), 978–988. doi:10.1093/aje/kwu229

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 461–474. doi:10.1037/0278-7393.33.3.461

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298. doi:10.1038/nn.2635
Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3(June). doi:10.3389/fpsyg.2012.00193

Kramer, J. H., Yaffe, K., Lengenfelder, J., & Delis, D. C. (2003). Age and gender interactions on verbal memory performance. *Journal of the International Neuropsychological Society*, 9(1), 97–102. doi:10.1017/S1355617703910113

Kühner, C., Bürger, C., Keller, F., & Hautzinger, M. (2007). Reliabilität und Validität des revidierten Beck-Depressionsinventars (BDI-II). *Der Nervenarzt*, 78, 651–656. doi:10.1007/s00115-006-2098-7

Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, 36(2), 173–187. doi:10.1086/597160

Li, Y., Gao, J., Enkavi, A. Z., Zaval, L., Weber, E. U., & Johnson, E. J. (2015). Sound credit scores and financial decisions despite cognitive aging. *Proceedings of the National Academy of Sciences*, 112(1), 65–69. doi:10.1073/pnas.1413570112

Lighthall, N. R., Huettel, S. A., & Cabeza, R. (2014). Functional compensation in the ventromedial prefrontal cortex improves memory-dependent decisions in older adults. *Journal of Neuroscience*, 34(47), 15648–15657. doi:10.1523/JNEUROSCI.2888-14.2014

Lim, Y. Y., Ellis, K. A., Harrington, K. D., Ames, D., Martins, R. N., Masters, C. L., Maruff, P. (2012). Use of the cogstate brief battery in the assessment of Alzheimer’s disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. *Journal of Clinical and Experimental Neuropsychology*, 34(4), 345–358. doi:10.1080/13803395.2011.643227

Lim, Y. Y., Pietrzak, R. H., Snyder, P. J., Darby, D., & Maruff, P. (2012). Preliminary data on the effect of culture on the assessment of Alzheimer’s disease-related verbal memory impairment with the international shopping list test. *Archives of Clinical Neuropsychology*, 27(2), 136–147. doi:10.1093/arclin/arcl02

Lim, Y. Y., Prang, K. H., Cysique, L. A. J., Pietrzak, R. H., Snyder, P. J., & Maruff, P. (2009). A method for cross-cultural adaptation of a verbal memory assessment. *Behavior Research Methods*, 41(4), 1190–1200. doi:10.3758/BRM.41.4.1190

Maruff, P., Lim, Y. Y., Darby, D. G., Ellis, K. A., Pietrzak, R. H., Snyder, P. J., ... Masters, C. L. (2013). Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer’s disease. *BMC Psychology*, 1(1), 30. doi:10.1186/2050-7283-1-30

Milosavljevic, M., Malmaud, J., & Huth, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *SSRN Electronic Journal*, 5(6), 437–449. doi:10.2139/ssrn.1901533

Morrow, D., & Chin, J. (2015). Decision making and health literacy among older adults. In: *Aging and decision making* (pp. 261–282). Elsevier. doi:10.1016/B978-0-12-417148-0.00013-3

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... Chertkow, H. (2005). The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, 16(5), 292–305. doi:10.1016/j.tics.2012.04.005

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1), 190–206. doi:10.1016/j.actpsy.2013.06.003

Oud, B., Krajbich, I., Miller, K., Cheong, J. H., Botvinick, M., & Fehr, E. (2016). Irrational time allocation in decision-making. *Proceedings of the Royal Society B: Biological Sciences*, 283(1822), 20151439. doi:10.1098/rspb.2015.1439

Palombo, D. J., Keane, M. M., & Verfaellie, M. (2015). How does the hippocampus shape decisions? *Neurobiology of Learning and Memory*, 125, 93–97. doi:10.1016/j.nlm.2015.08.005

Philiastides, M. G., & Ratcliff, R. (2013). Influence of branding on preference-based decision making. *Psychological Science*, 24(7), 1208–1215. doi:10.1177/0956797612470701

Polanía, R., Moïsa, M., Opitz, A., Grüenschow, M., & Ruff, C. C. (2015). The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nature Communications*, 6, 8090. doi:10.1038/ncomms9090
Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545–556. doi:10.1038/nrn2357

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General, 140*(3), 464–487. doi:10.1037/a0023810

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine, 44*(10), 2029–2040. doi:10.1017/S0033291713002535

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the montreal cognitive assessment (MoCA) in a population-based sample. *Neurology, 77*(13), 1272–1275. doi:10.1212/WNL.0b013e318230208a

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron, 90*(5), 927–939. doi:10.1016/j.neuron.2016.04.036

Strough, J., Löckenhoff, C. E., & Hess, T. M. (2015). The present, past, and future of research on aging and decision making. In: *Aging and decision making* (pp. 1–14). Elsevier. doi:10.1016/B978-0-12-417148-0.00001-7

Tiffin-Richards, F. E., Costa, A. S., Holschbach, B., Frank, R. D., Vassiliadou, A., Krüger, T., … Reetz, K. (2014). The montreal cognitive assessment (MoCA) - A sensitive screening instrument for detecting cognitive impairment in chronic hemodialysis patients. *PLoS ONE, 9*(10), e106700. doi:10.1371/journal.pone.0106700

United Nations, “Department of Economic and Social Affairs, Population Division (2015).” *World Population Ageing (ST/ESA/SER.A/390).* Retrieved from www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice. *Psychological Science, 18*(6), 516–523. doi:10.1111/j.1467-9280.2007.01932.x

Wimmer, G. E., Braun, E. K., Daw, N. D., & Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. *Journal of Neuroscience, 34*(45), 14901–14912. doi:10.1523/JNEUROSCI.0204-14.2014

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science, 338*(6104), 270–273. doi:10.1126/science.1223252
### Appendix

#### Table A1. Distributions of all collected measures.

| Measure                              | n  | M     | SD   | Range          |
|--------------------------------------|----|-------|------|----------------|
| Age                                  | 30 | 74.63 | 6.11 | 65–88          |
| Delayed ISLT score                   | 30 | 7.7   | 2.17 | 2–11           |
| Detection task (log10 ms)            | 29 | 2.54  | 0.08 | 2.4–2.7        |
| TMT-A (s)                            | 30 | 54.7  | 23.72| 22–140         |
| TMT-B (s)                            | 30 | 132.8 | 87.43| 51–440         |
| TMT ratio                            | 30 | 2.51  | 1.5  | 1.33–8.8       |
| TMT difference                       | 30 | 78.1  | 79   | 19–390         |
| BDI-II score                         | 30 | 6.4   | 3.79 | 0–13           |
| MoCA score                           | 30 | 24.67 | 3.11 | 14–29          |
| Percentage of intransitive triplets  | 30 | 2.35  | 2.13 | 0.35–10.31     |
| Percentage of errors                 | 29 | 17.72 | 8.16 | 1.86–36.66     |
| Mean reaction time (ms)              | 30 | 1825.02 | 645.63 | 1113.08–4028.82 |
| Percentage of choices changed in retest | 30 | 14.11 | 6.2  | 4.74–33.68     |

Note. ISLT = International Shopping List Test; TMT = Trail Making Test; BDI-II = Beck Depression Inventory = II; MoCA = Montreal Cognitive Assessment.
Table A2. Correlations of all measures.

| Measure                  | Age | Delayed ISLT | Detection task | TMT-A | TMT-B | TMT ratio | TMT difference | BDI-II | MoCA | % intransitive triplets | % of errors | Mean reaction time |
|--------------------------|-----|--------------|----------------|-------|-------|-----------|----------------|--------|------|------------------------|-------------|-------------------|
| Delayed ISLT             | -.52** | -            |                |       |       |           |                |        |      |                        |             |                   |
| Detection task           | .22 | -.11         |                |       |       |           |                |        |      |                        |             |                   |
| TMT-A                    | .53** | -.31         | .29            |       |       |           |                |        |      |                        |             |                   |
| TMT-B                    | .52** | -.46*        | .31            | .47** | -     |           |                |        |      |                        |             |                   |
| TMT ratio                | .2  | -.3          | .14            | -.13  | .81*** | -         |                |        |      |                        |             |                   |
| TMT difference           | .42* | -.42*        | .25            | .22   | .96*** | .93***     | -              |        |      |                        |             |                   |
| BDI-II                   | .51** | -.3          | .14            | .18   | .25   | .14       | .22            | -      |      |                        |             |                   |
| MoCA                     | -.35 | .33          | -.08           | -.22  | -.66*** | -.58***    | -.66***        | -.15   |      |                        |             |                   |
| % intransitive triplets  | .38* | -.02         | .36            | .13   | .26   | .17       | .25            | .24    | .03 |                        |             |                   |
| % errors                 | .13  | -.25         | .1             | -.02  | -.14  | -.14      | -.15           | .16    | .35 | .26                      |             |                   |
| Mean reaction time       | .26  | -.48**       | .57**          | .36   | .1    | -.12      | .01            | .22    | .15 | .27                      | .52**       |                   |
| % choices changed in retest | .39* | -.07         | .32            | .18   | .31   | .2        | .29            | .29    | -.03 | .95***                  | .24         | .23               |

Note. ISLT = International Shopping List Test; TMT = Trail Making Test; BDI-II = Beck Depression Inventory = II; MoCA = Montreal Cognitive Assessment. Missing values for the percentage of errors and the Detection task measures were excluded. * p < .05, ** p < .01, *** p < .001.
Figure A1. Empirical probabilities of errors occurring on various levels of rating differences. Circles represent empirical probabilities for individual participants, solid data points with error bars represent means. Error bars represent 95% confidence intervals. For rating differences of 9 and 10 means of percentages of errors were omitted due to low numbers of observations.

Figure A2. Frequencies of MoCA scores.
Choice analysis

Table A3. Linear regression analysis predicting percentages of errors from delayed ISLT and including participant with partially missing rating data.

| Estimate          | Percentage of errors |
|-------------------|----------------------|
| Delayed ISLT      | −1.27 (−1.83)        |
| MoCA score        | 1.25* (2.59)         |
| Constant          | −2.96 (−0.26)        |

Observations: 30

Note. Unstandardized estimates are presented with $t$-statistics in parentheses. ISLT = International Shopping List Test; MoCA = Montreal Cognitive Assessment. * $p < .05$.

Analyses predicting changes of choice in retest

Table A4. Linear regression analyses predicting percentages of choices changed in retest from delayed ISLT and logistic random intercept regression analysis predicting change of choice in retest from delayed ISLT, rating difference and their respective Interaction.

| Estimate          | Percentage of choices changed in retest | Choice reversals between test and retest$^*$ |
|-------------------|-----------------------------------------|-----------------------------------------------|
| Delayed ISLT      | −0.19 (−0.32)                           | 0.03 (0.63)                                   |
| MoCA score        | −0.01 (−0.02)                           | −0.02 (−0.5)                                  |
| Rating difference | −0.27*** (−10.5)                        | 0.03** (2.68)                                 |
| Rating difference $\times$ delayed ISLT |                         |                                               |
| Constant          | 15.73 (1.64)                            | −1.75 (−1.92)                                 |

Observations: 30

Note. Unstandardized estimates are presented with $t$-statistics in parentheses for the model predicting percentage of choices changed in retest and with $z$ statistics in parentheses for the model predicting choice reversals between test and retest. ISLT = International Shopping List Test; MoCA = Montreal Cognitive Assessment. *Rating differences equal to zero were excluded. Rating differences, delayed ISLT and MoCA scores were mean centered. ** $p < .01$, *** $p < .001$. 

Figure A3. Frequencies of delayed ISLT scores.