Troubled past: A critical psychometric assessment of the self-report Survey of Autobiographical Memory (SAM)

Roni Setton[1] · Amber W. Lockrow[1] · Gary R. Turner[2] · R. Nathan Spreng[1,3,4,5]

Abstract

The Survey of Autobiographical Memory (SAM) was designed as an easy-to-administer measure of self-perceived autobiographical memory (AM) recollection capacity. We provide a comprehensive psychometric evaluation of the SAM in younger and older adults. First, we evaluated the reliability of the SAM as a measure of self-perceived recollective capacity. Next, we tested whether the SAM was a valid measure of episodic and autobiographical memory performance, as assessed with widely used performance-based measures. Finally, we investigated associations between the SAM, cognitive measures and self-reported assessments of psychological functioning. The SAM demonstrated reliability as a self-report measure of perceived recollective capacity. High internal consistency was observed across subscales, with the exception of SAM-semantic. Evidence for independence among the subscales was mixed: SAM-episodic and SAM-semantic items showed poor correspondence with respective subscales. Good correspondence was observed between the future and spatial items and their SAM subscales. The SAM showed limited associations with AM performance as measured by the Autobiographical Interview (AI), yet was broadly associated with self-reported AI event vividness. SAM scores were weakly associated with performance-based memory measures and were age-invariant, inconsistent with known age effects on declarative memory. Converging evidence indicated that SAM-episodic and SAM-semantic subscales are not independent and should not be interpreted as specific measures of episodic or semantic memory. The SAM was robustly associated with self-efficacy, suggesting an association with confidence in domain general self-report abilities. We urge caution in the use and interpretation of the SAM as a measure of AM, pending revision and further psychometric validation.

Keywords Self-report · Autobiographical memory · Aging · Individual differences

Introduction

Autobiographical memory (AM), our repository of personal past experiences, is integral to everyday human function. There are competing theories about the organization of AM (Conway & Pleydell-Pearce, 2000, Conway et al., 2016; Hassabis & Maguire, 2007; Rubin & Umanath, 2015), but there is general agreement that AM encompasses many of the same dimensions of recollective experience that comprise declarative memory, namely episodic, semantic, and spatial memory, as well as future-oriented thinking. Briefly, episodic memory refers to the "what, where, and when" of past experience (Tulving, 1972; Tulving, 2002; Hassabis & Maguire, 2007); semantic memory consists of generalized knowledge about the world and oneself (Tulving, 1972; Renoult et al., 2012); spatial memory includes the ability to retain cognitive maps, navigate through remembered scenes, and recall relationships among objects, all of which are predicated on memory for one's spatial orientation in a given environment (Barnes, 1988; Healy & Joset-Alves, 2009; Jeffery, 2016); and future thinking, or prospection, involves using imagination to mentally project oneself into the future (see Schacter et al., 2017 for a review). While each domain has unique features, their interdependence is evidenced by shared neural resources within the default network (Addis et al., 2007;

✉ R. Nathan Spreng
nathan.spreng@gmail.com; nathan.spreng@mcgill.ca

[1] Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, 3801 University St., Montreal, QC H3A 2B4, Canada

[2] Department of Psychology, York University, Toronto, ON, Canada

[3] McConnell Brain Imaging Centre, McGill University, Montreal, QC, Canada

[4] Departments of Psychiatry and Psychology, McGill University, Montreal, QC, Canada

[5] Douglas Mental Health University Institute, Montreal, QC, Canada

Ⓐ Springer

Buckner & Carroll, 2007; Spreng et al., 2009). Together, these functions subserve AM and lay the foundation for how we think about ourselves, bond with others, and solve problems in day-to-day life (e.g., Bluck & Alea, 2002). Self-perceptions of episodic memory, often in the form of subjective memory complaints, have been linked to a higher risk of developing dementia (e.g., Mitchell et al., 2014). Self-perceptions of AM abilities may offer additional unique insight into the experience of remembering as a marker of cognitive health.

The survey of autobiographical memory (SAM) is an easy-to-administer self-report measure of AM abilities (Palombo et al., 2013). Participants are instructed to think about their memory abilities in general and provide appraisals for abilities spanning specific events, factual knowledge, routes and landmarks, and imagined future events. The original questionnaire consisted of 102 questions based on established instruments on the phenomenology of AM (e.g., the Memory Experiences Questionnaire, Sutin & Robins, 2007) and literature on naturalistic memory. Palombo and colleagues created the 26-item SAM by using multivariate methods to reduce survey responses from 598 healthy adults (ages 18–65) to four dimensions, retaining reliable items to form SAM-episodic, SAM-semantic, SAM-spatial, and SAM-future subscales. These subscales recapitulated well-established divisions of recollective experience within AM (Tulving, 1972; Tulving, 2002; Hassabis & Maguire, 2007). An individual differences approach was then used to assess discriminant ($N = 598$) and criterion validity (Ns = 89 and 52 younger adults) of the SAM and its subscales. SAM-episodic scores were lower for individuals who reported a history of depression compared to those who did not, in line with findings of impoverished episodic memory related to major depressive disorder (e.g., Williams et al., 2007; Söderlund et al., 2014). Men also had higher scores than women for SAM-spatial, replicating a prevalent finding in the spatial memory literature (see Levine et al., 2016 for a review). Relationships between the SAM and both episodic memory and AM were used to demonstrate criterion validity: SAM-episodic was positively related to scene recollection, but not familiarity, on a Remember-Know task ($N = 98$; Yonelinas, 2001), and SAM-spatial was positively related to AM details conveying place information on the Autobiographical Interview ($N = 52$; Levine et al., 2002). These findings were interpreted as early evidence that the SAM could be a valid trait-like measure of autobiographical recollection tendencies.

Picco et al. (2020) undertook a psychometric evaluation of the SAM to test its factor structure and explore relationships among individual items. They first replicated the multiple correspondence analysis on the 26-item SAM ($N = 1879$), identifying four dimensions akin to those found in Palombo et al. (2013). The authors then tested a number of different confirmatory factor analysis models ($N = 1017$) and determined that models containing four factors (for each SAM subscale) or

four factors plus an additional general factor (SAM-total) provided equally adequate fits. Network analyses performed on the larger sample ($N = 2896$) revealed that questions on the SAM related most to other questions within a subscale, a finding which was stable across genders and groups with and without a history of anxiety. While this study provided further evaluation of the reliability and underlying structure of the SAM, validity was not assessed.

Validation of the SAM as a self-report measure of AM has primarily centered on its relationship to imagery and vividness. Five people who demonstrated exceptional ability to accurately remember events from the past without an overt mnemonic strategy (LePort et al., 2012) reported higher scores on SAM-episodic compared to healthy controls (Sheldon et al., 2016; Palombo et al., 2018). Performance on face–name association and visual memory tests were also elevated, while performance on classic laboratory tests of episodic memory was comparable to controls (LePort et al., 2012). These results provided evidence for an association between SAM-episodic scores and AM ability. Conversely, three individuals who reported an impoverished sense of remembering (Palombo et al., 2015) had lower SAM-episodic scores despite intact objective autobiographical and episodic memory (Sheldon et al., 2016; Palombo et al., 2018). Relatedly, SAM-episodic has shown correspondence to self-reported vividness (rated on a scale from 1 to 5) of specific autobiographical events following recollection, but not to recollection itself (Sheldon et al., 2016). These findings lend credence to SAM-episodic as a measure sensitive to the self-reported capacity to consciously recollect episodic details of past events. However, they do not address the SAM's ability to measure episodic, semantic, spatial, and future aspects of memory performance.

Surprisingly little evidence for a clear link between the SAM and AM has been reported. This was highlighted in a recent evaluation of subjective versus objective performance measures. In a sample of 217 adults, four times as large as the original validation study, no relationship was observed between the SAM and AM (Clark & Maguire, 2020). Armson et al. (2020) reported an association between SAM-episodic and episodic event details recalled by participants one week after a staged audio tour. Further, participants who reported greater episodic capacity for these unrehearsed events on the SAM made more gaze fixations, indicative of greater visual exploration, when recalling episodic details of the audio tour. These data suggested that individuals with higher SAM-episodic may rely more on visual imagery during episodic recall and re-experiencing of the recent past. However, similar associations were found with SAM-semantic, suggesting that the SAM-episodic scale, even when limited to unrehearsed event recall, is non-specific as a measure of vivid episodic re-experiencing.

Developed as a trait measure of AM, invariant to age, the SAM is now being administered to large samples of older adults, though most validation studies have involved healthy young adults. We are aware of only one SAM study involving older adults (Fan et al., 2020a). This report examined whether trait AM, as assessed by the SAM, moderated age-related decreases in cognition. As expected, older age was associated with worse performance on laboratory tests of episodic memory. No age relationships were observed for self-reported cognitive functioning nor the SAM. However, better self-reported cognitive function was predicted by SAM-episodic as well as the interaction between age and SAM-episodic: older adults with higher than average SAM-episodic reported more frequent cognitive complaints with increasing age, whereas those with lower than average SAM-episodic reported fewer cognitive complaints with age. This finding was taken as evidence that older adults who endorse less episodic AM have more practice with compensatory memory strategies and fare better amid age-related episodic decline. SAM-episodic did not play a similar moderating role on age and objective episodic memory.

To date, evidence suggests that the SAM may serve as a trait measure of self-reported recollective experience rather than a valid measure of episodic or autobiographical memory per se. To our knowledge, a formal study that tests these predictions in younger and older adults has not been reported. To address this gap, here we present a comprehensive validation study of the psychometric properties of the SAM. We administered the SAM to samples of healthy younger and older adults, analyzing them separately and together where appropriate. If the SAM is a psychometrically reliable and valid measure, we predicted that it would demonstrate (i) internal consistency, (ii) convergent and divergent validity, and (iii) stability across different age groups. Further, we predicted that as a valid measure of autobiographical recollective capacity, SAM performance would (iv) recapitulate known age reductions in episodic AM and increases in semantic AM (e.g., Levine et al., 2002; Spreng et al., 2018).

SAM reliability was assessed using several internal consistency metrics. Specifically, we calculated Cronbach's alpha, item total correlations, and factor analysis scores. Correlations among subscales and gender differences between them were carried out and compared to the initial validation. SAM validity was examined through associations between SAM scores and measures of memory and fluid cognition, AM (including self-reported vividness), and non-memory-related self-report measures including personality and social cognition. Performance-based measures of memory (episodic, semantic, AM) were used to examine convergent validity. Performance-based measures of fluid cognition were used to examine divergent validity. Self-report measures of personality and social cognition were used to determine the extent to which associations were attributable to method variance alone

(Clark and Maguire, 2020). The stability of the SAM was evaluated by comparing internal consistency and validity results for each age cohort. We then pooled our samples to compare age effects on episodic and semantic memory measures and SAM scores. We also used pooled data, controlling for age and gender, to re-examine the validity of the SAM as an individual differences measure in the full study sample. Finally, we repeated all analyses with simple average scores for all SAM subscales to evaluate the utility of the SAM's recommended proprietary scoring procedure. The overarching goal of the study was to conduct a comprehensive psychometric assessment of the SAM (i) as a self-report survey of perceived recollective capacity and (ii) as a valid measure of autobiographical memory ability.

Materials and Methods

Participants

A total of 209 young and 193 older healthy adults from Ithaca, New York and Toronto, Canada completed the SAM. Participants were screened during an eligibility interview for having a history of neurological or other medical illness known to impact cognition, acute or chronic psychiatric illness, ongoing or recent treatment with psychotropic medication, and significant changes to health status within 3 months. As part of the below procedures, participants completed the Mini-Mental State Examination (MMSE; Folstein et al., 1975) to briefly estimate cognitive status. Nine participants (3 younger, 6 older) were excluded for having MMSE scores below 27/30 along with scores of fluid processing (NIH Toolbox; Gershon et al., 2013) below an age-adjusted national percentile of 25%. Beck's Depression Inventory (BDI; Beck et al., 1996) or the Geriatric Depression Scale (GDS; Yesavage & Brink, 1983) served as a nonclinical metric for history of depressive symptoms. Three participants (1 young, 2 older) were excluded for scoring in the range of "severe depression." One older participant was excluded for presenting with signs of confabulation during testing procedures, and two younger participants were excluded for incorrectly answering attention checks throughout the web-based portion of the procedure. The final sample consisted of 203 younger (18–34 years, 119 female, $M = 22.27$, $SD = 3.32$) and 184 older adults (58–92 years, 98 female, $M = 69.41$, $SD = 6.90$; see Table 1). Participants received $10/hour or course credit for their time.

We assume that the associations of interest, namely between the SAM and other cognitive measures, are of the same magnitude as those typically observed in psychology research: an r of approximately .25, or a d of 0.52 (Fraley & Marks, 2007; Hemphill, 2003; Meyer et al., 2001). To detect

Table 1 Table 1

Sample Demographics

	Younger Adults	Older Adults
N	203 (119 female)	184 (98 female)
Age		
Range	18-34	58-92
M	22.27	69.41
SD	3.32	6.9
Education		
Range	12-23	12-26
M	14.99	17.24
SD	1.87	2.89
NIH Fluid Cognition		
Range	88.74-153.37	75.04-122.79
M	119.61	93.93
SD	12.85	7.08
NIH Crystallized Cognition		
Range	99.07-153.95	107.28-153.95
M	125.09	135.86
SD	9.31	10.88
Episodic Memory (performance-based)		
Range	-1.07-1.53	-2.13-1.07
M	0.58	-.64
SD	0.43	0.67
Semantic Memory (performance-based)		
Range	-2.62-1.35	-3.22-1.85
M	-.38	0.42
SD	0.72	0.78
Fluid Cognition (performance-based)		
Range	-.83-1.56	-2.21-.66
M	0.51	-.57
SD	0.47	0.433

approximately 70% of all effect sizes based on known estimates (Fraley & Marks, 2007), a sample of at least 150 participants is recommended for examining individual differences. Our sample thus provides 95% confidence in detecting a correlation magnitude of 0.15 (Mar et al., 2013).

Procedure

Participants completed the SAM as part of a large battery of behavioral measures over several testing sessions both in lab and online. Each session lasted between one and three hours. Assessments were administered in person during one of the sessions or online via Qualtrics. Attention checks (e.g., "To what extent (1–5) do you agree with this statement: I always walk on my hands when ascending the stairs.") were embedded within web-based measures to ensure compliance.

Survey of autobiographical memory

The SAM is a self-report instrument used to measure self-perceived AM abilities (Palombo et al., 2013). Participants completed the 26-item version, rating their general memory abilities on a five-point Likert scale from strongly disagree to strongly agree. SAM-episodic, SAM-semantic, SAM-spatial, SAM-future, and SAM-total scores were calculated according to the original protocol (courtesy of Brian Levine) to capture the multidimensional facets of subjective re-experiencing. Raw item-level data, reverse-coded where appropriate, were retained for further analysis.

Performance-based measures of memory and fluid cognition

Participants also completed performance-based measures of memory and fluid cognition. Thirty-three participants (17 younger, 16 older) were excluded due to technical difficulties or greater than 50% missing data, leaving a subset of 354 (186 younger, 168 older) participants. Episodic memory was measured with the Auditory Verbal Learning Test (Rey) as implemented in the NIH Cognition Toolbox (Gershon et al., 2013), the Associative Recall Paradigm (Brainerd et al., 2014), and Verbal Paired Associates from the Weschler Memory Scale-IV (Weschler, 2009). Semantic memory was measured with Oral Reading Recognition and Picture Vocabulary Tests from the NIH Cognition Toolbox (Gershon et al., 2013), and Shipley-2 Vocabulary (Shipley et al., 2009). Fluid processing was measured with scores from Picture Sequence Memory, List Sorting Working Memory, Dimensional Change Card Sort, Flanker Inhibitory Control and Attention, and Pattern Comparison Processing Speed Tests as part of the NIH Cognition Toolbox (Gershon et al., 2013), Shipley-2 Blocks (Shipley et al., 2009), Trail-Making Test (Reitan, 1958), Symbol Digit Modality Test (Smith, 1982), and Reading Span Task (Daneman & Carpenter, 1980). Composite scores of episodic memory, semantic memory, and fluid cognition were derived from scores on these tasks (see Section 2.4.1 and Table S1).

A Remember-Know Source recognition memory paradigm (Tulving, 1985; Gardiner et al., 1997) was also completed by 258 of these participants. Forty-four participants were excluded due to insufficient variation in response, leaving data from 214 participants (127 younger, 87 older). Remember-Know was not included as part of the episodic memory composite in order to separately explore its association with SAM-episodic, as detailed in Section 2.4.1 below.

Autobiographical memory

A subset of participants (145 younger, 112 older) completed the Autobiographical Interview (AI; Levine et al., 2002). The AI is a naturalistic measure of AM where participants recall

one specific memory from each of the following life stages: childhood, teenage years, early adulthood, middle adulthood, and late adulthood. Younger adults provide a memory from the first three stages, whereas older adults provide a memory from all five. Participants recalled each memory in as much detail as possible (free recall) and were then lightly prompted to provide additional details (general probe). After all memory recollections, participants were then questioned about each memory to cue further episodic remembering (specific probe). Participants also rated the vividness ("How clearly can you visualize this event?") and rehearsal ("How often do you think or talk about this memory?") of each memory on a scale from 1 to 6 (vividness: 1 = not at all, 6 = extremely; rehearsal: 1 = once every few years, 6 = once per week). As per the original protocol, memories were segmented and scored for episodic-like (internal) and non-episodic (external) details. Internal details comprised information about the event unfolding, location, time, sensory descriptions, emotions, and thoughts related to the event. External details included unrelated events, general information about the world or oneself, repetitions, and other nonscorable utterances such as metacognitive statements. All interviews were double-scored and showed high inter-rater reliability (interclass correlation of internal and external details, respectively: $r(255) = .87$, $p < .000$; $r(255) = .88$, $p < .000$).

Personality and social cognition

A subset of participants (186 younger, 168 older) also completed measures of personality and social cognition including the Big Five Aspects Scale (DeYoung et al., 2007), NIH Emotion Toolbox (Gershon et al., 2013), Interpersonal Reactivity Index (Davis, 1983), Toronto Empathy Questionnaire (Spreng et al., 2009), and Reading the Mind in the Eyes (Baron-Cohen et al., 1997). Complete BDI and GDS data were available for 115 younger participants and 166 older participants.

Reliability of the SAM

We determined the reliability of the SAM by inspecting internal consistency, factor structure, and gender differences separately for each age group.

Internal consistency of the SAM was examined to determine consistency among subscales and constituent items. Item-level data were used to compute Cronbach's α, item-deleted Cronbach's α, and item total correlations for each subscale and the measure as a whole. Cronbach's α indicates the strength of item relations. Item-deleted Cronbach's α provides a value for Cronbach's α minus one item to assess how individual items influence the internal consistency of a scale. Item total correlation provides a value for the relationship between an individual item and the average of all other items

in the scale. Together these tests identify items where performance deviates from other items, thus providing a measure of internal consistency.

Next, we performed a series of factor analyses on the item-level data to reproduce the underlying structure of the SAM. An exploratory factor analysis model was initially estimated with 25 factors to visualize a scree plot of eigenvalues and determine that a four-factor solution was appropriate (see section 3.1. and Fig. 1). A model was then estimated with four factors using maximum likelihood extraction and varimax (orthogonal) rotation (Sakaluk & Short, 2017). A series of confirmatory factor analysis models were fit in line with Picco et al. (2020): one-factor, three-factor, four-factor, and a high-order model comprising one general factor and four subfactors. Models fit within each age group were estimated with robust maximum likelihood estimation to account for ordinal variables and smaller sample sizes (Li, 2016). Additional models with participants from both age cohorts were fit with diagonally weighted least squares estimation. In all models, intercorrelation between factors was permitted unless otherwise noted.

Pearson correlations were conducted between all SAM scores (subscales and total) with 95% confidence intervals and an alpha of .05. To test whether previously reported gender differences for SAM-spatial were replicable (Palombo et al., 2013), a mixed ANOVA on SAM scores was modeled with gender as a between-subjects factor and SAM subscale as a within-subjects factor (Greenhouse-Geisser corrections for violations of sphericity reported where appropriate). Tukey's post hoc t tests were run for group comparisons.

We additionally tested for a relationship between years of education and SAM scores in older adults. Spearman's ρ correlations were calculated between each SAM score and years of education with 95% confidence intervals and an alpha of .05.
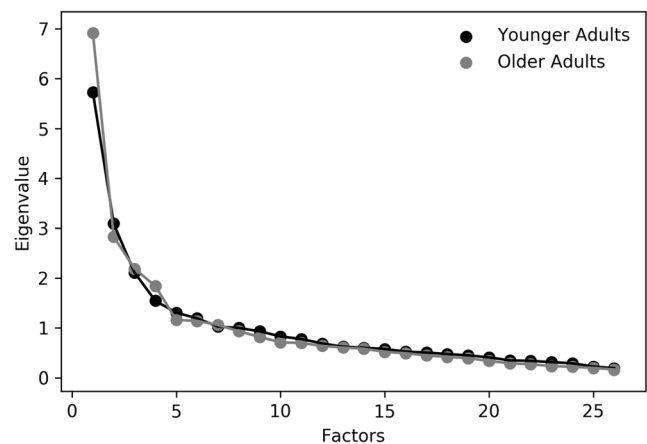
Fig. 1 Scree plots from exploratory factor analyses. Item-level responses from the SAM were entered into Exploratory Factor Analyses (EFA) in younger and older adult samples separately. EFA was performed with a principal factors solution using 25 factors for visualization purposes

Validity of the SAM

To observe whether the SAM related to other validated behavioral instruments, we performed correlations between SAM scores and performance-based measures of memory and fluid cognition, a measure of AM, and measures of personality and social cognition. Correlations were conducted in favor of other multivariate analyses to observe simple associations and facilitate interpretation. Correlations were carried out first within each age group, then across the entire sample.

Associations with memory and fluid cognition

Scores from 18 measures (Table S1) were Z-scored and averaged to create composite indices of episodic memory, semantic memory, and fluid cognition. Missing cells were imputed with age group means prior to Z-scoring. At least one cell was imputed for 51 participants, but doing so did not notably alter results (Table S3).

Using composite scores for different memory domains should improve the predictive validity of the SAM, which aims to capture AM abilities in general rather than for any specific task (Beaudoin & Desrichard, 2011). To create composite scores, data from younger and older participants were concatenated. Samples were then repartitioned by age cohort to explore associations between SAM scores and composite scores separately. Ninety-five percent confidence intervals were used to detect significant Spearman's ρ correlations at an alpha of .05 and a Bonferroni adjustment of $p < .017$ based on three tests for each SAM score.

We separately examined associations between the SAM and a Remember-Know task. According to dual-process accounts of recognition memory (e.g., Yonelinas, 2002; Yonelinas et al., 2010), Remember-Know paradigms can be used to isolate autonoetic consciousness, the sense of re-experiencing a specific event and the context in which it was encoded ("recollection"), from knowing, memory in the absence of contextual detail ("familiarity"; Wheeler et al., 1997). Support for SAM-episodic was previously found in its relationship to recollection, but not familiarity, in a Remember-Know task (Palombo et al., 2013). We therefore use 95% confidence intervals to detect significant Spearman's ρ correlations at an uncorrected alpha of .05 between SAM-episodic and three Remember-Know variables: recollection, familiarity (as in Stamenova et al., 2017), and source memory.

Associations with autobiographical memory

Associations between the AI and the SAM were explored to observe the extent to which self-reported AM corresponds with detailed autobiographical recollection and self-reported vividness. We examined relationships between the SAM and several dependent variables derived from scored interviews:

internal detail count, external detail count, total detail count, word count, internal details as a proportion of all details (internal proportion), and internal and external details as a proportion of word count (internal and external density scores). Variations on internal and external detail scores were used to explore whether verbosity had a disproportionate influence on any relationship with the SAM. Detail scores reflect an average across all AI memories from free recall and general probe (3 for young adults, 5 for older adults). Select subcategories of internal and external details were also associated with the SAM, as described below. Details from specific probe were not included, as the focus here was on spontaneous (free) recollection performance, uninfluenced by specific interviewer questions or the expectation to remember certain details. Associations including specific probe (as in Palombo et al., 2013) can be found in Supplemental Material (Table S4).

Given putative associations between SAM subscales and specific dimensions of declarative memory, we tested a priori hypotheses about the relationships between the SAM and AI detail scores. First, we expected internal detail variables to positively relate to SAM-episodic and SAM-future, since remembering the past and simulating the future draw on similar mechanisms (Addis et al., 2007). Although the SAM's factor structure might suggest that SAM-episodic and SAM-future subscales are independent of each other, they share a significant amount of variance and are highly correlated (Palombo et al., 2013; Picco et al., 2020; see Results and Fig. 2 below). Second, we reasoned that internal detail variables would relate to SAM-spatial since the only reported association between SAM and AM was between internal place details on the AI and the dimension from which SAM-spatial was derived (Palombo et al., 2013). We looked specifically at the relationship between the SAM and the subcategory of internal place details in case the association was specific to place details. We also inspected associations with internal perceptual details—which include information about spatial orientation—since SAM-spatial has largely been validated with performance on spatial navigation tasks (Clark & Maguire, 2020; Selarka et al., 2019). Finally, we hypothesized that external detail variables would positively relate to SAM-semantic since semantic details comprise the majority of external details on the AI (Levine et al., 2002; Bastin et al., 2013). We also examined associations with semantic details alone, since external details comprise additional non-episodic information that may mask existing relationships.

Relationships between self-reported vividness on the AI and all SAM subscales were also tested. SAM-episodic was predicted to show a positive relationship in light of previously reported associations (e.g., Sheldon et al., 2016) and its sensitivity to visual imagery in atypical cases of autobiographical remembering (Sheldon et al., 2016; Palombo et al., 2015; Palombo et al., 2018).
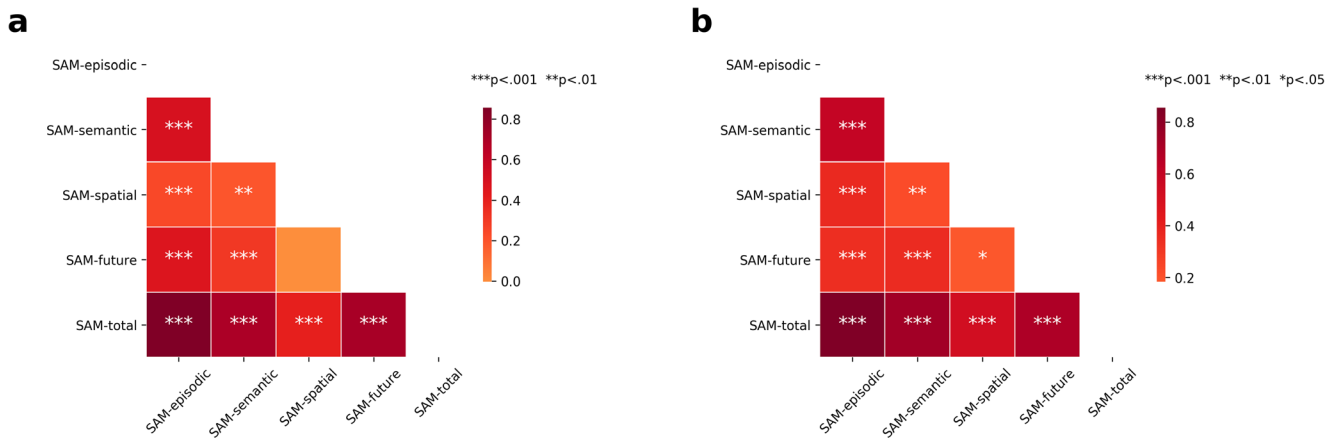
Fig. 2 Correlations between SAM scores. Lower triangles of covariation matrices containing SAM subscale and total scores are depicted for **a)** younger adults and **b)** older adults

For predicted relationships, we report Spearman's ρ correlations with 95% confidence intervals at an alpha of .05 (uncorrected). For all other associations, 95% confidence intervals were used to detect significant Spearman's ρ correlations at an alpha of .05 and a Bonferroni adjustment of $p < .006$ based on eight tests for each SAM score.

Associations with personality and social cognition

Associations between SAM scores and measures of personality and social cognition were explored using 95% confidence intervals to detect significant Spearman's ρ correlations at an alpha of .05. A Bonferroni adjustment of $p < .002$ was implemented based on 28 tests for each SAM score. For the smaller subset with BDI and GDS data, correlations were conducted with no correction.

Age group differences

We examined age group differences on the SAM and compared them to age differences in memory and fluid cognition. Data were pooled across samples ($N = 387$), and a mixed ANCOVA was modeled with age group as a between-subjects variable, SAM subscale as a within-subjects variable, and their interaction (Greenhouse-Geisser corrections for violations of sphericity reported where appropriate) on SAM scores. Gender was included as a covariate of no interest. Tukey's post hoc t tests were run for group comparisons. Associations to examine the validity of the SAM were then conducted on the full sample controlling for age (as a continuous variable) and gender.

Scoring

The SAM was originally validated with a multivariate weighting scoring procedure that remains unpublished and available only upon request from the authors. It is unclear

whether the scoring complexity conserves unique information about the SAM beyond that of a simpler procedure. We therefore evaluated whether the SAM performed similarly when scored in a manner more similar to other behavioral scales. In each sample, raw item-level responses (respecting reverse-scored items) were averaged within subscales and across the whole measure to create average scores (as in Fan et al., 2020b). To assess the correspondence between multivariate and average SAM scores, Pearson correlations were conducted between each multivariate SAM score and its corresponding average score with 95% confidence intervals and an alpha of .05 (Table 5). All associations inspecting the validity of the SAM were then repeated with the average SAM scores in lieu of the multivariate SAM scores. Averages, rather than sums, were calculated since the number of questions varied across SAM subscales, and would facilitate easier comparison of ANOVA results from analyses with different scoring procedures.

Software

Statistical analyses were conducted in python 3.6 (packages included factor_analyzer, pingouin, and scipy; Van Rossum & Drake, 2009) and R version 3.3.3 (packages included pwr, levaan, and lme4; R Core Team, 2013).

Results

Reliability of the SAM

Younger adults

We considered item-level performance to evaluate whether subscales contained items that measured the same construct. Values of Cronbach's α were first calculated to indicate how items related to one another with respect to their subscales and

the full measure. We next computed item-deleted Cronbach's α to observe whether removing a single item changed the Cronbach's α value. Item total correlations further specified how each item contributed to the reliability of its subscale and the scale as a whole.

Cronbach's α values for all SAM scores satisfied accepted values between .7 and .9 (Nunnally, 1978; George & Mallery, 2003; Hair, 2010) with the exception of SAM-semantic (Table 2). Two spatial items ("17. I have a hard time judging the distance (e.g., in meters or kilometers) between familiar landmarks." and "20. I use specific landmarks for navigating.") reduced Cronbach's α and had low item total correlations (Table 3). Comparable results were reported in Picco et al. (2020).

We next examined the factor structure of the SAM with an exploratory factor analysis. SAM data were considered suitable for structure detection with high values on Bartlett's test of sphericity ($\chi^2(325) = 1862.00$, $p < .001$), and the Kaiser-Meyer-Olkin measure of sampling adequacy was .80. Maximum likelihood extraction with an orthogonal (varimax) rotation using a four-factor solution was chosen to reproduce the four dimensions of the SAM. Maximum likelihood extraction ensures that factors only retain shared construct variance and are not contaminated by error variance (Sakaluk & Short, 2017). We used orthogonal rotation to remain consistent with the multiple correspondence analysis, which calculates orthogonal dimensions, used to create the SAM (Palombo et al., 2013).

The scree plot shown in Fig. 1 (black line) shows four factors with eigenvalues greater than 1, indicating that a four-factor solution was appropriate. The minimum adequate loading of items is considered .32 (Tabachnick & Fidell, 2007), but we raise this criterion to .4 given our re-examination of an existing scale and the high loadings of most items. Factor 1 contained high loadings of SAM-future items, explaining 12.10% of the variance. Factor 2 loaded six SAM-episodic and two SAM-semantic items, explaining 11.53% of the variance. One of these SAM-episodic items ("1. Specific events are difficult for me to recall.") cross-loaded more highly with Factor 4. Another SAM-episodic item ("3. When I remember events, in general I can recall objects that were in the environment.") did not load above .4 on any factor. Factor 3 had high loadings for most SAM-spatial items and explained 9.72% of the variance. Two SAM-spatial items, questions 17 and 20, observed to reduce Cronbach's α above, did not load

with any factor. Factor 4 loaded two SAM-semantic items and two SAM-episodic items, explaining 5.67% of the variance. Three out of the six SAM-semantic items ("9. I can learn and repeat facts easily, even if I don't remember where I learned them.", "11. After I have met someone once, I easily remember his or her name.", "12. I can easily remember the names of famous people (sports figures, politicians, celebrities).") did not load above .4 for any factor. Full factor loadings are listed in Table 3. Cumulative variance explained was 39.03%.

We also tested the current four-factor solution by fitting a series of confirmatory factor analysis models with robust maximum likelihood estimation assuming uncorrelated factors. Despite the mixed grouping of items across SAM-episodic and SAM-semantic subscales observed above, fit statistics indicated similarly moderate fits for both three- and four-factor models (see Table S2). A high-order model with SAM-total as a general factor and each of the SAM subscale as subfactors also showed a moderate fit.

Next, we turn to the psychometric properties of SAM subscale and total scores. Distributions for all SAM scores are tabulated in the left panel of Table 5. Pearson correlations between SAM subscale and total scores showed high correspondence (Fig. 2a). The only unrelated pair was SAM-future and SAM-spatial ($r(201) = -.03$, $p = .971$).

Gender differences across subscales were tested with a mixed ANOVA modeling gender, subscale, and their interaction on SAM scores. We observed a main effect of subscale ($F(4, 804) = 3.77$, $p < .05$ with Greenhouse-Geisser correction, $\eta^2_p = .02$) and an interaction between subscale and gender ($F(4, 804) = 2.41$, $p < .05$, $\eta^2_p = .01$). Post hoc pairwise t tests revealed that men had higher SAM-spatial scores ($M = 101.02$, $SD = 14.71$) than women ($M = 95.48$, $SD = 15.90$; Cohen's $d = .36$, $p < .05$; Fig. 3, left panel).

Item-level responses from the SAM were entered into exploratory factor analyses (EFA) in younger and older adult samples separately. EFA was performed with a principal factors solution using 25 factors for visualization purposes

Older adults

Above we assessed the reliability of the SAM in a younger cohort through internal consistency, factor structure, and gender differences. Here, we do the same in an older sample. We refer back to results from the younger group to determine the stability of the SAM across samples.

Table 2 Reliability of the SAM: Cronbach's α by age group

| Age Group	SAM-episodic	SAM-semantic	SAM-spatial	SAM-future	SAM-total
Younger Adults	0.794	0.568	0.768	0.851	0.840
Older Adults	0.809	0.615	0.723	0.897	0.880

Table 3 Reliability of the SAM: Cronbach's α item deleted, item total correlation, and exploratory factor analysis results in younger adults (N=203)

SAM Item	Subscales		Full Scale		EFA			
	Cronbach's α Item Deleted	Item Total Correlation	Cronbach's α Item Deleted	Item Total Correlation	Factor 1	Factor 2	Factor 3	Factor 4
Episodic								
1. Specific events are difficult for me recall. (R)	.77	.52	.83	.48	.07	**.46**	.01	**.60**
2. When I remember events, I have a hard time determining the order of details in the event. (R)	.78	.42	.83	.42	.13	.36	.02	**.45**
3. When I remember events, in general I can recall objects that were in the environment.	.78	.41	.83	.41	.19	.39	.06	.13
4. When I remember events, in general I can recall what I was wearing	.78	.46	.83	.41	.28	**.43**	.03	-.06
5. I am highly confident in my ability to remember past events.	.75	.66	.83	.63	.24	**.72**	.05	.26
6. When I remember events, I remember a lot of details.	.74	.69	.83	.58	.23	**.81**	-.07	.14
7. When I remember events, in general I can recall which day of the week it was.	.79	.37	.84	.38	.05	**.40**	.22	.03
8. When I remember events, in general I can recall people, what they looked like, or what they were wearing.	.77	.49	.83	.46	.13	**.50**	.13	.05
Semantic								
9. I can learn and repeat facts easily, even if I don't remember where I learned them.	.52	.31	.84	.19	-0.2	.04	.09	.38
10. After I have read a novel or newspaper, I forget the facts after a few days. (R)	.56	.22	.84	.24	.03	.11	.06	**.46**
11. After I have met someone once, I easily remember his or her name.	.51	.35	.84	.34	.12	.31	.05	.10
12. I can easily remember the names of famous people (sports figure, politicians, celebrities).	.52	.31	.84	.25	-.01	.25	.11	.08
13. I have a hard time remembering information I have learned at school or work. (R)	.53	.29	.84	.28	.14	.09	-.05	**.58**
14. I am very good at remembering information about people that I know (e.g., the names of a co-worker's children, their personalities, places friends have visited etc.).	.49	.37	.83	.48	.12	**.48**	.15	.17
Spatial								
15. In general, my ability to navigate is better than most of my family/friends.	.70	.62	.84	.32	.00	.16	**.74**	-.05
16. After I have visited an area, it is easy for me to find my way around the second time I visit.	.68	.70	.83	.43	.07	.23	**.80**	-.02
17. I have a hard time judging the distance (e.g., in meters or kilometers) between familiar landmarks. (R)	.79	.31	.85	.12	-.05	-.01	.37	.10
18. I get lost easily, even in familiar areas. (R)	.69	.71	.83	.41	-.03	.20	**.78**	.11
19. If my route to work or school was blocked, I could easily find the next faster way to get there.	.72	.57	.84	.29	-.03	.10	**.63**	.01
20. I use specific landmarks for navigating.	.80	.20	.84	.18	.20	.01	.22	.02
Future								
21. When I imagine an event in the future, the event generates vivid mental images that are specific in time and place.	.83	.64	.83	.44	**.69**	.23	-.05	.00
22. When I imagine an event in the future, I can picture the spatial layout.	.82	.69	.83	.46	**.78**	.08	.08	.08
23. When I image an event in the future, I can pictured people and what they look like.	.82	.68	.83	.43	**.72**	.18	-.08	.08
24. When I imagine an event in the futures, I can imagine how I may feel.	.85	.52	.84	.37	**.53**	.14	.08	.04
25. When I imagine an event in the future, I can picture images (e.g., people, objects, etc)	.82	.68	.83	.47	**.72**	.21	-.02	.05
26. I have a difficult time imagining specific events in the future. (R)	.83	.62	.83	.40	**.62**	.20	-.08	.13

Note. (R) indicates items that were reverse coded. Bold text denotes factor loading above 0.4.
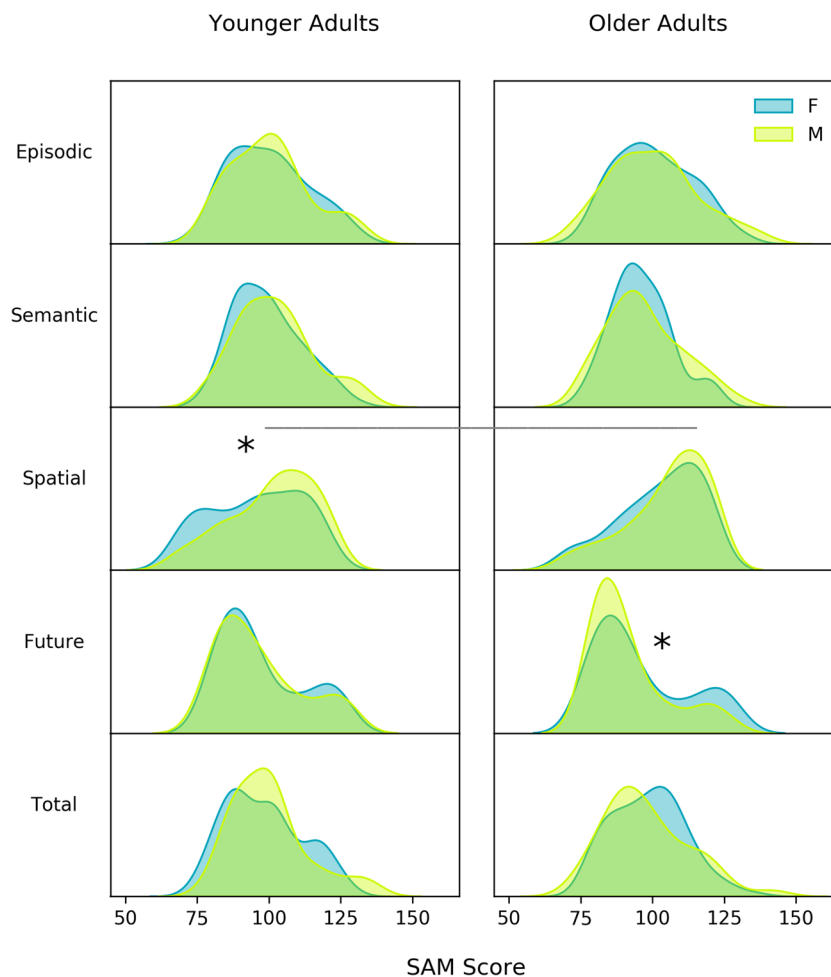
Fig. 3 Gender and age differences in the SAM. Distributions shown for SAM subscale scores in the younger (left panel) and older (right panel) adult cohorts by gender. Lines indicate age group differences and * denotes gender differences. Gray lines illustrate higher scores in older adults compared to younger adults when controlling for gender

Cronbach's α values in older adults were comparable to those in younger adults, with all α values falling within a range of .7 and .9 except SAM-semantic (Table 2). All values were marginally higher in the older cohort except SAM-spatial. Cronbach's α increased when two SAM-episodic items were excluded ("1. Specific events are difficult for me to recall." and "3. When I remember events, in general I can recall objects that were in the environment"). Notably, question 1 was the same item observed to cross-load higher with SAM-semantic items in younger adults. The same two SAM-spatial items raised α values when excluded ("17. I have a hard time judging the distance (e.g., in meters or kilometers) between familiar landmarks." and "20. I use specific landmarks for navigating."). These four items showed relatively low item total correlations with respect to their subscales and SAM-total, indicating a poor relationship with the remaining 22 items (Table 4).

The factorability of the SAM was tested again in this cohort to observe whether the underlying structure of the SAM changed with an older sample. Bartlett's test of sphericity was significant ($\chi^2(325) = 2108$, $p < .001$), and the Kaiser-Meyer-Olkin measure of sampling adequacy was .84. The gray line in Fig. 1 shows that four factors had eigenvalues greater than 1. Full factor loadings for a four-factor exploratory factor analysis model are shown in Table 4. Factor 1 loaded SAM-future items highly together, explaining 14.56% of the variance. Factor 2 loaded six SAM-episodic items and one SAM-semantic item to explain 13.56% of the variance. Factor 3 loaded four SAM-spatial items highly together to explain 9.62% of the variance. As in the younger sample, questions 17 and 20 did not load highly on any factor. Factor 4 loaded two SAM-episodic items and one SAM-semantic item to explain 7.48% of the variance. Four out of six SAM-semantic items ("9. I can learn and repeat facts easily, even if I don't remember where I learned them.", "10. After I have read a novel or newspaper, I forget the facts after a few days.", "11. After I have met someone once, I easily remember his or her name.", "12. I can easily remember the names of famous people (sports figures, politicians, celebrities).") did not load above .4 with any factor. This was similar

Table 4 Reliability of the SAM: Cronbach's α item deleted, item total correlation, and exploratory factor analysis results in older adults (N=184)

SAM Item	Subscales		Full Scale		EFA			
	Cronbach's α Item Deleted	Item total Correction	Cronbach's α Item Deleted	Item total Correction	Factor 1	Factor 2	Factor 3	Factor 4
Episodic								
1. Specific events are difficult for me to recall. (R)	.82	.33	.88	.40	.06	.15	.09	**.68**
2. When I remember events, I have a hard time determining the order of details in the events. (R)	.80	.41	.88	.45	.11	.21	.16	**.57**
3. When I remember events, in general I can recall objects that were in the environment.	.81	.39	.88	.38	.19	**.42**	.16	-.02
4. When I remember events, in general I can recall what I was wearing.	.78	.56	.88	.48	.17	**.63**	.08	-.04
5. I am highly confident in my ability to remember past events.	.77	.67	.87	.60	.00	**.70**	.19	.36
6. When I remember events, I remember a lot of details.	.76	.71	.87	.60	.08	**.76**	.06	.29
7. When I remember events, in general I can recall which day of the week it was.	.78	.56	.88	.49	.09	**.63**	.09	.05
8. When I remember events, in general I can recall people, what they looked like, or what they were wearing.	.78	.58	.87	.52	.24	**.61**	.00	.05
Semantic								
9. I can learned and repeat facts easily, even if I don't remember where I learned them.	.55	.41	.88	.44	.22	.26	.05	.30
10. After I have read a novel or newspaper, I forget the facts after a few days. (R)	.58	.33	.88	.33	.14	.18	-0.2	.37
11. After I have met someone once, I easily remember his or her name.	.59	.31	.88	.33	.12	.39	.05	.06
12. I can easily remember the names of famous people (sports figures, politicians, celebrities	.50	.51	.88	.46	.23	.34	.03	.28
13. I have a hard time remembering information I have learned at school or work. (R)	.61	.23	.88	.23	.05	-.06	.12	**.53**
14. I am very good at remembering information about people that I know (e.g., the names of a co-worker's children, their personalities, places friends have visited etc.).	.59	.30	.88	.45	.21	**.48**	.10	.13
Spatial								
15. In general, my ability to navigate is better than most of my family/friends.	.61	.67	.88	.29	-.06	.09	**.82**	.05
16. After I have visited an area, it is easy for me to find my way around the second time I visit.	.60	.71	.88	.48	.09	.28	**.86**	.02
17. I have a hard time judging the distance (e.g., in meters or kilometers) between familiar landmarks. (R)	.76	.23	.88	.23	.16	-.09	.26	.30
18. I get lost easily, even in familiar areas. (R)	.66	.54	.88	.31	-.01	.05	**.61**	.22
19. If my route to work or school was blocked, I could easily find the next faster way to get there.	.64	.65	.88	.47	.14	.19	**.66**	.14
20. I use specific landmarks for navigating.	.78	.04	.88	.26	.33	.24	.03	-.14
Future								
21. When I imagine an event in the future, the event generates vivid mental images that are specific in time and place.	.88	.74	.87	.56	**.74**	.27	-.02	.11
22. When I imagine an event in the future, I can picture the spatial layout.	.87	.79	.87	.65	**.80**	.17	.21	.19
23. When I imagine an event in the future, I can picture people and what they look like.	.87	.79	.87	.55	**.81**	.17	.02	.12
24. When I imagine an event in the future, I can imagine how I may feel.	.89	.62	.88	.46	**.64**	.18	-.01	.08
25. When I imagine an event in the future, I can picture images(e.g., people, objects, etc)	.87	.80	.87	.65	**.79**	.24	.12	.16
26. I have a difficult time imagining specific events in the future. (R)	.90	.61	.88	.44	**.63**	.07	-.01	.23

Note: (R) indicates items that were reverse coded. Bold text denotes factor loadings above 0.4.

to the finding in the younger sample, where three out of six SAM-semantic items did not load above .4. Cumulative variance explained was 45.23%, six percentage points higher than for younger adults.

As with younger adults, confirmatory factor analyses showed moderate fits for the three-factor, four-factor, and high-order models, although fit indices were slightly lower in older adults (see Table S2). Moreover, the four-factor model slightly outperformed the three-factor and high-order models, as demonstrated by a lower standardized root mean square residual. Results were qualitatively similar when models were estimated on the full sample with diagonally weighted least squares (see Table S2).

Next we examined the psychometric properties of SAM subscale and total scores in older adults (distributions in Table 5). Pairwise correlations between SAM scores were highly positive (Fig. 2b). A mixed ANOVA modeling gender, subscale, and their interaction was conducted to test for gender differences on SAM scores. A main effect of subscale ($F(4, 724) = 26.10$, $p < .001$ with Greenhouse-Geisser correction, $\eta^2_p = .13$) and an interaction effect between gender and subscale ($F(4,724) = 3.20, p < .05, \eta^2_p = .02$) were observed. Post hoc pairwise t tests revealed that SAM-future scores were higher in women ($M = 95.92$, $SD = 16.30$) than in men ($M = 91.44$, $SD = 13.46$; Cohen's $d = .30, p < .05$; Fig. 3 right panel). We additionally tested whether years of education had an

effect on SAM scores, since older adults had long completed standard education. No associations were detected (all p values > 0.10).

Validity of the SAM

Associations with memory and fluid cognition

Younger adults First, we examined how SAM scores related to performance-based composites of episodic memory, semantic memory, and fluid cognition. Consistent with domain specificity, SAM-semantic was positively related to semantic memory scores (Table 6). No other SAM scores showed significant relationships with the three composite scores.

Dual-process accounts of episodic memory assert that many of the tests included in the episodic memory composite are not process-pure measures of recollection (Yonelinas, 2002; Yonelinas et al. 2010), leading us to separately examine associations between SAM-episodic and Remember-Know recollection, familiarity, and source memory. No significant relationships were observed (recollection: $\rho(125) = .07$, $p = .468$; familiarity: $\rho(125) = -.06$, $p = .507$, source: $\rho(125) = .03$, $p = .715$).

Additional post hoc correlations were conducted between SAM-semantic and the individual measures of semantic

Table 5 Distribution of standard SAM scores and averaged SAM scores

Subscale	Standard Scores			Average Scores			Correlation, r (p)
	Range	M	SD	Range	M	SD	
Younger Adults (N=203)							
Episodic	73.00-133.83	100.19	13.84	1.25-4.75	3.33	0.72	.96 (.000)
Semantic	75.84-134.63	100.41	12.60	1.83-4.83	3.45	0.64	.94 (.000)
Spatial	68.61-121.00	97.77	15.62	1.00-5.00	3.71	0.80	.94 (.000)
Future	77.98-126.45	97.08	14.77	1.00-5.00	3.83	0.80	.87 (.000)
Total	73.77-136.79	98.92	16.29	1.97-4.73	3.58	0.50	.90 (.000)
Older Adults (N=184)							
Episodic	73.00-136.87	100.34	14.24	1.13-5.00	3.33	0.75	.96 (.000)
Semantic	75.57-130.10	96.60	11.90	1.50-4.67	3.24	0.67	.93 (.000)
Spatial	68.61-121.00	103.80	14.21	1.67-5.00	4.03	0.69	.95 (.000)
Future	77.98-126.45	93.75	15.18	1.00-5.00	3.65	0.89	.86(.000)
Total	72.25-142.42	93.21	13.67	2.10-4.83	3.56	0.56	.92 (.000)
Full Sample (N=387)							
Episodic	73.00-136.88	100.26	14.01	1.13-5.00	3.33	0.74	.96 (.000)
Semantic	75.57-134.63	98.60	12.40	1.50-4.83	3.35	0.66	.94 (.000)
Spatial	68.61-121.00	100.64	15.25	1.00-5.00	3.86	0.77	.95 (.000)
Future	77.98-126.45	95.50	15.04	1.00-5.00	3.74	0.85	.86 (.000)
Total	72.25-142.42	98.58	13.46	1.97-4.83	3.57	0.52	.91(.000)

Note: Standard scores reflect scores produced by the multivariate weighting scheme (Palombo et al., 2013). Average scores reflect scores produced by taking the average of raw responses (reverse-coded where appropriate) for a given subscale or the whole measure.

Table 6 Validity of the SAM: correlations with episodic, semantic, and fluid performance scores

Performance Measure	SAM-episodic	SAM-semantic	SAM-spatial	SAM-future	SAM-total
Younger Adults (N=186)					
Episodic Memory	.00 (.952)	.02(.772)	-.13(.090)	.07 (.359)	.00 (.969)
Semantic Memory	.10 (174)	**.20(.007)***	.03 (.672)	-.04 (.590)	.12 (.117)
Fluid Cognition	-.08(.280)	-.02 (.751)	.00 (.986)	-.09(.209)	-.10(.186)
Older Adults(N=168)					
Episodic Memory	-.07 (.392)	-.06 (.433)	.03 (.704)	.10 (.189)	.02 (.815)
Semantic Memory	-0.05 (.527)	-.04 (.603)	-.08(.309)	.14 (.067)	.01 (.909)
Fluid Cognition	.09 (.254)	.02 (.783)	.17 (.031)	.07 (.346)	.10 (.218)
Full Sample (N=354)					
Episodic Memory	-.02 (.661)	-.02 (649)	-.01 (915)	.05 (.336)	.01 (.795)
Semantic Memory	.02 (.757)	.10 (.070)	-.02 (.711)	.06 (.229)	.06 (247)
Fluid Cognition	.01 (.817)	-.04 (.432)	.09 (.105)	-.04(.406)	-.02 (.742)

Note. Correlations shown are Spearman ρ with p values in parentheses. Correlations for the full sample include age and gender as covariates.

*survives Bonferroni correction at $p<.017$. Bold text indicates significance.

memory. Out of three measures, SAM-semantic reliably co-varied only with Oral Reading Recognition ($\rho(184) = .21$, $p < .01$).

Composite scores of memory and fluid cognition largely involved verbal measures. We extracted the single visuospatial measure, Shipley Blocks, and conducted a post hoc correlation with SAM-spatial as an additional validity test (see Clark & Maguire, 2020 for a thorough assessment). No relationship was observed.

Older adults SAM scores were not significantly related to any of the three composite scores (Table 6). SAM-episodic was not related to recollection ($\rho(85) = -.10$, $p = .373$), familiarity ($\rho(85) = -.04$, $p = .725$), or source memory ($\rho(85) = .07$, $p = .550$). Post hoc correlations were conducted to inspect whether SAM-semantic scores could predict scores on individual measures of semantic memory and whether SAM-spatial associated with an individual measure of visuospatial memory. In contrast to younger adults, here SAM-semantic scores were not related to semantic memory, and SAM-spatial showed a positive relationship with Shipley Blocks ($\rho(166) = .17$, $p < .05$).

Associations with autobiographical memory

Younger adults Next, we examined relationships between SAM scores and the AI (Table 7). Younger adults provided memories that were, on average, vivid ($M = 4.54/6$, $SD = .61$) and not well rehearsed ($M = 2.32/6$, $SD = .80$). We hypothesized that internal detail variables would positively relate to SAM-episodic, SAM-spatial, and SAM-future scores if the SAM was a valid measure of AM. We also hypothesized that SAM-semantic would positively relate to external detail scores. We observed a positive relationship between SAM-

spatial and internal detail count. No relationships were detected between SAM-semantic and external detail scores. Instead, a positive relationship was observed between SAM-semantic and internal detail count. Associations including specific probe (Table S4) were qualitatively similar, with some noteworthy differences. SAM-semantic, SAM-spatial, and SAM-total all showed positive associations with internal detail count. These subscales were also associated with total details. AI measures that control for total verbal output (proportion and density scores) were not correlated with SAM subscales.

We also predicted that subcategories of the AI would associate with the SAM: namely that place and perceptual details would relate to SAM-spatial and that semantic details would relate to SAM-semantic. Similar to Palombo et al. (2013), SAM-spatial positively correlated with internal place detail count ($\rho(143) = .17$, $p < .05$), but not with internal place density ($\rho(143) = -.04$, $p = .631$). All SAM scores except SAM-future positively related to internal perceptual detail count (SAM-episodic, $\rho(143) = .18$, $p < .05$; SAM-semantic, $\rho(143) = .24$, $p < .005$; SAM-spatial, $\rho(143) = .23$, $p < .01$; SAM-total, $\rho(143) = .25$, $p < .005$). However, only SAM-episodic and SAM-total retained significant relationships with internal perceptual density (SAM-episodic, $\rho(143) = .22$, $p < .01$; SAM-total, $\rho(143) = .22$, $p < .01$). With respect to semantic details, only SAM-spatial showed a positive correlation with semantic detail count ($\rho(43) = .23$, $p < .01$), but not with semantic density ($\rho(143) = .09$, $p = .297$).

If the SAM is a valid measure of vividness, we hypothesized that SAM-episodic—corresponding to the re-experiencing, rather than the recollection, of an event—would uniquely relate to self-reported vividness on the AI. Self-reported vividness of specific autobiographical events from the AI was positively related to SAM-episodic, SAM-semantic, SAM-future, and SAM-total scores.

Table 7 Validity of the SAM: Correlations with the autobiographical interview

Detail Score	SAM-episodic	SAM-semantic	SAM-spatial	SAM-future	SAM-total
Younger Adults (N=145)					
Internal Details	.11 (.172)	**.23 (.005)***	**.22 (.009)**	.04 (.640)	.19 (.025)
External Details	.08 (.333)	.10 (.216)	.19 (.022)	-.06(.447)	.08 (.331)
Internal Proportion	.00 (.960)	.09 (.291)	-.06 (.447)	.10 (.234)	.05 (.535)
Total Details	.12 (.170)	.21 (.011)	.21 (.010)	.02 (.788)	.17 (.044)
Word Count	.05 (.582)	.19 (.025)	.21 (.013)	.03 (.766)	.13 (.125)
Internal Density	.12 (.149)	.06 (.466)	-.08 (.323)	.01 (.903)	.08 (.340)
External Density	.06 (.445)	-.09 (.307)	.06 (.496)	-.14 (.100)	-.03 (.677)
Self-Report Vividness	**.37 (.000)***	**.23 (.005)***	.17 (.048)	**.28 (.001)***	**.41 (.000)***
Older Adults (N=112)					
Internal Details	.16 (.099)	.15 (.110)	.07 (.463)	**.25 (.008)**	.24 (.010)
External Details	.14 (.157)	.08 (.399)	-.06 (.554)	.09 (.349)	.14 (.154)
Internal Proportion	-.05 (.578)	.04 (.664)	.11 (.242)	.09 (.322)	.02 (.828)
Total Details	.16 (.090)	.16 (.102)	.01 (.885)	.21 (.030)	.23 (.017)
Word Count	.16 (.096)	.20 (.037)	-.03 (.756)	.19 (.051)	.21 (024)
Internal Density	-.03 (.730)	-.10 (.303)	.15 (.126)	.02 (.804)	-.02 (799)
External Density	.04 (.655)	-.11 (.234)	-.04 (.715)	-.09 (.331)	-.04 (.701)
Self-Report Vividness	**.23 (.017)**	.18 (.062)	-.01 (.942)	.14 (.144)	.24 (.013)
Full Sample (N=257)					
Internal Details	**.13 (.031)**	**.19 (.002)***	**.16 (.013)**	**.12 (.05)**	**.21 (.001)***
External Details	.10 (.101)	.09 (.158)	.07 (.242)	.02 (.728)	.10 (.100)
Internal Proportion	-.01 (.866)	.06 (.323)	.02 (.699)	.08 (.212)	.05 (.436)
Total Details	.13 (.040)	**.18 (.004)***	.13 (.044)	.10 (.106)	**19 (.002)***
Word Count	.08 (.190)	**.18 (.005)***	.09 (.136)	.10 (.112)	.15 (.015)
Internal Density	.07 (.237)	.00 (.954)	.03 (.646)	-.01 (.843)	.05 (.396)
External Density	.04 (.548)	-.09 (.172)	.00 (945)	-.11 (.092)	-.05 (.467)
Self-Report Vividness	**.29 (.000)***	**.20 (.001)***	.08 (.180)	**.22 (.001)***	**.32 (.000)***

Note: Correlations shown are Spearman ρ with *p* values in parentheses. Correlations for the full sample include age and gender as covariates.

*survives Bonferroni correction at p<.006. Bold text indicates significance.

Older adults We hypothesized the same pattern of associations in older adults. Older adults recalled memories that were vivid ($M = 4.79/6$, $SD = .81$) and not well rehearsed ($M = 1.81/6$, $SD = .63$). One positive association emerged between SAM-future and internal detail count (Table 7), but not with proportion or density scores which control for verbal output. No relationships were observed with internal place, internal perceptual, or semantic detail scores (counts and density). Unlike in the younger sample, only the relationship between SAM-episodic and self-reported vividness ratings was found to be significant. Results were highly similar when associations with the AI included specific probe (Table S4).

Associations with personality and social cognition

Finally, SAM scores were related to several non-memory measures within each age group. Motivated by the finding that individuals reporting a history of depression had lower SAM-episodic scores (Palombo et al., 2013), we examined

relationships between SAM scores and depression indices. Associations with personality and social cognition were conducted to determine whether method variance belies association with the SAM (Clark and Maguire, 2020). We did not have domain-specific predictions for these analyses. Instead, we predicted that SAM associations with other self-report measures would exceed those observed with performance-based measures. We summarize the broad patterns observed below and in Tables S5–S7, noting that several associations had higher magnitudes than anticipated.

Younger adults In corroboration with previous results, SAM-episodic scores were negatively related to BDI scores (Table 8). The relationship remained even after controlling for age, education, and gender.

Out of the 28 assessments of personality and social cognition, SAM scores were correlated with two measures in younger adults (Table S5). SAM-episodic, SAM-semantic, and SAM-total scores were all positively related to extraversion

Table 8 Validity of the SAM: Correlations with depression scales by age group

Measure	SAM-episodic	SAM-semantic	SAM-spatial	SAM-future	SAM-total
BDI (N=115)	**-.22 (.019)***	.06 (.546)	-0.04 (.644)	-.07 (.453)	-.12 (.205)
GDS (N=166)	**-.17 (.028)***	**-.32 (.000)***	**-.23 (.003)***	-.07 (.377)	**-.25 (.001)***

Note: BDI=Beck Depression Inventory; GDS=Geriatic Depression Scale. Correlations shown are Spearman ρ with *p* values in parentheses.

**p*<.05. Bold text indicates significance.

and self-efficacy. SAM-total scores were also positively related to openness.

Older adults SAM-episodic, SAM-semantic, SAM-spatial, and SAM-total scores were negatively related to GDS (Table 8). These associations remained significant after controlling for age, education, and gender.

Many more associations with personality and social cognition were observed in older adults (see Table S6 for full listing). These associations mainly comprised relationships between SAM scores and measures of personality, positive/negative affect, and empathy. The SAM was related to all five dimensions of personality. SAM-episodic, SAM-future, and SAM-total scores showed positive associations with openness, conscientiousness, extraversion, and agreeableness. SAM-spatial was negatively associated with neuroticism, anger hostility, fear affect, perceived rejection, perceived stress, sadness, and loneliness. SAM-semantic scores also showed a negative relationship with loneliness. Conversely, SAM-spatial scores were positively related to positive affect, general life satisfaction, meaning and purpose, emotional support, and self-efficacy. SAM-semantic and SAM-total scores were also positively related to meaning and purpose, emotional support, and self-efficacy. Positive associations were also observed between SAM-future and SAM-total scores and measures of empathy, including empathic concern, perspective taking, and fantasy subscales of the Interpersonal Reactivity Index (IRI) and the Toronto Empathy Questionnaire. SAM-episodic, SAM-semantic, and SAM-total scores were negatively associated with IRI perceived distress.

Age group comparisons

We compared SAM scores between younger and older adults to test whether the SAM subscales would detect and reproduce well-documented age-related changes in cognition. A mixed ANOVA modeling age group as a between-subjects factor and SAM subscale as a within-subjects factor was conducted on SAM scores. Gender was included as a covariate of no interest. We observed a significant main effect of subscale ($F(4,1540) = 14.74, p < .001$ with Greenhouse-Geisser correction, $\eta^2_p = .03$) and a significant interaction between age group and subscale ($F(4,1540) = 13.82, p < .001, \eta^2_p = .03$). Post hoc *t* tests demonstrated that the interaction was driven

by age group differences in SAM-spatial (Fig. 3). Specifically, older adults had higher SAM-spatial scores ($t(383) = 6.03$, $p < .005$; see Table 5).

These age group differences in self-reported memory from the SAM stand in contrast to age group differences in performance-based measures of episodic memory, semantic memory, and AM. Younger adults had higher episodic memory scores ($T(273) = 19.42, p < .001$; Cohen's $d = 2.11$), whereas older adults had higher semantic memory scores ($T(341) = 10.08, p < .001$; Cohen's $d = 1.08$). Similarly, age group differences on the SAM did not correspond to age group differences in episodic (internal) and semantic (external) details in AM. Younger adults recounted more internal details ($T(254.81) = 3.49, p < .001$) on the AI, even when controlling for verbosity (internal proportion: $T(208.78) = 9.54, p < .001$; internal density: $T(254.77) = 8.61, p < .001$). In contrast, older adults communicated more external details ($T(176.5) = 3.76, p < .001$), even when controlling for verbosity ($T(197.18) = 7.88, p < .001$). Older adults also reported higher vividness ratings ($T(201.64) = 2.64, p < .01$), but lower frequency of rehearsal ($T(254.86) = 5.71, p < .001$). No age group differences in verbosity were present (total details: $T(246.61) = .76, p = .45$; word count: $T(236.64) = .65, p = .52$). All differences remained when controlling for gender.

Associations with memory and fluid cognition

Associations between SAM scores and performance-based measures were then revisited controlling for age and gender to determine whether SAM scores were associated with cognition in a larger, more highly powered sample. For performance-based measures of fluid cognition and memory ($N = 354$), no significant relationships emerged with SAM scores (Table 6). No significant relationships were observed when SAM-semantic was related to individual semantic memory measures or when SAM-spatial was related to performance on the Shipley Blocks mental rotation task.

Associations with autobiographical memory

We also repeated associations with our performance-based measure of AM, the AI, and self-reported vividness ratings on the AI controlling for age and gender ($N = 257$; Table 7).

We had a priori hypotheses that internal detail variables would positively relate to SAM-episodic, SAM-spatial, and SAM-future scores if the SAM was a valid measure of individual differences in AM. We observed positive relationships between internal detail count and all SAM scores. Verbosity scores also showed significant relationships with SAM-semantic and SAM-total. Notably, no relationships were observed between the SAM and AI detail scores that controlled for verbosity. We also predicted that external detail variables would positively relate to SAM-semantic scores, but no relationships were observed. Associations with the AI including specific probe were nearly identical (Table S4).

In our targeted analyses of associations with select subcategories on the AI, all SAM scores except SAM-future were related to perceptual detail count (SAM-episodic, $\rho(254) = .14$, $p < .05$); SAM-semantic, $\rho(254) = .20$, $p < .005$; SAM-spatial, $\rho(254) = .15$, $p < .05$; SAM-total $\rho(254) = .21$, $p < .001$). When controlling for verbosity (perceptual density), only relationships with SAM-episodic and SAM-total remained (SAM-episodic, $\rho(254) = .14$, $p < .05$; SAM-total, $\rho(254) = .16$, $p < .01$). Place and semantic detail counts and densities were not related to any SAM score.

If the SAM corresponded specifically to individual differences in properties of AM, we predicted that SAM-episodic would positively relate to vividness, even in the absence of a direct link to recollection. Self-reported vividness demonstrated a positive association with SAM-episodic. This association was not specific, however; positive relationships were also observed with SAM-semantic, SAM-future, and SAM-total scores.

Associations with personality and social cognition

Personality and social cognition relationships to the SAM in the full sample controlling for age and gender incorporated most of the associations from each sample with some additions ($N = 354$; Table S7). In terms of personality, openness and agreeableness were related to all SAM scores except SAM-spatial, and neuroticism was negatively related to SAM-episodic, SAM-spatial, and SAM-total scores. Notably, positive associations with friendship and general life satisfaction surfaced: friendship was associated with SAM-episodic and SAM-total, while general life satisfaction was related to SAM-semantic and SAM-total. Similarly, meaning and purpose was positively related to all SAM scores but SAM-spatial. Perceived stress was negatively related to SAM-semantic. IRI perceived distress was also negatively related to SAM-spatial.

Scoring

The scoring of the SAM is proprietary and based on a complex multidimensional weighting of items. We created new SAM scores using average Likert responses to probe whether the weighted scoring scheme added value to simpler alternatives. Average scores were highly positively correlated with the standard multidimensional weighting scores (Table 5, right panel). Associations with performance-based measures, the AI, and personality and social cognition were nearly identical with average SAM scores.

General Discussion

The present study evaluated the psychometric properties of the SAM as (i) a self-report survey of perceived recollective capacity and (ii) a measure of AM ability. We investigated the psychometric properties of the SAM on reliability, validity, stability across samples (gender and age groups), and the ability of this self-report measure to recapitulate known age effects in AM. The results converged to indicate that the SAM is a reliable assessment of self-perceived recollective ability. However, validity of the SAM as a measure of AM was not supported. Reliability, which was moderate among most subscales except for SAM-semantic, was informed by testing the SAM's internal consistency, latent factor structure, relationships among subscales, and gender differences. Factor analyses confirmed that three- and four-factor solutions had moderate fits, yet loadings of SAM-episodic and SAM-semantic items on several factors urges against considering SAM-episodic and SAM-semantic as independent subscales. SAM-semantic was associated with performance-based semantic memory in younger adults. No other associations with performance-based laboratory measures were observed. Critically, we examined associations between self-reported AM on the SAM and AM recalled on the AI. In the full cohort, the total number of episodic-like (internal) details was positively associated with all SAM subscales and total score. However, none of these associations remained significant when controlling for verbal output in the AI event narratives. The SAM also related to self-reported vividness of recalled autobiographical events, a phenomenological property of AM that shares method variance with the SAM. Against predictions, this association was observed for SAM-episodic, SAM-semantic, and SAM-future subscales. Beyond comparisons with memory measures, the SAM consistently covaried with measures of personality, self-efficacy, and depressive symptoms. With respect to age comparisons, performance of the SAM was not stable across the younger and older cohorts, demonstrating disparate patterns of associations with other measures. Finally, age effects on the SAM did not recapitulate known patterns of age-related cognitive change. We review these findings below and discuss implications for the SAM as a measure of self-perceived recollective capacity and episodic AM.

Reliability and validity of the SAM

SAM-episodic

Overall, SAM-episodic had high internal consistency as a measure of self-perceived capacity to consciously recollect episodic details of past events, but was not a valid measure of episodic memory ability. Previously, SAM-episodic was related to recollection, but not familiarity, of scenes in a Remember-Know recognition memory test (Palombo et al., 2013; Rudebeck et al., 2009; Yonelinas, 2001). In the absence of a similar relationship with SAM-semantic, and considering the foundational role that scene construction plays in memory re-instantiation and imagination (Hassabis & Maguire, 2007), the authors interpreted this finding as evidence that SAM-episodic had unique overlap with objective recollection. In our results, neither recollection nor familiarity on Remember-Know was related to SAM-episodic. It should be noted that scene and object recollection could not be disambiguated in the version administered here. Others have disputed the utility of Remember-Know paradigms as valid measures of episodic memory (Rubin & Umanath, 2015), finding that scores associate more highly with a belief in accuracy of memory than with a sense of reliving (Rubin et al., 2003; Rubin & Siegler, 2004). Here, SAM-episodic strongly correlated with self-efficacy in both younger and older adults, suggesting that self-reported episodic abilities may scale more with confidence than actual episodic memory. Our results replicated the original finding that lower SAM-episodic scores relate to a greater frequency of depressive symptoms (Palombo et al., 2013) in young adults. For older adults, relationships were observed between depressive symptoms and SAM-episodic, SAM-semantic, SAM-spatial, and SAM-total scores.

For young adults on the AI, SAM-episodic was reliably related to the subcategory of perceptual details, and this relationship persisted when controlling for verbosity. No relationships between the SAM and episodic (internal) AM details were observed for older adults. Within the full sample, SAM-episodic was positively related to total episodic details, but this relationship was not significant when controlling for verbosity on the AI. The absence of association reported here between SAM-episodic and episodic-like AM details is consistent with earlier reports (Palombo et al., 2013; Hebscher et al., 2018; Clark & Maguire, 2020). Notably, SAM-episodic was strongly associated with vividness, a secondary self-report measure of AM, in younger adults and across the entire sample. SAM-episodic has since been used to validate the Autobiographical Recollection Test (ART), a self-report assessment of several recollective properties including reliving, vividness, visual imagery, scene, narrative coherence, life story relevance, and rehearsal (Berntsen et al., 2019). Taken together, these findings provide support for the SAM as a measure of self-perceived recollective capacity.

SAM-semantic

SAM-semantic showed poor internal consistency and demonstrated only a single significant association with performance-based measures of memory. SAM-semantic has been considered to be a discrete and valid measure of self-perceived capacity to recollect semantic knowledge. This assessment is based largely on earlier reports of diverging associations between SAM-semantic and SAM-episodic scores. Coutanche and Koch (2017) observed a marginal association between SAM-semantic and a vocabulary measure in young adults. Here we report that only one of three semantic tasks, a reading task, was associated with SAM-semantic in younger adults; no relationships were observed for older adults. SAM-semantic was positively related to episodic-like details on the AI. No associations were observed with semantic-like (external) details in younger adults and across the full sample. A similarly unpredicted pattern was observed for specific subcategories of the AI: SAM-semantic showed a positive relationship with perceptual details and no relationship to semantic details. The impact of verbosity on observed associations suggests that SAM-semantic may track narrative length more than specific detail type. It is possible that our AI scoring procedure, which did not distinguish between general and personal semantics, obscured a relationship to SAM-semantic. Updated scoring of non-episodic details has shown to be effective in distinguishing between various forms of dementia (Strikwerda-Brown et al., 2018; Renoult et al., 2020). These findings leave open the possibility that SAM-semantic may correspond more to general than personal semantics.

SAM-semantic was also positively associated with self-reported vividness in younger adults and in the full sample. Previously, vividness of autobiographical recollection has been associated with SAM-episodic but not SAM-semantic (Sheldon et al., 2016). Recollective properties on the ART, such as vividness, have also shown a positive association with SAM-semantic, although lower in magnitude compared to that of SAM-episodic (Bernsten et al., 2019). It is unclear whether SAM-episodic and SAM-semantic correspond to similar or different features of AM phenomenology (i.e., vividness and narrative coherence, respectively). Nevertheless, our findings highlight that SAM-episodic and SAM-semantic subscales likely measure overlapping capacities. We elaborate on this point in the following section.

Relationship between SAM-episodic and SAM-semantic

Scores on SAM-episodic and SAM-semantic were highly correlated. Unlike findings from Picco et al. (2020), a three-factor solution to the SAM was almost identical to, and slightly exceeded, the four-factor solution. While the interrelationship between these subscales corroborates the original validation (Palombo et al., 2013), it is unclear from these data whether

self-perceived recollection of episodic and semantic information should be considered dissociable capacities. Does the shared variance reflect the interaction of episodic and semantic memory constructs or imprecise measurement? Compelling evidence suggests that episodic and semantic memory are not as distinct as first conceived (Tulving, 1972). Semantic memory develops early (Wheeler et al., 1997; Newcombe et al., 2007) and provides a scaffold from which episodic processes can emerge (see Irish & Piguet, 2013 for a review) once a sense of subjective time, self, and autonoetic awareness have also developed (Tulving, 2002; Martin-Ordas et al., 2014). Some degree of shared variance between episodic and semantic memory could therefore be expected. However, SAM-episodic and SAM-semantic both showed scant relationships to laboratory episodic and semantic memory tasks, perhaps suggestive of imprecise measurement. Since AM has been shown to recruit different neural resources than traditional laboratory-measured episodic and semantic memory (Gilboa, 2004; McDermott et al., 2009), we assessed the SAM's ability to measure episodic and semantic AM capacities more directly with objective performance on the AI.

We predicted that SAM subscales, capturing self-reported capacity to recollect AM, would be associated with specific measures of AM and vividness derived from the AI. We further predicted that the strength of associations would exceed associations with less ecologically valid, laboratory tests of episodic memory. A large body of work suggests that more ecologically valid tasks should augment the correlation between subjective and objective memory (Berry, West, & Dennehey, 1989; Schmidt, Berg, & Deelman, 2001; Sunderland, Watts, Baddeley, & Harris, 1986 as in Crumley et al., 2014). While SAM scores were associated with both AM and vividness, there was little evidence of specificity in these associations. SAM-episodic and SAM-semantic were both positively correlated with the number of episodic-like details and self-reported vividness of retrieved autobiographical events on the AI. This suggests that differences between self-perceived capacity to recollect episodic and semantic autobiographical information does not align with differences in an individual's episodic and semantic AM ability. This finding ran counter to our predictions, as the AI presently remains a gold-standard measure of episodic and semantic AM (Williams & Broadbent, 1986; Levine et al., 2002). As such, covariance between these AI measures and the analogous subscales on the SAM was expected, but unsupported by our data.

An alternative prediction is that individuals recall episodic event details using different strategies, such as visual imagery. Failing to consider these factors may conceal more nuanced differences in perceived versus actual recollective abilities. Consistent with this idea, Armson and colleagues (2021) demonstrated that SAM-episodic was positively associated with greater visually guided exploration during episodic versus non-episodic recollection. These findings suggest that SAM-episodic captures a greater reliance on visual imagery specific to episodic recollection. However, a similar association was observed between visual exploration behaviors and SAM-semantic scores. It is possible that visual imagery may be a non-specific recollective scaffold, where one builds confidence in their self-perceived recollective capacity without indexing specific episodic and semantic features of AM.

Taken together, our findings suggest that SAM-episodic and SAM-semantic assess self-perceived, yet domain-general, recollection capacity. There is little evidence that these subscales assess distinct features of recollective experience as predicted by common models of AM. Our findings underscore a critical need to revisit the mental capacities assessed by the SAM-episodic and SAM-semantic subscales and their relationship to individual differences in trait mnemonics (Sheldon et al., 2016; Fan et al., 2020a, Fan et al., 2020b). We suggest caution when interpreting these subscales as measures of distinct or independent cognitive constructs.

SAM-spatial

High internal consistency was observed for both younger and older samples, despite two poor-performing questions. Younger men reported higher spatial abilities than younger women, in line with previous work on the SAM and spatial memory (Palombo et al., 2013; see Levine et al., 2016 for a review). Whether the absence of a gender difference in the older sample reflects less of a disparity in objective spatial memory with age is unclear: some studies report a persistent gender gap into older age (León et al., 2016) while others report decline of a male advantage for spatial memory into later life (Lacreuse et al., 1999).

SAM-spatial has consistently shown correspondence to objective performance in spatial tasks, but a reliable association to spatial AM has yet to emerge. Although here we did not identify relationships with performance-based composite measures, SAM-spatial displayed a significant relationship with a single visuospatial measure in older adults. SAM-spatial has shown predictive validity for spatial navigation in cohorts of younger adults (Clark & Maguire, 2020; Selarka et al., 2019). Unlike SAM-episodic and SAM-semantic subscales, SAM scores of perceived spatial recollective capacity may adhere to visuospatial abilities, and the association may persist across the adult lifespan. Recent work provides further evidence for this idea. SAM-spatial scores were related to individual differences in spatial imagery across the adult lifespan (Fan et al., 2020b), consistent with the observed association with place features of AM (Palombo et al., 2013). While these studies suggest that SAM-spatial demonstrates the strongest association between perceived and actual ability of all the subscales, it is important to note that scaling by verbosity eliminated relationships with AM in our data.

Further, SAM-spatial was unrelated to AM vividness in our samples. Vividness may be independent of individual differences in perceived capacity for spatial recollection. Of note, Clark and Maguire (2020) observed an association between self-reported vividness on the AI and SAM-spatial in younger adults. It is likely that Clark and Maguire's larger sample of younger adults had more predictive power to detect an association that we observed only as trending in younger adults. SAM-spatial was unrelated to vividness in older adults or the full sample. It may be that visual and spatial imagery decouple with age, but we do not directly test that here.

Together, these findings suggest that SAM-spatial has predictive validity with respect to spatial tasks and spatial imagery, but not spatial AM. However, isolating spatial aspects of AM using the AI is particularly challenging. It is difficult to reliably demarcate spatial from more salient sensory details in everyday experiences. In addition, the prominence of landmark-based place details on the AI pertains to the items on SAM-spatial with the least internal consistency, potentially denigrating associations with SAM-spatial. Further validation testing with specific narrative prompts could be carried out for a more precise evaluation of SAM-spatial.

SAM-future

Our data suggest that the SAM-future subscale had excellent reliability. Older women rated their abilities higher than older men. Related work has failed to find behavioral evidence for gender differences in other mentalizing abilities such as creativity or divergent thinking (Reese et al., 2001; Abraham et al., 2014), but less is known about future thinking. Women do tend to perform better than men on episodic memory tasks and recount their memories with greater episodic specificity during AM interviews (Herlitz & Rehnman, 2008; Pillemer et al., 2003; Asperholm et al., 2019). This is posited to relate to women's better memory for vocabulary and abstract items that exist along the verbal–spatial continuum, such as faces, and perhaps even the future. While speculative, we suggest that these gender differences may be accentuated in older adulthood, where individuals have a larger repertoire of memories on which to draw and repurpose for imagining future events.

We were unable to assess the validity of the SAM-future subscale vis-à-vis performance measures of future thinking, as we did not include objective measures of future thinking, prospection, or imagination. Further, as the AI did not include a future thinking prompt, validation with a measure of AM was limited. Future validation work involving this subscale of the SAM would benefit from the implementation of the proposed adaptations to the AI, emphasizing future-oriented thinking (Addis et al., 2007).

SAM-total

SAM-total demonstrated excellent reliability, but there was no evidence supporting its validity as a global measure with respect to objective memory. Significant SAM-total relationships with AM were observed for total number of episodic-like details; however, this association was eliminated when controlling for verbosity. SAM-total was also associated with self-reported vividness on the AI, both in younger adults and across the full sample. As a summative score of the four subscales, SAM-total presumably represents an overall rating of self-perceived recollection capacity (Palombo et al., 2013). However, given the failure to find any associations with performance-based measures of memory, SAM-total cannot be considered a valid measure of AM. Further, given the lack of association between SAM-total scores and the cognitive measures implemented here, caution is warranted in using the SAM as an assay of cognitive abilities more generally.

Stability in different samples

Early validation studies of the SAM suggested it was a valid trait-like measure of autobiographical recollection (Palombo et al., 2013). A valid trait or individual differences measure should be unaffected by group assignment. This criterion generally held for internal consistency, latent factor structure, and associations with performance-based measures of memory and cognition across our two age groups. However, significant age differences emerged in associations between SAM scores and other self-report measures. More associations between SAM scores and personality, social cognition, and depressive indices were observed for older versus younger adults. Consistent with this idea, personality, self-efficacy, and depressive symptoms have been found to be more strongly correlated with subjective memory than objective memory and general cognitive function in older adults (Bandura, 1989; Perrig-Chiello, et al., 2000; Snitz et al., 2015; Herreen & Zajac, 2017). Yet a highly similar pattern of SAM associations with personality and social cognition emerged when pooling both samples and controlling for age and gender. Controlling for age when relating subjective memory abilities on the SAM to other self-report measures, such as personality, may underestimate or inflate observed effects.

These findings suggest mixed performance of the SAM across cohorts, but the SAM's relationship to AM was notably stable. We review the associations with AM, elaborate on age-group effects, and discuss their implications below.

Replication of age effects on memory

A core motivation for the study was to investigate the validity of the SAM as a measure of AM. As we know that AM undergoes a fundamental transition with age (Levine et al.

2002), we reasoned that these differences would be reflected in age differences across the SAM subscales. We leveraged our well-powered samples of younger and older adults with SAM, episodic memory, semantic memory, and AM data to inspect whether we could replicate well-known age group differences using the SAM. Specifically, we tested whether we could reproduce age-related decreases in episodic memory or episodic-like AM recall and increases in semantic memory or semantic-like AM recall (e.g., Park & Reuter-Lorenz, 2009; Levine et al., 2002) in the SAM. We did not find any age group difference in SAM-episodic or SAM-semantic; rather, older adults endorsed greater spatial abilities than younger adults. SAM-spatial has previously shown high predictive validity with navigation tasks (Clark & Maguire, 2020; Selarka et al., 2019), in which younger adults tend to outperform older adults (e.g., Moffat et al., 2001). Our findings could not be attributed to sample-specific properties, since performance-based measures of episodic, semantic, and AM revealed group differences in the expected directions. A group difference specific to SAM-spatial also rules out the possibility that older adults simply show overconfidence when rating their memory abilities (Dodson et al., 2007; Toth et al., 2011). Older adults do retain exceptional navigation for highly familiar environments (Rosenbaum et al., 2012), and could be drawing on preserved semantic memory to recall highly rehearsed spatial maps when rating spatial memory abilities. Indeed, age-related navigation decrements are often self-reported in the context of unfamiliar routes and places (Burns, 1999; Moffat, 2009). As discussed above, it is possible that SAM-spatial measures distinct features of spatial memory in different populations, especially if individuals reflect on abilities maintained rather than lost, to appraise their memories. A further possibility is that older adults, despite objective declines in AM, retain specific phenomenological aspects of AM. Alzheimer's disease patients, who show a greater discrepancy between objective and subjective AM than healthy older adults (El Haj & Antoine, 2017), are still able to experience emotionality and salience of AM despite loss of other recollective features (El Haj et al., 2016). In a similar manner, normally aging older adults may demonstrate preferential retention of spatial imagery. Conversely, older adults may simply overestimate spatial abilities more than other memory abilities (West et al., 2002).

Pooling data from both age groups and controlling for age, we would predict a reduction in the influence of latent age-related variables, thereby maximizing our power to detect reliable associations with performance-based memory measures. However, across the full sample, no significant associations were observed. All SAM subscales showed a positive relationship with the total number of episodic-like details on the AI. However, these were eliminated when controlling for verbosity in autobiographical narratives, measured as proportion or density scores. The importance of controlling for verbosity is corroborated by SAM associations with word count, total details, and self-reported extraversion and self-efficacy (Table S7). Here too, self-reported AM vividness was positively related to all SAM subscales except SAM-spatial.

In sum, the prediction that SAM is a valid measure of AM was not supported. Our data do demonstrate that the SAM may measure general recollective capacity more than discrete aspects of re-experiencing.

SAM as a measure of cognitive health: Assessing mental imagery and vividness

The lack of association between perceived ability and performance, as observed here for the SAM and performance-based memory measures, is pervasive across the psychological sciences. The disparity may be explained by the dynamic interplay of self-efficacy and monitoring (National Research Council, 1994; Moores et al., 2006; Bernacki et al., 2014) or poor reliability of performance-based tasks (Dang et al., 2020). Self-perceptions of memory ability nevertheless remain an important indicator of cognitive health. Subjective memory complaint is a critical risk factor for cognitive decline (Schmand et al., 1996; Schofield et al., 1997; Steinberg et al., 2013; Mitchell et al., 2014), often preceding amnestic mild cognitive impairment, an early stage of Alzheimer's disease (e.g., Buckley et al., 2016; Neto & Nitrini, 2016; Norton et al., 2017). For this reason, many instruments have already been devised including the Memory Functioning Questionnaire (Gilewski & Zelinski, 1988; Gilewski et al., 1990), Memory Assessment Clinics Self-Rating Scale (Crook & Larrabee, 1990; Crook et al., 1992), Memory Self-Efficacy Questionnaire (Berry et al., 1989), Spatial Self-Efficacy Questionnaire (West et al., 2002), Metamemory in Adulthood Questionnaire (Dixon & Hultsch, 1983a, 1983b), Memory Controllability Inventory (Lachman et al., 1995), and Personal Beliefs About Memory Inventory (Lineweaver & Hertzog, 1998). Much of this work has focused on episodic memory, and not AM, showing small but reliable relationships to traditional laboratory episodic memory tasks (Beaudouin & Desrichard, 2011; Crumley et al., 2014; Hertzog & Pearman, 2014). Yet impaired senses of visual imagery and re-experiencing are also characteristic symptoms of amnesia (Rubin & Umanath, 2015). Memory complaints paired with imagery degradation may be an early marker of dementia (El Haj et al., 2016). In this context, the SAM, as another self-reported measure of perceived recollective capacity, may serve as a marker of cognitive health.

Since reports of impoverished memory recollection cannot be verified as accurate, a sense of re-experiencing may serve as an approximation. Even if the SAM does not track with memory per se, it could show promise as a diagnostic tool in its relationship to AM vividness. Leading theories suggest that

AM is supported by scene construction via the hippocampus (Ranganath & Ritchey, 2012) and requires an accompanying sense of re-experiencing, or vividness (Hassabis & Maguire, 2007; Maguire & Mullally, 2013). During mental elaboration of AMs, brain activity in the hippocampus was found to track subjective vividness whereas brain activity in lateral temporal and parietal regions tracked greater quality of objective episodic recollection (Thakral et al., 2020). Vividness, by way of the hippocampus, may therefore signal the availability of episodic traces located in other parts of cortex (D'Angiulli et al., 2013; Moscovitch et al., 2005). A sense of mental visual imagery, in contrast, may not be necessary for remembering, but has been shown to be highly correlated with ratings of vividness (Rubin et al., 2003; Rubin & Umanath, 2015; Clark & Maguire, 2020). Both imagery and vividness appear to be constructs tapped by the SAM.

Imagery

Individuals with severely deficient AM often report an impoverished sense of mental imagery, or aphantasia, accompanied by a specific difficulty for tasks that draw on visual memory. These individuals show less engagement of brain areas involved in AM while remembering, although remembering itself is not significantly impaired (Palombo et al., 2015). Critically, they demonstrate lower SAM-episodic scores than controls (Sheldon et al., 2016). More recent work with aphantastic individuals has also demonstrated lower SAM-future and SAM-semantic scores (Dawes et al., 2020). Individuals with highly superior AM display higher SAM-episodic scores than controls (Sheldon et al., 2016). Together, these cases suggest that the SAM may capture aspects of imagery associated with AM. Indeed, both SAM-episodic and SAM-future predict imagery ability on an objective scene construction task, albeit not as reliably as other self-report vividness questionnaires (Clark & Maguire, 2020).

Accumulating evidence suggests that SAM-episodic and SAM-spatial may relate to distinct aspects of imagery specific to objects and spatial relationships, respectively (Fan et al., 2020b). As previously mentioned, SAM-spatial predicts performance on spatial navigation tasks but not imagery tasks (Clark & Maguire, 2020). Moreover, spatial imagery, as measured on the Object-Spatial Imagery and Verbal Questionnaire Spatial subscale, also reliably predicted navigation performance. This is noteworthy since object and spatial imagery have been shown to differentially contribute to phenomenological aspects of AM and mental reconstruction of events (Aydin, 2018). In fact, Alzheimer's disease patients have shown impairments in spatial imagery with preserved object imagery and AM (El Haj et al., 2019). Although the SAM was designed to measure episodic, semantic, spatial, and future aspects of AM, it may instead capture distinct aspects of AM imagery.

Vividness

In terms of a more direct link to AM, vividness has previously been shown to positively scale with SAM-episodic but not SAM-semantic (Sheldon et al., 2016). This prompted authors to explore the divergent functional connectivity patterns associated with "perceptual" and "conceptual" trait mnemonics. Our own results indicate that vividness relates to SAM-episodic and SAM-semantic, as well as SAM-future and SAM-total. Recently Fan et al. (2020a) related SAM-episodic to age and cognitive function in a large sample of older adults and explained their findings in terms of memory strategies related to vividness. However, as reviewed above, such an interpretation is problematic given the low reliability of the SAM-semantic subscale, unclear distinction between SAM-episodic and SAM-semantic, and validity concerns raised by the findings reported here.

Self-perceived memory ability may not be a trait

There is little evidence to support that memory appraisal is a stable trait. Higher correlations between subjective and objective memory have been observed in older adults compared to younger adults (Parisi et al., 2012; Crumley et al., 2014), and are thought to be linked to higher education or greater insight into memory abilities (Zelinski et al., 2001; Crumley et al., 2014). While education was not related to the SAM in our results, these findings suggest that the correspondence between perceived and actual abilities is dynamic, rather than trait-like. Indeed, *changes* to subjective memory are more robustly predictive of *changes* to performance on memory tasks (Snitz et al., 2015). The updating of beliefs, which necessitates data from multiple time points, lays the foundation for the most predictive models of learning outcomes (e.g., Higashi et al., 2019) and plays an integral role in successful forecasting (Mellers et al., 2015). Although subjective memory is correlated with trait personality, if memory changes over time, it follows that self-perceived memory and mnemonics (Bouazzaoui et al., 2010) should change accordingly. Changes perceived as drastic, such as with subjective memory complaints, are extremely useful to detect cognitive impairment. Given the SAM's ease of administration, researchers may learn more about individual differences in autobiographical mnemonics with data from multiple time points. A critical first step in this direction would be to determine test-retest reliability of the SAM in future psychometric assessments.

Scoring

The SAM is scored based on a multiple correspondence analysis used in the SAM's conception. Here we replicated all results obtained with the multivariate weighting procedure using an easy-to-calculate average of the Likert scores

(respecting reverse scoring). Recent work used similar average scores, which implies that the weighted scoring scheme may be unnecessary (Fan et al., 2020b). Proprietary scoring can impede the thorough examination needed for the SAM to provide meaningful insights into how the sense of remembering impacts cognition.

Conclusion

The SAM demonstrated reliability as a self-report measure of perceived recollective capacity. However, assessing psychometric validity in the measurement of this construct is challenging, and validation with observable memory performance is a matter of debate. Consistent with earlier reports, our findings suggest that SAM subscale scores may reflect phenomenological aspects of autobiographical recall, and vividness specifically. However, we observed the most robust associations between SAM scores and a measure of self-efficacy, suggesting that the SAM may more reliably index confidence in domain-general self-report abilities. Critically, our findings demonstrate that the SAM is not a psychometrically valid measure of trait mnemonics as assessed with widely used performance-based memory measures. Scores on SAM-episodic and SAM-semantic subscales should not be interpreted as independent or specific, which stands in contrast to how these constructs are considered under standard models of declarative memory. Based on our findings, and in agreement with previous research (Clark & Maguire, 2020), we urge caution in the use of the SAM as a measure of AM pending revision and further psychometric validation.

References

Abraham, A., Thybusch, K., Pieritz, K., & Hermann, C. (2014). Gender differences in creative thinking: Behavioral and fMRI findings. *Brain Imaging and Behavior*, 8(1), 39–51. https://doi.org/10.1007/s11682-013-9241-4

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, 45(7), 1363–1377.

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, 145(8), 785–821. https://doi.org/10.1037/bul0000197

Aydin, C. (2018). The differential contributions of visual imagery constructs on autobiographical thinking. *Memory*, 26(2), 189–200. https://doi.org/10.1080/09658211.2017.1340483

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. Developmental Psychology, 25, 729–735. https://doi.org/10.1037/0012-1649.25.5.729

Barnes, C. A. (1988). Aging and the physiology of spatial memory. *Neurobiology of Aging*, 9(C), 563–568. https://doi.org/10.1016/S0197-4580(88)80114-3

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(7), 813–822. https://doi.org/10.1111/j.1469-7610.1997.tb01599.x

Bastin, C., Feyers, D., Jedidi, H., Bahri, M. A., Degueldre, C., Lemaire, C., … Salmon, E. (2013). Episodic autobiographical memory in amnestic mild cognitive impairment: What are the neural correlates? *Human Brain Mapping*, 34(8), 1811–1825. https://doi.org/10.1002/hbm.22032

Beaudoin, M., & Desrichard, O. (2011). Are Memory Self-Efficacy and Memory Performance Related? A Meta-Analysis. *Psychological Bulletin*, 137(2), 211–241. https://doi.org/10.1037/a0022106

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). Manual for the Beck Depression Inventory-II. ychological Corporation

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2014). Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. *Metacognition and Learning*, 10(1), 99–117. https://doi.org/10.1007/s11409-014-9127-x

Berntsen, D., Hoyle, R. H., States, U., & Rubin, D. C. (2019). The Autobiographical Recollection Test (ART): A Measure of Individual Differences in Autobiographical Memory. *Journal of Applied Research in Memory and Cognition*, 8, 305–318.

Berry, J. M., West, R. L., & Dennehey, D. M. (1989). Reliability and Validity of the Memory Self-Efficacy Questionnaire. *Developmental Psychology*, 25(5), 701–713. https://doi.org/10.1037/0012-1649.25.5.701

Bluck, S., & Alea, N. (2002). Exploring the functions of autobiographical memory: Why do I remember the autumn? In J.D. Webster, & B.K. Haight (Eds.). *Critical advances in reminiscence: From theory to application* (pp. 61-75). Springer.

Bouazzaoui, B., Isingrini, M., Fay, S., Angel, L., Vanneste, S., Clarys, D., & Taconnat, L. (2010). Aging and self-reported internal and external memory strategy uses: The role of executive functioning. *Acta Psychologica*, 135(1), 59–66. https://doi.org/10.1016/j.actpsy.2010.05.007

Brainerd, C. J., Reyna, V. F., Gomes, C. F., Kenney, A. E., Gross, C. J., Taub, E. S., Spreng, R. N., & Alzheimer's Disease Neuroimaging Initiative (2014). Dual-retrieval models and neurocognitive impairment. *Journal of experimental psychology. Learning, memory, and cognition*, 40(1), 41–65. https://doi.org/10.1037/a0034057

Buckley, R. F., Maruff, P., Ames, D., Bourgeat, P., Martins, R. N., Masters, C. L., … Ellis, K. A. (2016). Subjective memory decline predicts greater rates of clinical progression in preclinical Alzheimer's disease. *Alzheimer's and Dementia*, 12(7), 796–804. https://doi.org/10.1016/j.jalz.2015.12.013

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Burns, P. C. (1999). Navigation and mobility of older drivers. Journal ofGerontology: Social Sciences, 54B, S49–S55.

Clark, I. A., & Maguire, E. A. (2020). Do questionnaires reflect their purported cognitive functions? *Cognition*, *195*. https://doi.org/10.1101/583690

Conway, M. A., Loveday, C., & Cole, S. N. (2016). The remembering–imagining system. *Memory Studies*, *9*(3), 256–265. https://doi.org/10.1177/1750698016645231

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. https://doi.org/10.1037//0033-295X.

Coutanche, M. N., & Koch, G. E. (2017). Variation across individuals and items determine learning outcomes from fast mapping. *Neuropsychologia*, *106*(June), 187–193. https://doi.org/10.1016/j.neuropsychologia.2017.09.029

Crook, T. H., Feher, E. P., & Larrabee, G. J. (1992). Assessment of Memory Complaint in Age-Associated Memory Impairment: The MAC-Q. *International Psychogeriatrics*, *4*(2), 165–176. https://doi.org/10.1017/S1041610292000991

Crook, T. H., & Larrabee, G. J. (1990). A self-rating scale for evaluating memory in everyday life. *Psychology and Aging,* *5*(1), 48–57. https://doi.org/10.1037/0882-7974.5.1.48

Crumley, J. J., Stetler, C. A., & Horhota, M. (2014). Examining the relationship between subjective and objective memory performance in older adults: A meta-analysis. *Psychology and Aging*, *29*(2), 250–263. https://doi.org/10.1037/a0035908

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450. Retrieved from https://proxy.library.mcgill.ca/login?url=https://search.proquest.com/docview/1297349869?accountid=12339

D'Angiulli, A., Runge, M., Faulkner, A., Zakizadeh, J., Chan, A., Morcos, S., 2013. Vividness of visual imagery and incidental recall of verbal cues, when phenomenological availability reflects long-term memory accessibility. Front. Psychol. 4, 1–18.

Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures Weakly Correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-65705-7

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Dixon, RA, & Hultsch, DF. (1983a). Metamemory and memory for text relationships in adulthood: A cross-validation study. *Journal of Gerontology, 38*, 689–694.

Dixon, RA, & Hultsch, DF. (1983b). Structure and development of metamemory in adulthood. *Journal of Gerontology, 38*, 682–688.

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollec- tion account. Psychology and Aging, 22, 122–133.

El Haj, M., & Antoine, P. (2017). Discrepancy between subjective autobiographical reliving and objective recall: The past as seen by Alzheimer's disease patients. *Consciousness and Cognition*, *49*, 110–116.

El Haj, M., Gallouj, K., & Antoine, P. (2019). Mental imagery and autobiographical memory in Alzheimer's disease. *Neuropsychology, 33*(5), 609–616. https://doi.org/10.1037/neu0000521

El Haj, M., Kapogiannis, D., & Antoine, P. (2016). Phenomenological reliving and visual imagery during autobiographical recall in Alzheimer's disease. Journal of Alzheimer's Disease, 52(2), 421–431. https://doi.org/10.3233/JAD-151122.

Fan, C., Romero, K., & Levine, B. (2020a). Older adults with lower autobiographical memory abilities report less age-related decline in everyday cognitive function. *BMC Geriatrics*, *20*(308), 1–12. https://doi.org/10.31234/osf.io/zqs78

Fan, C., Abdi, H., & Levine, B. (2020b). On the relationship between autobiographical episodic memory and spatial navigation. *Memory & Cognition*. https://doi.org/10.31234/osf.io/pnbfz

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State" A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fraley, R. C., & Marks, M. J. (2007). The null hypothesis significance testing debate and its implications for personality research. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 149–169). Guilford Press.

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1997). On reporting recollective experiences and "direct access to memory systems". *Psychological Science*, *8*(5).

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Allyn & Bacon.

Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology*, *80*(11 Suppl 3), S2–S6. https://doi.org/10.1212/WNL.0b013e3182872e5f

Gilboa A. (2004). Autobiographical and episodic memory–one and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia, 42*(10), 1336–1349. https://doi.org/10.1016/j.neuropsychologia.2004.02.014

Gilewski, M. J., & Zelinski, E. M. (1988). Memory Functioning Questionnaire (MFQ). *Psychopharmacology Bulletin*, *24*(4), 665–670. https://doi.org/10.4324/9781003076391-63

Gilewski, M. J., Zelinski, E. M., & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging*, *5*(4), 482–490. https://doi.org/10.1037/0882-7974.5.4.482

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*: Pearson College Division.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Healy, S. D., & Jozet-Alves, C. (2009). Spatial Memory. *Encyclopedia of Animal Behavior*, 304–307. https://doi.org/10.1016/B978-0-08-045337-8.00145-5

Hebscher, M., Levine, B., & Gilboa, A. (2018). The precuneus and hippocampus contribute to individual differences in the unfolding of spatial representations during episodic autobiographical memory. *Neuropsychologia*, *110*(March 2017), 123–133. https://doi.org/10.1016/j.neuropsychologia.2017.03.029

Hemphill, J. F. (2003). Interpreting the magnitude of correlation coefficients. *American Psychologist, 58*, 78–79.

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

Herreen, D., & Zajac, I. (2017). The Reliability and Validity of a Self-Report Measure of Cognitive Abilities in Older Adults: More Personality than Cognitive Function. *Journal of Intelligence*, *6*(1), 1. https://doi.org/10.3390/jintelligence6010001

Hertzog, C., & Pearman, A. (2014). Memory Complaints in Adulthood and Old Age. In T. J. Perfect & D. S. Lindsay (Eds.), *The Sage Handbook of Applied Memory* (pp. 1–25). SAGE Publications Ltd. https://doi.org/10.4135/9781446294703.n24

Higashi, H., Minami, T., & Nakauchi, S. (2019). Cooperative update of beliefs and state-transition functions in human reinforcement learning. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-53600-9

Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in Behavioral Neuroscience*, *7*(April), 27. https://doi.org/10.3389/fnbeh.2013.00027

Jeffery, K. J. (2016). *Spatial memory. The Curated Reference Collection in Neuroscience and Biobehavioral Psychology* (Second Edition, Vol. 3). Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.21077-8

Lachman, M. E., Bandura, M., Weaver, S. L., & Elliott, E. (1995). Assessing Memory Control Beliefs: The Memory Controllability Inventory. *Aging, Neuropsychology, and Cognition*, *2*(1), 67–84. https://doi.org/10.1080/13825589508256589

Lacreuse, A., Herndon, J. G., Killiany, R. J., Rosene, D. L., & Moss, M. B. (1999). Spatial cognition in rhesus monkeys: Male superiority declines with age. *Hormones and Behavior*, *36*(1), 70–76. https://doi.org/10.1006/hbeh.1999.1532

LePort, A. K. R., Mattfeld, A. T., Dickinson-Anson, H., Fallon, J. H., Stark, C. E. L., Kruggel, F., … McGaugh, J. L. (2012). Behavioral and neuroanatomical investigation of Highly Superior Autobiographical Memory (HSAM). *Neurobiology of Learning and Memory*, *98*(1), 78–92. https://doi.org/10.1016/j.nlm.2012.05.002

León, I., Tascón, L., & Cimadevilla, J. M. (2016). Age and gender-related differences in a spatial memory task in humans. *Behavioural Brain Research*, *306*, 8–12. https://doi.org/10.1016/j.bbr.2016.03.008

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(2), 127–155. https://doi.org/10.1002/wcs.1380

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689. https://doi.org/10.1037//0882-7974.17.4.677

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lineweaver, T. T., & Hertzog, C. (1998). Adults' efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology, and Cognition*, *5*(4), 264–296. https://doi.org/10.1076/anec.5.4.264.771

Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: a manifesto for change. *Journal of experimental psychology. General*, *142*(4), 1180–1189. https://doi.org/10.1037/a0033650

Mar, R. A., Spreng, R. N., & Deyoung, C. G. (2013). How to produce personality neuroscience research with high statistical power and low additional cost. *Cognitive, Affective and Behavioral Neuroscience*, *13*(3), 674–685. https://doi.org/10.3758/s13415-013-0202-6

Martin-Ordas, G., Atance, C. M., & Caza, J. S. (2014). How do episodic and semantic memory contribute to episodic foresight in young children? *Frontiers in Psychology*, *5*(JUL), 1–11. https://doi.org/10.3389/fpsyg.2014.00732

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, *47*(11), 2290–2298. https://doi.org/10.1016/j.neuropsychologia.2008.12.025

Mellers, B., Stone, E., Atanasov, P., Rohrbaugh, N., Emlen Metz, S., Ungar, L., … Tetlock, P. (2015). The psychology of intelligence analysis: Drivers of prediction accuracy in world politics. *Journal of Experimental Psychology: Applied*, *21*(1), 1–14. https://doi.org/10.1037/xap0000040

Meyer, G. J., Finn, S. E., Eyde, L., Kay, G. G., Moreland, K. L., Dies, R. R., ... Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128–165.

Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(6), 439–451. https://doi.org/10.1111/acps.12336

Moffat, S. D. (2009). Aging and spatial navigation: What do we know and where do we go? *Neuropsychology Review*, *19*(4), 478–489. https://doi.org/10.1007/s11065-009-9120-3

Moffat, S. D., Zonderman, A. B., & Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, *22*(5), 787–796. https://doi.org/10.1016/S0197-4580(01)00251-2

Moores, T. T., Cha-Jan, J. C., & Smith, D. K. (2006). Clarifying the Role of Self-Efficacy and Metacognition as Predictors of Performance: Construct Development and Test. *Data Base for Advances in Information Systems*, *37*, 125–132. https://doi.org/10.1145/1161345.1161360

Moscovitch, M., Rosenbaum, R.S., Gilboa, A., Addis, D.R., Westmacott, R., Grady, C., McAndrews, M.P., Levine, B., Black, S., Winocur, G., et al., 2005. Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. Journal of Anat. 207, 35–66.

National Research Council. (1994). Learning, Remembering, Believing: Enhancing Human Performance. The National Academies Press. 10.17226/2303.

Neto, A. S., & Nitrini, R. (2016). Subjective cognitive decline: The first clinical manifestation of Alzheimer's disease? *Dementia & Neuropsychologia*, *10*(3), 170–177. https://doi.org/10.1590/s1980-5764-2016dn1003002

Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development Of Episodic And Autobiographical Memory: A Cognitive Neuroscience Perspective. *Advances in Child Development and Behavior*, *35*, 37–85. https://doi.org/10.1016/B978-0-12-009735-7.50007-4

Norton, D. J., Amariglio, R., Protas, H., Chen, K., Aguirre-Acevedo, D. C., Pulsifer, B., … Quiroz, Y. T. (2017). Subjective memory complaints in preclinical autosomal dominant Alzheimer disease. *Neurology*, *89*(14), 1464–1470. https://doi.org/10.1212/WNL.0000000000004533

Nunnally, J. C. (1978). *Psychometric theory*: New York : McGraw-Hill, 1978. 2nd edition.

Palombo, D. J., Alain, C., Söderlund, H., Khuu, W., & Levine, B. (2015). Severely deficient autobiographical memory (SDAM) in healthy adults: A new mnemonic syndrome. *Neuropsychologia*, *72*, 105–118. https://doi.org/10.1016/j.neuropsychologia.2015.04.012

Palombo, D. J., Sheldon, S., & Levine, B. (2018). Individual differences in autobiographical memory. *Trends in Cognitive Sciences*, 1–15. https://doi.org/10.1016/j.tics.2018.04.007

Palombo, D. J., Williams, L. J., Abdi, H., & Levine, B. (2013). The survey of autobiographical memory (SAM): A novel measure of trait mnemonics in everyday life. *Cortex*, *49*(6), 1526–1540. https://doi.org/10.1016/j.cortex.2012.08.023

Parisi, J. M., Gross, A. L., Rebok, G. W., Saczynski, J. S., Cook, S. E., Langbaum, J. B. S., … Fredrick, W. (2012). Perceptions : Findings from the ACTIVE Study. *Psychologica Aging*, *26*(3), 518–524. https://doi.org/10.1037/a0022458

Park, D.C. & Reuter-Lorenz, P. (2009). The adaptive brain : Aging and neurocognitive scaffolding. *Annal Review Psychologica*, *60*, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656.The

Perrig-Chiello P, Perrig W, Stahelin H. (2000). Differential aspects of memory self-evaluation in old and very old people. *Aging & Mental Health*, *4*(2), 130–135.

Picco, S., Pedreira, M. E., & Fernández, R. S. (2020). Psychometric validation of the survey of autobiographical memory: confirmatory factor analysis and network analysis. *Memory*, *28*(8), 1037–1050. https://doi.org/10.1080/09658211.2020.1812662

Pillemer, D. B., Wink, P., DiDonato, T. E., & Sanborn, R. L. (2003). Gender differences in autobiographical memory styles of older adults. *Memory*, *11*(6), 525–532. https://doi.org/10.1080/09658210244000117

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*. https://doi.org/10.1038/nrn3338

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Reese, H. W., Lee, L. J., Cohen, S. H., & Puckett, J. M. (2001). Effects of intellectual variables, age, and gender on divergent thinking in adulthood. International Journal of Behavioral Development, 25, 491–500.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8,* 271–276. https://doi.org/10.2466/PMS.8.7.271-276

Renoult, L., Davidson, P. S. R., Palombo, D. J., Moscovitch, M., & Levine, B. (2012). Personal semantics: At the crossroads of semantic and episodic memory. *Trends in Cognitive Sciences*, *16*(11), 550–558. https://doi.org/10.1016/j.tics.2012.09.003

Renoult, L., Armson, M. J., Diamond, N. B., Fan, C. L., Jeyakumar, N., Levesque, L., Oliva, L., McKinnon, M., Papadopoulos, A., Selarka, D., St Jacques, P. L., & Levine, B. (2020). Classification of general and personal semantic details in the Autobiographical Interview. *Neuropsychologia, 144*, 107501. https://doi.org/10.1016/j.neuropsychologia.2020.107501

Rosenbaum, R. S., Winocur, G., Binns, M. A., & Moscovitch, M. (2012). Remote spatial memory in aging: All is not lost. *Frontiers in Aging Neuroscience*, *4*(SEP). https://doi.org/10.3389/fnagi.2012.00025

Rubin, D. C., Schrauf, R. R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. Memory & Cognition, 31, 887–901. doi:https://doi.org/10.3758/BF03196443

Rubin, D. C., & Siegler, I. C. (2004). Facets of personality and the phenomenology of autobiographical memory. Applied Cognitive Psychology, 18, 913–930.

Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*(1), 1–23. https://doi.org/10.1037/a0037907

Rudebeck, S. R., Scholz, J., Millington, R., Rohenkohl, G., Johansen-Berg, H., & Lee, a. C. H. (2009). Fornix Microstructure Correlates with Recollection But Not Familiarity Memory. *Journal of Neuroscience*, *29*(47), 14987–14992. https://doi.org/10.1523/JNEUROSCI.4707-09.2009

Sakaluk, J. K., & Short, S. D. (2017). A Methodological Review of Exploratory Factor Analysis in Sexuality Research: Used Practices, Best Practices, and Data Analysis Resources. *Journal of Sex Research*, *54*(1), 1–9. https://doi.org/10.1080/00224499.2015.1137538

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic future thinking: mechanisms and functions. *Current Opinion in Behavioral Sciences*, *17*, 41–50. https://doi.org/10.1016/j.cobeha.2017.06.002

Schmand, B., Jonker, C., Hooijer, C., & Lindeboom, J. (1996). Subjective memory complaints may announce dementia. *Neurology*, *46*(1), 121–125. https://doi.org/10.1212/WNL.46.1.121

Schmidt, I. W., Berg, I. N. A. J., & Deelman, B. G. (2001). Relations between subjective evaluations of memory and objective memory performance. *Perceptual and Motor Skills*, *93*, 761–776.

Schofield, P. W., Marder, K., Dooneief, G., Jacobs, D. M., Sano, M., & Stern, Y. (1997). Association of subjective memory complaints with subsequent cognitive decline in community-dwelling elderly individuals with baseline cognitive impairment. *American Journal of Psychiatry*, *154*(5), 609–615. https://doi.org/10.1176/ajp.154.5.609

Selarka, D., Rosenbaum, R. S., Lapp, L., & Levine, B. (2019). Association between self-reported and performance-based navigational ability using internet-based remote spatial memory assessment. *Memory*, *27*(5), 723–728. https://doi.org/10.1080/09658211.2018.1554082

Sheldon, S., Farb, N., Palombo, D. J., & Levine, B. (2016). Intrinsic medial temporal lobe connectivity relates to individual differences in episodic autobiographical remembering. *Cortex*, *74*, 206–216. https://doi.org/10.1016/j.cortex.2015.11.005

Shipley, W. C., Gruber, C. P, Martin, T. A., & Klein, A. M. (2009). *Shipley-2 manual*. Western Psychological Services.

Smith, A. (1982). *Symbol Digit Modalities Test (SDMT) manual (revised)*.

Snitz, B. E., Small, B. J., Wang, T., Chang, C.-C. H., Hughes, T. F., & Ganguli, M. (2015). Do subjective memory complaints lead or follow objective cognitive change? A five-year population study of temporal influence. *Journal of International Neuropsychological Society*, *21*(9), 732–742. https://doi.org/10.1017/S1355617715000922

Söderlund, H., Moscovitch, M., Kumar, N., Daskalakis, Z. J., Flint, A., Herrmann, N., & Levine, B. (2014). Autobiographical episodic memory in major depressive disorder. *Journal of Abnormal Psychology*, *123*(1), 51–60. https://doi.org/10.1037/a0035610

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment*, *91*(1), 62–71. https://doi.org/10.1080/00223890802484381

Spreng, R. N., Lockrow, A. W., DuPre, E., Setton, R., Spreng, K., & Turner, G. R. (2018). Semanticized autobiographical memory and the default - executive coupling hypothesis of aging. *Neuropsychologia, 110*, 37–43. https://doi.org/10.1016/j.neuropsychologia.2017.06.009

Stamenova, V., Gao, F., Black, S. E., Schwartz, M. L., Kovacevic, N., Alexander, M. P., & Levine, B. (2017). The effect of focal cortical frontal and posterior lesions on recollection and familiarity in recognition memory. *Cortex*, *91*, 316–326. https://doi.org/10.1016/j.cortex.2017.04.003

Steinberg, S. I., Negash, S., Sammel, M. D., Bogner, H., Harel, B. T., Livney, M. G., … Arnold, S. E. (2013). Subjective memory complaints, cognitive performance, and psychological factors in healthy older adults. *American Journal of Alzheimer's Disease and Other Dementias*, *28*(8), 776–783. https://doi.org/10.1177/1533317513504817

Strikwerda-Brown, C., Mothakunnel, A., Hodges, J. R., Piguet, O., & Irish, M. (2018). External details revisited - A new taxonomy for coding 'non-episodic' content during autobiographical memory retrieval. *Journal of Neuropsychology*. https://doi.org/10.1111/jnp.12160

Sunderland, A., Watts, K., Baddeley, A. D., & Harris, J. E. (1986). Subjective memory assessment and test performance in elderly adults. *Journals of Gerontology*, *41*(3), 376–384. https://doi.org/10.1093/geronj/41.3.376

Sutin, A. R., & Robins, R. W. (2007). Phenomenology of autobiographical memories: The memory experiences questionnaire. *Memory*, *15*(4), 390–411. https://doi.org/10.1080/09658210701256654

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (Seventh Ed). : Pearson. https://doi.org/10.1007/978-1-4757-2514-8_3

Thakral, P. P., Madore, K. P., & Schacter, D. L. (2020). The core episodic simulation network dissociates as a function of subjective experience and objective content. *Neuropsychologia*, *136*(November 2019).

Toth, J. P., Daniels, K. A., & Solinger, L. A. (2011). What you know can hurt you: Effects of age and prior knowledge on the accuracy of judgments of learning. Psychology and Aging, 26, 919–931.

Tulving, E. (1972). Episodic and semantic memory. In Organization of Memory, ed. E Tulving, W Donaldson, pp. 381–403. : Academic.

Tulving, E. (1985). Memory and consciousness. Canadian Psychology/ Psychologie canadienne, 26(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E. (2002). Episodic memory: From Mind to Brain. *Annual Review of Psychology*, *53*, 1–25.

Van Rossum, G. & Drake, F.L. (2009). Python 3 Reference Manual. CreateSpace, .

Wechsler, D. (2009). *Wechsler Memory Scale (Fourth Edition).* : Pearson.

West, R.L., Welch, D.C., & Knabb, P.D. (2002) Gender and Aging: Spatial Self-Efficacy and Location Recall, Basic and Applied Social Psychology, 24:1, 71-80, https://doi.org/10.1207/S15324834BASP2401_7

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*, *133*(1), 122–148. https://doi.org/10.1037/0033-2909.133.1.122

Williams, J. M., & Broadbent, K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology, 95*(2), 144–149. https://doi.org/10.1037/0021-843X.95.2.144

Yesavage, J. A., & Brink, T. L. (1983). Development and validation of geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *80*(1), 37–49. https://doi.org/10.2307/1957152

Yonelinas A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 356(1413), 1363–1374. https://doi.org/10.1098/rstb.2001.0939

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity : A Review of 30 Years of Research. *Journal of Memory and Language*, *46*, 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Zelinski, E. M., Burnight, K. P., & Lane, C. J. (2001). The relationship between subjective and objective memory in the oldest old: Comparisons of findings from a representative and a convenience sample. *Journal of Aging and Health*, *13*(2), 248–266.