



Published in final edited form as:

Memory. 2014 April ; 22(3): 163–170. doi:10.1080/09658211.2013.774021.

Examining the Long-Term Stability of Overgeneral Autobiographical Memory

Jennifer A. Sumner^{a,*}, Susan Mineka^a, Richard E. Zinbarg^{a,b}, Michelle G. Craske^c, Suzanne Vrshek-Schallhorn^a, and Alyssa Epstein^c

^aNorthwestern University, Department of Psychology, 2029 Sheridan Road, Suite 102, Evanston, IL 60208, USA

^bThe Family Institute at Northwestern University, 618 Library Place, Evanston, IL 60201, USA

^cUniversity of California, Los Angeles, Department of Psychology, Franz Hall Box 951563, Los Angeles, CA 90094, USA

Abstract

Overgeneral autobiographical memory (OGM) is a proposed trait-marker for vulnerability to depression, but relatively little work has examined its long-term stability. This study investigated the stability of OGM over several years in 271 late adolescents and young adults participating in a larger longitudinal study of risk for emotional disorders. The Autobiographical Memory Test (AMT) was administered twice, with test-retest intervals ranging from approximately 3 to 6 years. There was evidence of significant but modest stability in OGM over several years. Specifically, Spearman rank correlations (ρ s) between the proportions of specific and categoric memories generated on the two AMTs were .31 and .32, respectively. We did not find evidence that the stability of OGM was moderated by the length of the test-retest interval. Furthermore, the stability coefficients for OGM for individuals with and without a lifetime history of major depressive disorder (MDD) were relatively similar in magnitude and not significantly different from one another (ρ s = .34 and .42 for the proportions of specific and categoric memories for those with a history of MDD; ρ s = .31 for both the proportions of specific and categoric memories for those without a history of MDD). Implications for the conceptualization of OGM are discussed.

Keywords

overgeneral autobiographical memory; autobiographical memory specificity; long-term stability; depression

Over the past two and a half decades, it has become evident that not only *what* people remember but also *the way* in which they recall their personal past has important implications for psychological functioning. In particular, the degree of specificity of autobiographical memory recall has emerged as a cognitive phenomenon with special relevance to depression (see Sumner, 2012; Williams et al., 2007, for reviews). When asked to generate a specific memory that refers to an event that lasted less than one day and occurred at a particular time and place (e.g., “when I played tennis last Friday”), individuals with depression tend to be less specific and/or more overgeneral in their recall than nondepressed controls. Such overgeneral autobiographical memory (OGM) has been associated with diagnoses of major depressive disorder (MDD; Williams et al., 2007), and it predicts a worse course of depression (e.g., Hermans et al., 2008; see Sumner, Griffith, &

*Corresponding author: Tel.: 1 847 491 3843; fax: 1 847 491 7859; Jennifer.Sumner@gmail.com.

Mineka, 2010, for a meta-analysis) and subsequent onsets of major depressive episodes (Kleim & Ehlers, 2008; Sumner et al., 2011). OGM has been proposed as a marker of vulnerability to depression. Consequently, OGM has become the subject of much research in recent years, especially because of its potential clinical implications for prevention and intervention (e.g., Raes, Williams, & Hermans, 2009a).

The vast majority of studies on OGM have employed a particular cuing methodology known as the Autobiographical Memory Test (AMT; Williams, & Broadbent, 1986). On the AMT, individuals are presented with cue words (usually emotion words of positive and negative valence), and are asked to retrieve a unique specific memory in response to each cue within a given time limit (generally 30 or 60s). Despite the widespread use of the AMT, the psychometric properties of this methodology have only begun to gain attention in recent years (e.g., Griffith et al., 2009; 2011; Heron et al., 2012). Consequently, information on the test-retest reliability of the AMT is somewhat lacking. To the extent that OGM is a trait-marker for depression, it should exhibit significant stability over time.

To date, a few studies have examined the stability of OGM by administering the AMT at two points in time and comparing performance on the two tests. For example, in six samples (five undergraduate student samples and one MDD patient sample), there was a fair degree of stability in OGM across test-retest intervals ranging from 1 to 5 months (Raes, Hermans, & Williams, 2009b). In these studies, Spearman rank correlation coefficients for the numbers of specific memories on the AMT ranged from .53 to .68. Peeters, Wessel, Merkelbach, and Boon-Vermeeren (2002) obtained similar results in a sample of outpatients with MDD; the majority of Spearman correlations based on comparing AMT performance at baseline and at 3 and 7 months later ranged from .43 to .79, although it was not specified how the magnitudes of the stability coefficients were related to the length of the test-retest interval in this sample. Hermans et al. (2008) found evidence of greater stability of OGM in patients with MDD over 3 to 4 weeks ($r = .88$ for the numbers of specific memories). It is noteworthy, however, that Hermans et al. (2008) used the same exact AMT at both OGM assessments, and this may, in part, explain the large magnitude of their stability estimate. Most prior work has not administered the same version of the AMT twice in order to avoid practice effects (e.g., Raes et al., 2009). Together, these findings suggest that OGM—as measured with the AMT—is relatively stable over periods ranging up to several months. Nevertheless, information is lacking on the extent to which OGM is stable over longer periods of time. If OGM is indeed a trait marker of risk for depression as proposed (e.g., Williams et al., 2007), then it should exhibit some stability over a considerable span of time. An examination of the stability of OGM over several years would contribute further to our understanding of this aspect of OGM.

Another question that remains is whether the stability of OGM varies as a function of depression history. Prior studies have examined the stability of OGM in either nonclinical undergraduate samples (e.g., Raes et al., 2009b, Studies 1–5) or samples with MDD (e.g., Hermans et al., 2008). The lack of information on the long-term stability of OGM for individuals with and without a history of MDD is therefore another limitation of the extant literature. Additional research on how the stability of OGM might vary as a function of depression has the potential to refine our understanding of OGM and how it relates to psychopathology. For example, to the extent that OGM is a trait-marker for vulnerability to depression, and not an indicator of a current depressed state, then it should be similarly stable in individuals with and without a history of depression.

The goal of the current study was to investigate the long-term stability of OGM as measured with the AMT using data from the Northwestern-UCLA Youth Emotion Project (YEP), a two-site 8-year longitudinal study of risk factors for emotional disorders (e.g., Zinbarg et al.,

2010). In the YEP, the AMT has been administered twice, with test-retest intervals ranging from approximately 3 to 6 years. In addition to investigating the stability of OGM over a period of several years, we examined the extent to which the stability of OGM might vary as a function of 1) the length of the test-retest interval, or 2) a lifetime history of MDD.

Method

Participants

Participants were from a larger sample of late adolescents and young adults in an 8-year longitudinal study of risk factors for emotional disorders (the Youth Emotion Project, YEP). At recruitment, participants were juniors at one of two public high schools, one in suburban Chicago and one in suburban Los Angeles. At screening, participants completed the Eysenck Personality Questionnaire neuroticism scale (EPQ-R-N; Eysenck & Eysenck, 1975), and they were categorized as low-, medium-, or high-scorers (broken down by tertiles). High-EPQ-R-N-scorers were oversampled to obtain a behavioral high-risk sample for the development of emotional disorders (59% of the original sample of 627 participants were high-EPQ-R-N-scorers). Participants were recruited in three waves of cohorts between 2003 and 2005.

In the YEP, the AMT was administered twice (as described below). The current study used data from the 271 participants who completed both AMTs. There were approximately 4.5 years between the two AMT administrations for these individuals ($M = 4.7$, $SD = 0.8$), although the interval between the two AMTs ranged from 2.8 to 6.3 years. At the first AMT, these 271 participants ranged in age from 16 to 18 years ($M = 17.1$, $SD = 0.4$). Participants in this subsample were predominantly female (68%) and ethnically diverse (14% African American, 3% Asian, 51% Caucasian, 15% Hispanic, 11% Multiracial, and 6% Other). Participants who completed both AMTs did not differ significantly from those who completed the first, but not second, AMT in terms of gender, race/ethnicity, EPQ-R-N score at screening, OGM on the first AMT, or lifetime history of MDD at the time of the first AMT, $ps > .12$.

Measures

The Autobiographical Memory Test (AMT)—The AMT (Williams & Broadbent, 1986) is a cuing paradigm used to elicit autobiographical memories. For both AMT administrations in the YEP, participants were presented with cue words and instructed to retrieve a unique specific autobiographical memory in response to each cue. Participants were given 30 seconds to produce a response on each trial. To ensure understanding of the task, there were up to seven practice items on both AMTs; participants were required either to retrieve two consecutive specific memories or to complete all seven practice trials before moving on to the experimental trials. Participants were only given feedback on the practice items.

The first AMT was administered individually in person, and an experimenter presented cues that were printed in a booklet one at a time. The first AMT had 16 experimental trials that alternated between positive (*happy, loved, successful, energetic, comfortable, brave, safe, calm*) and negative (*lonely, failure, sad, hopeless, afraid, angry, tense, worried*) cue words. The second AMT was administered over the telephone. There were 12 experimental trials on the second AMT that again alternated between positive (*safe, ambitious, peace, hope, brave, interested*) and negative (*disappoint, inferior, hurt, frustrated, tense, regret*) cue words. For both AMTs, responses were made orally and tape-recorded. For each AMT, positive and negative cue words did not differ significantly in usage frequency. Additionally, collapsing across cue valence, words used in the first AMT had higher concreteness (based on ratings

from the MRC Psycholinguistic Database; Coltheart, 1981) than words used in the second AMT.

On each trial of the two AMTs, responses were scored as falling into one of five categories: specific memory, extended memory (a memory for an event that lasted longer than one day, e.g., “my summer vacation”), categoric memory (a memory for a summary or class of events, e.g., “parties with my friends”), semantic associate (a response containing general semantic information but no personal memory, e.g., “my dog”), or omission (no response). Errors on the task were also recorded (e.g., if a participant did not respond within the 30s time limit). Response coding had good inter-rater reliability for both AMT administrations. Mean kappas for within-site ($N = 73$) and cross-site ($N = 37$) reliability for the first AMT were both .78. Mean kappas for within-site ($N = 46$) and cross-site ($N = 46$) reliability for the second AMT were both .77. As is common in OGM research (e.g., Hermans et al., 2008), we used the proportion of specific memories and the proportion of categoric memories as our measures of OGM for each AMT. Consistent with recommendations from psychometric analyses of the AMT (e.g., Griffith et al., 2009; Heron et al., 2012), we collapsed across positive and negative cue words. Both AMTs were found to have good internal consistency (Cronbach’s $\alpha = .77$ and $.71$ for the first and second AMT, respectively).

Structured Clinical Interview for DSM-IV (SCID)—Advanced graduate students and Bachelor’s-level research assistants administered the SCID (First, Spitzer, Gibbon, & Williams, 2002) to assign Axis I diagnoses at baseline and annual follow-up assessments. Interviewer training included extensive self-study, didactics, role-playing, tests of diagnostic ability based on interview audio recordings, and observation of live interviews by doctoral-level supervisors. Interviewers’ assessments were supervised by doctoral-level clinical psychologists. Interviewers assigned clinician severity ratings (CSRs, rated on a 0–8 scale) to indicate the degree of impairment and/or distress associated with diagnoses; clinically significant diagnoses had CSRs greater than or equal to 4. SCID data were used to classify participants as those with or without a lifetime history of clinically significant MDD by the time of the second AMT administration. SCID diagnoses of clinically significant MDD have had generally good inter-rater reliability in the YEP (e.g., kappas for clinically significant MDD diagnoses ranged from .56 to 1.00 for the baseline and first five annual follow-up assessments).

Procedure

At the baseline assessment, a lifetime SCID was administered to assess Axis I psychopathology (additional measures not relevant to the current study were also administered at baseline; see Sutton et al., 2011). The SCID was also administered at annual follow-up assessments in order to assess Axis I psychopathology for the time since the previous interview. Approximately 9 months after the baseline assessment ($M = 9.0$, $SD = 2.0$), the first AMT was administered to a random subsample of 333 individuals as part of a larger battery of cognitive tasks. With the second AMT, all participants remaining in the YEP were offered the opportunity to complete this task. The second AMT was administered between 3.2 and 7.1 years after the baseline assessment ($M = 5.4$, $SD = 0.8$), and a total of 455 individuals completed the task. Of the 333 individuals who originally completed the first AMT, 271 also completed the task at the second administration. We thus examined data from these 271 participants in the present study.

Whereas the first AMT was administered approximately 9 months after the baseline assessment for participants in all three cohorts in 2003 through 2005, respectively, the second AMT was administered during the same period of time (2008 – 2011) to all

participants. Thus, the length of time between baseline and the second AMT administration varied as a function of participant cohort. Given this variability, we were interested in examining whether the stability of OGM might be moderated by the length of the test-retest interval (e.g., whether stability estimates might decrease as the test-retest interval increased).

Results

Overall, participants generally provided a majority of specific memories on both AMTs. The mean proportion of specific memories retrieved by participants on the first AMT ($M = .79$, $SD = .18$) was significantly greater than the mean proportion of specific memories retrieved on the second AMT ($M = .69$, $SD = .22$), $t(270) = 7.02$, $p < .0001$. In addition, the mean proportion of categoric memories retrieved on the first AMT ($M = .06$, $SD = .13$) was significantly greater than the mean proportion of categoric memories retrieved on the second AMT ($M = .02$, $SD = .05$), $t(270) = 5.02$, $p < .0001$. Participants made significantly more errors on the second AMT ($M_{proportion\ of\ errors} = .22$, $SD = .19$) than on the first AMT ($M_{proportion\ of\ errors} = .08$, $SD = .10$), $t(270) = 11.31$, $p < .0001$. Examples of errors on the AMT included not responding within the 30s time limit and repeating a memory response that was provided earlier in the task.

To examine the stability of OGM, we computed Spearman rank correlation coefficients (ρ_s) between our measures of OGM from the first and second AMTs. Spearman rank correlation coefficients were reported rather than Pearson correlation coefficients given some skewness in the OGM variables. However, results with Pearson correlation coefficients were similar to those with Spearman rank correlation coefficients (see footnotes 1 and 3 for details). For the entire sample of individuals who completed both AMTs ($N = 271$), the Spearman correlation between the proportions of specific memories was .31, and the Spearman correlation between the proportions of categoric memories was .32, both $ps < .0001$.¹ These estimates suggest that there was significant but modest stability in OGM over a period of several years.

Because the test-retest interval varied for different participant cohorts, we examined whether the relationship between the OGM measures from the two AMTs differed as a function of the length of the test-retest interval. We regressed performance on the second AMT on 1) performance on the first AMT, 2) the length of the test-retest interval in years, and 3) the interaction of first AMT performance with length of the test-retest interval. Performance on the first AMT and the length of the test-retest interval in years (both grand mean centered) were entered on the first step of the regression, and the interaction term was entered on the second step. The interaction of the proportion of specific memories generated on the first AMT with the length of the test-retest interval was not statistically significant, $b = -0.02$, $t(267) = -0.19$, $p = .85$. Furthermore, the additional variance in the proportion of specific memories generated on the second AMT that was accounted for by the interaction term was trivially small, R^2 change $< .0001$. Similarly, the interaction of the proportion of categoric memories generated on the first AMT with the length of the test-retest interval was not statistically significant, $b = 0.01$, $t(267) = 0.34$, $p = .73$. Again, the variance in the proportion of categoric memories generated on the second AMT that was accounted for by the interaction term was negligible, R^2 change $< .0001$. These findings are depicted in Figure 1. As shown in the graph, the stability coefficients for OGM were quite similar in magnitude for the 3-, 4-, and 5 plus-year test-retest intervals, and they did not differ significantly from one another. Together, these results fail to provide evidence that the stability of OGM was moderated by the length of the test-retest interval.

¹Results with Pearson correlation coefficients were similar to those with Spearman rank correlation coefficients: $r = .35$ for the proportion of specific memories and $r = .32$ for the proportion of categoric memories, both $ps < .0001$.

We also examined whether the stability of OGM differed as a function of a history of depression by comparing the stability coefficients for participants with ($n = 61$) and without ($n = 184$) a lifetime history of one or more major depressive episodes by the time of the second AMT administration.² The Spearman rank correlations between the proportions of specific memories were similar in magnitude for those with ($\rho = .34, p = .008$) and without a history of MDD ($\rho = .31, p < .0001$), and they did not differ significantly, $z = 0.22, p = .83$.³ Additionally, the correlations between the proportions of categoric memories were both statistically significant for those with ($\rho = .42, p = .001$) and without a history of MDD ($\rho = .31, p < .0001$). Although the stability coefficient based on categoric memories for participants with a history of MDD was larger in magnitude than the corresponding coefficient for those without a history of MDD, the two correlations did not differ significantly, $z = 0.84, p = .40$. Overall, these findings suggest that the long-term stability of OGM was relatively similar for individuals with and without a history of depression.

Discussion

Together, these findings suggest that there is significant but modest stability in OGM over time. Prior research has demonstrated that OGM is relatively stable over the course of up to several months (Hermans et al., 2008; Peeters et al., 2002; Raes et al., 2009b), and we have extended such work by investigating the stability of OGM over the course of several years. Even though some of the test-retest coefficients from previous studies are larger in magnitude than those obtained in our study, it is perhaps not surprising given that we had much longer test-retest intervals (i.e., 3 to 6 years) compared to prior research. Furthermore, to our knowledge we are the first to compare the long-term stability of OGM for individuals with and without a lifetime history of MDD. In general, those with and without a history of depression exhibited relatively similar patterns of stability in OGM over several years.

We believe that our findings suggest that there is an enduring component to OGM, even if the magnitudes of the stability coefficients are not very high. Recent theorizing on the stability of psychological constructs suggests that the patterns of continuity that exist in test-retest coefficients may be more important than the magnitude of those coefficients (Fraley & Roberts, 2005). Fraley and Roberts (2005) advise that, when conceptualizing the stability of a psychological construct, researchers should examine whether the test-retest coefficients approach either zero or a non-zero asymptote as the test-retest interval increases. Taken together with the extant literature, our findings suggest that the stability of OGM remains relatively stable with increasing time between the two assessments, rather than approaching zero. Indeed, not only did performance on the first AMT not interact significantly with the length of the test-retest interval to predict performance on the second AMT but, as shown in Figure 1, the stability coefficients for OGM were quite similar in magnitude over the range of test-retest intervals in this study.

The notion of an enduring component to OGM is also supported by the finding that the stability coefficients that we obtained in this study were rather similar in magnitude to the stability coefficients for other psychological constructs that are generally accepted to be trait-like. For example, Fraley and Roberts (2005) computed meta-analytic stability

²We excluded 26 participants with a history of clinically significant bipolar disorder ($n = 5$), posttraumatic stress disorder ($n = 7$), acute stress disorder ($n = 1$), dysthymia ($n = 14$), and/or psychotic symptoms ($n = 2$) by the time of the second AMT administration from these analyses (some participants had more than one disorder).

³Results with Pearson correlation coefficients were similar to those with Spearman rank correlation coefficients for those with and without a history of MDD. Pearson correlations between the proportions of specific memories were similar in magnitude for those with ($r = .36, p < .005$) and without a history of MDD ($r = .34, p < .0001$). Similarly, the Pearson correlations between the proportions of categoric memories were both statistically significant for those with ($r = .37, p < .005$) and without a history of MDD ($r = .28, p < .0001$).

coefficients for the personality trait of neuroticism for ages 1 through 30 based on 124 longitudinal samples. Of the available meta-analytic test-retest correlations, there were several estimates that closely resembled the late adolescent/young adult participant age range studied in the YEP and the particular test-retest intervals that were examined in the current investigation. These meta-analytic stability coefficients for neuroticism were as follows: $r = .55$ for measures of neuroticism between ages 16 and 20, $r = .33$ for measures between ages 17 and 20, $r = .42$ for measures between ages 17 and 21, $r = .22$ for measures between ages 17 and 24, $r = .56$ for measures between ages 18 and 21, $r = .55$ for measures between ages 18 and 22, and $r = .74$ for measures between ages 18 and 23. Our results for stability coefficients in the overall sample ($\rho = .31$ and $r = .35$ for the proportion of specific memories, and $\rho = .32$ and $r = .32$ for the proportion of categoric memories) certainly fall within the range of meta-analytic stability coefficients obtained for a personality trait in young adulthood. Thus, we believe that our findings can provide evidence that is consistent with the presence of a trait-like component to OGM.

Nevertheless, it is also important to consider influences that contribute to instability in measurements of OGM over time. For example, although there is little support for a reliable relationship between current depressive symptoms and OGM (e.g., Hermans et al., 2008), there is initial evidence that current mood state may influence OGM. In a sample of healthy volunteers, a negative mood induction—and decreases in happiness from before to after the induction—were associated with reductions in memory specificity from pre- to post-induction, even when covarying depressive symptoms at baseline (Au Yeung, Dalgleish, Golden, & Schartau, 2006). Although these results do not rule out the possibility of a trait-like component to OGM, they do indicate that current mood state can influence OGM at a given assessment. Subsequent research has also provided support for state influences on OGM, including various contextual factors (Debeer, Raes, Williams, & Hermans, 2011; Raes, Schoofs, Griffith, & Hermans, 2012). Future studies should thus take certain state influences into consideration as well as depression history when examining OGM and its stability over time. Additional research is also needed to better understand the extent to which factors that have been associated with memory performance, such as sleep and alcohol use (e.g., Nilsson, Bäckman, & Karlsson, 1989), might contribute to variation in OGM over time.

Furthermore, our study is not without limitations. First, we used different methodologies with the two AMTs (face-to-face administration with the first AMT vs. telephone administration with the second AMT) due to participants graduating from high school and leaving the Chicago or Los Angeles area for college and/or work. In addition, the two AMTs differed in the cue words used, both in terms of the words themselves and the number of cues. For example, as noted above, the cue words used in the second AMT were less concrete than those used in the first AMT. Greater cue word concreteness and imageability have been found to be associated with lower OGM (e.g., Williams, Healy, & Ellis, 1999), and this may have accounted, at least in part, for the differences in mean levels of OGM at the two assessments. Most studies of the test-retest reliability of the AMT have also used different sets of cue words (e.g., Peeters et al., 2002; Raes et al., 2009b), but this approach introduces an additional source of variability in AMT performance across the two administrations. Method variance is thus another factor that can contribute to observed instability in OGM over time, and using alternate forms of the AMT results in a conservative estimate of stability. It is likely that we would have observed higher stability in OGM if we had used the exact same AMT at both assessments. Indeed, evidence of the greatest stability in OGM comes from a study that used the same AMT at both time points (as well as one of the shortest test-retest intervals; Hermans et al., 2008). Another limitation of the current investigation is the lack of a measure of IQ. IQ has been found to be inversely

related to OGM (e.g., Wessel, Merckelbach, & Dekkers, 2002), and we were not able to take this relationship into account in this study.

OGM has been proposed as a marker for vulnerability to depression (Williams et al., 2007), and our results—taken together with others in the literature—suggest that there is an enduring component to this cognitive phenomenon. Such findings are consistent with the notion that OGM is not merely state-dependent, and they provide support for its promise as a target for prevention and intervention.

Acknowledgments

We would like to thank the National Institutes of Health for supporting our research (Grant# R01 MH065652 to Drs. Mineka and Zinbarg, R01 MH065651 to Dr. Craske, and F31 MH088014 to Jennifer Sumner). We also acknowledge the assistance of the many students who helped with data collection.

References

- Au Yeung C, Dalgleish T, Golden A, Schartau P. Reduced specificity of autobiographical memories following a negative mood induction. *Behaviour Research and Therapy*. 2006; 44:1481–1490.10.1016/j.brat.2005.10.011 [PubMed: 16356472]
- Coltheart M. The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*. 1981; 33A:497–505.10.1080/14640748108400805
- Debeer E, Raes F, Williams JMG, Hermans D. Context dependent activation of reduced autobiographical memory specificity as an avoidant coping style. *Emotion*. 2011; 11:1500–1506.10.1037/a0024535 [PubMed: 22142214]
- Eysenck, HJ.; Eysenck, SBG. *Manual of the Eysenck Personality Questionnaire (adult and junior)*. London: Hodder & Stoughton; 1975.
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. *Structured Clinical Interview for DSM-IV-TR Axis I disorders, research version, non-patient edition. (SCID-I/NP)*. New York: Biometrics Research, New York State Psychiatric Institute; 2002.
- Fraley RC, Roberts BW. Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*. 2005; 112:60–74.10.1037/0033-295X.112.1.60 [PubMed: 15631588]
- Griffith JW, Sumner JA, Debeer E, Raes F, Hermans D, Mineka S, Zinbarg RE, Craske MG. An item-response theory/confirmatory factor analysis of the Autobiographical Memory Test. *Memory*. 2009; 17:609–623.10.1080/09658210902939348 [PubMed: 19585345]
- Griffith JW, Sumner JA, Raes F, Barnhofer T, Debeer E, Hermans D. Current psychometric and methodological issues in the measurement of overgeneral autobiographical memory. *Journal of Behavior Therapy and Experimental Psychiatry*. 2011 Advanced online publication. 10.1016/j.jbtep.2011.05.008
- Hermans D, Vandromme H, Debeer E, Raes F, Demyttenaere K, Brunfaut E, Williams JMG. Overgeneral autobiographical memory predicts diagnostic status in depression. *Behaviour Research and Therapy*. 2008; 46:668–677.10.1016/j.brat.2008.01.018 [PubMed: 18342835]
- Heron J, Crane C, Gunnell D, Lewis G, Evans J, Williams JMG. 40,000 memories in young teenagers: Psychometric properties of the Autobiographical Memory Test in a UK cohort study. *Memory*. 2012; 20:300–320.10.1080/09658211.2012.656846 [PubMed: 22348421]
- Kleim B, Ehlers A. Reduced autobiographical memory specificity predicts depression and posttraumatic stress disorder after recent trauma. *Journal of Consulting and Clinical Psychology*. 2008; 76:231–242.10.1037/0022-006X.76.2.231 [PubMed: 18377120]
- Nilsson L, Bäckman L, Karlsson T. Priming and cued recall in elderly, alcohol intoxicated, and sleep deprived subjects: A case of functionally similar memory deficits. *Psychological Medicine*. 1989; 19:423–433.10.1017/S0033291700012460 [PubMed: 2762444]
- Peeters F, Wessel I, Merckelbach H, Boon-Vermeeren M. Autobiographical memory specificity and the course of major depressive disorder. *Comprehensive Psychiatry*. 2002; 43:344–350.10.1053/comp.2002.34635 [PubMed: 12216009]

- Raes F, Schoofs H, Griffith JW, Hermans D. Rumination relates to reduced autobiographical memory specificity in formerly depressed patients following a self-discrepancy challenge: The case of autobiographical memory specificity reactivity. *Journal of Behavior Therapy and Experimental Psychiatry*. 2012; 43:1002–1007.10.1016/j.jbtep.2012.03.003 [PubMed: 22561051]
- Raes F, Williams JMG, Hermans D. Reducing cognitive vulnerability to depression: A preliminary investigation of MEmory Specificity Training (MEST) in inpatients with depressive symptomatology. *Journal of Behavior Therapy and Experimental Psychiatry*. 2009a; 40:24–38.10.1016/j.jbtep.2008.03.001 [PubMed: 18407245]
- Raes, F.; Williams, JMG.; Hermans, D. On the test-retest reliability of the autobiographical memory test. In: Palfroft, LB.; Lopez, MV., editors. *Personality assessment: New research*. New York: Nova Science Publishers, Inc; 2009b. p. 391-397.
- Sumner JA. The mechanisms underlying overgeneral autobiographical memory: An evaluative review of evidence for the CaR-FA-X model. *Clinical Psychology Review*. 2012; 32:34–48.10.1016/j.cpr.2011.10.003 [PubMed: 22142837]
- Sumner JA, Griffith JW, Mineka S. Overgeneral autobiographical memory as a predictor of the course of depression: A meta-analysis. *Behaviour Research and Therapy*. 2010; 48:614–625.10.1016/j.brat.2010.03.013 [PubMed: 20399418]
- Sumner JA, Griffith JW, Mineka S, Rekart KN, Zinbarg RE, Craske MG. Overgeneral autobiographical memory and chronic interpersonal stress as predictors of the course of depression in adolescents. *Cognition and Emotion*. 2011; 25:183–192.10.1080/02699931003741566 [PubMed: 21432666]
- Sutton JM, Mineka S, Zinbarg RE, Craske MG, Griffith JW, Rose RD, Mor N. The relationships of personality and cognitive styles with symptoms of depression and anxiety. *Cognitive Therapy and Research*. 2011; 35:381–393.10.1007/s10608-010-9336-9 [PubMed: 21841850]
- Wessel I, Merckelback H, Dekkers T. Autobiographical memory specificity, intrusive memory, and general memory skills in Dutch-Indonesian survivors of the World War II era. *Journal of Traumatic Stress*. 2002; 15:227–234.10.1023/A:1015207428675 [PubMed: 12092915]
- Williams JMG, Barnhofer T, Crane C, Hermans D, Raes F, Watkins E, Dalgleish T. Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*. 2007; 133:122–148.10.1037/0033-2909.133.1.122 [PubMed: 17201573]
- Williams JMG, Broadbent K. Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*. 1986; 95:144–149.10.1037/0021-843X.95.2.144 [PubMed: 3711438]
- Williams JMG, Healy HG, Ellis NC. The effect of imageability and predictability of cues in autobiographical memory. *The Quarterly Journal of Experimental Psychology*. 1999; 52A:555–579.10.1080/027249899390963 [PubMed: 10504900]
- Zinbarg RE, Mineka S, Craske MG, Griffith JW, Sutton J, Rose RD, Waters AM. The Northwestern-UCLA youth emotion project: Associations of cognitive vulnerabilities, neuroticism and gender with past diagnoses of emotional disorders in adolescents. *Behaviour Research and Therapy*. 2010; 48:347–358.10.1016/j.brat.2009.12.008 [PubMed: 20070951]

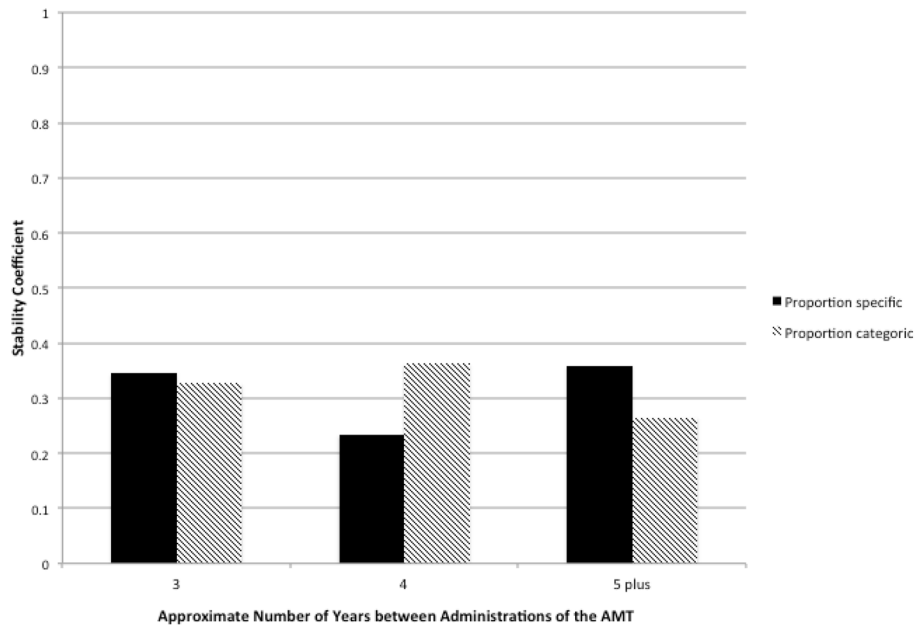


Figure 1. Spearman rank (ρ) stability coefficients for overgeneral autobiographical memory plotted as a function of the approximate length of the test-retest interval. The 3-year test-retest interval includes participants with 30–47 months between the two AMTs ($n = 47$), the 4-year test-retest interval includes participants with 48–59 months between the two AMTs ($n = 109$), and the 5 plus-year test-retest interval includes participants with 60–75 months between the two AMTs ($n = 115$). There were no significant differences between the stability coefficients for the different test-retest intervals, $ps > .31$.