Check for updates

Received: 26 June 2023 | Revised: 21 June 2024 | Accepted: 6 August 2024

DOI: 10.1002/acp.4240

RESEARCH ARTICLE

WILEY

How much trait variance is captured in autobiographical memory ratings?

Ezgi Aytürk[1] | Berivan Ece[2] | Nilüfer Göktaş[1] | Sami Gülgöz[1]

[1]Department of Psychology, Koç University, Istanbul, Turkey

[2]Department of Medical Social Sciences, Northwestern University, Evanston, Illinois, USA

Correspondence
Ezgi Aytürk, Department of Psychology, Koç University, Rumeli Feneri Yolu, Sarıyer, İstanbul 34450, Turkey.
Email: ezayturk@ku.edu.tr

Abstract

This study examined the trait-variance in autobiographical memory recollection (AMR). Participants ($N = 397$) provided ratings on eight autobiographical memories elicited using cue words. Multilevel random coefficients modeling revealed that 43% of the variance in memory ratings was due to individual differences, while the remaining 57% was related to other factors. Scores on the Brief Autobiographical Recollection Test (Brief ART; Berntsen et al.), a measure of metacognitive evaluations of autobiographical memory, were significantly correlated with mean memory ratings. However, they only explained 34% of the variance in mean ratings, and 15% of the variance in the ratings of individual memories, suggesting that metacognitive evaluations do not fully coincide with actual memory experiences. These findings suggest that AMR is a complex phenomenon influenced by both stable individual differences and situational factors. Multilevel modeling provides a rigorous approach to gaining a more nuanced understanding of AMR by disentangling various sources of variance.

KEYWORDS
autobiographical memory, individual differences, metacognition, multilevel modeling, random coefficients modeling

1 | INTRODUCTION

Autobiographical memory refers to "the retrieval and encoding of personal experiences and knowledge, including the cognitive and emotional components associated with those experiences" (American Psychological Association, 2020, p. 137). In other words, it contains our personal past experiences (e.g., Conway & Pleydell-Pearce, 2000). In autobiographical memory recollection (AMR) research, participants are usually instructed to remember personal memories and rate their recollections in terms of several event characteristics such as importance and vividness (e.g., Berntsen & Rubin, 2002; Rubin & Berntsen, 2009) using measurement instruments such as the Autobiographical Memory Test (AMT; Williams & Broadbent, 1986), Autobiographical Memory Interview (Kopelman et al., 1990), and the Autobiographical Interview (Levine et al., 2002). These studies focus on specific memories as the unit of analysis, and any dependencies among multiple memories of the same individual are treated as a nuisance. Recent research (e.g., Berntsen et al., 2019; Gehrt et al., 2022, 2023; Rubin, 2021) has suggested that AMR is a stable, trait-like individual difference variable, such that, some individuals tend to remember their past vividly in general while others do not, or some individuals are more likely to experience a sense of reliving or mental time travel while remembering a past event compared to others. As such, AMR may also be studied at the level of individuals.

Studies on AMR take different approaches in operationalizing the trait-AMR. The first approach involves obtaining multiple AMR ratings

Ezgi Aytürk and Berivan Ece should be considered joint first authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2024 The Author(s). *Applied Cognitive Psychology* published by John Wiley & Sons Ltd.

from participants, computing a "by-person" summary statistic (Murayama et al., 2014) such as within-person means or within-person average correlations among multiple memory ratings, and finally using these by-person summary values as measures of trait-AMR (e.g., Gehrt et al., 2022, 2023; Rubin, 2021). These trait-AMR measures are then used to explore whether there is a trait-like effect in AMR; or the relationship between individual differences in AMR and other variables of scientific interest. The second approach to studying individual differences in AMR is using measurement instruments that are developed to measure trait-AMR. The recently developed Autobiographical Recollection Test (ART; Berntsen et al., 2019) is one such instrument that has gained popularity. In the ART, individuals rate how their AMR is in general on seven dimensions such as level of vividness, coherence, and feeling of reliving. This second approach is less time-consuming and more efficient than having participants rate multiple autobiographical memories, provided that scores on the individual difference questionnaires such as the ART coincide with the by-person summaries of the ratings of actual memories (e.g., Gehrt et al., 2022, 2023).

In a recent study by Gehrt et al. (2023), participants completed the Danish version of the ART, reported three memories elicited with cue words, and rated them with a modified version of the Brief ART adapted for specific events. For each individual, a mean score was computed for each memory characteristic and then compared with the ART scores. The authors estimated pairwise correlations between ratings of each memory characteristic, averaged them and compared with the correlation between mean memory characteristics and the ART scores. The ART scores were positively correlated with aggregates of each memory characteristic. They concluded that Danish ART is a valid measure of trait-like characteristics in autobiographical memory. This was consistent with their earlier finding (Gehrt et al., 2022) that ART scores were positively correlated with aggregate ratings of characteristics of specific memories of participants as measured by the Autobiographical Memory Questionnaire (AMQ; Rubin et al., 2003), across various memory-eliciting conditions, with or without a one-week lag between the administration of the ART and memory ratings. Gehrt et al. (2022) suggested that individuals' scores on ART, indeed, are "reliably related to how they remember specific events", and that their analyses provide "evidence of the ART (Berntsen et al., 2019) as a trait-like measure of autobiographical memory characteristics that *can reliably related to how specific autobiographical memories are recollected and future events imagined*" (Gehrt et al., 2022, p. 93). This has led to claims that ART scores could be used in applied settings, such as legal contexts to assess eyewitness credibility or to evaluate age-related declines in memory (Gehrt et al., 2023).

Despite the potential benefits of assessing individual differences in AMR, the existing evidence base is limited in several important ways. First, there is a lack of consistent evidence that AMR is a stable, trait-like characteristic. Second, even if AMR is a stable individual difference, it is not clear that it can be reliably assessed using psychometric tools such as the ART scale.

2 | DISENTANGLING THE SOURCES OF VARIANCE IN AMR

The characteristics of a recollection experience are influenced by a multitude of factors, including both person-specific and situational factors (e.g., Rubin et al., 2004). Person-effects, also known as individual differences, in AMR, refer to the variability among individuals in certain memory-related aspects (e.g., Janssen et al., 2015; Vannucci et al., 2015). For instance, individuals may differ in their overall memory quality or in the strength of specific effects, such as the influence of event valence on memory quality. Situation-effects, on the other hand, are associated with variations stemming from event characteristics, such as the distinction between positive and negative events, traumatic experiences, recent versus remote events, memories of special occasions, and memories triggered by specific cue words (e.g., Berntsen & Rubin, 2002; Guler & Mackovichova, 2019). An additional layer of variance exists among memories of the same type for a given individual. For example, there can be variations in memory quality across events with a positive valence or in memory quality of events recalled in response to a particular cue word.

Experimental psychology primarily focuses on within-subject variance due to the event characteristics, and treats other sources of variance, such as individual differences as noise (Borsboom, Kievit, Cervone & Hood, 2009). The goal of these studies is often to generalize findings across a population of memories of the same memory type (e.g., Raes et al., 2003). Conversely, correlational psychology, which is concerned with individual differences, focuses on between-subject variance in AMR and treats within-subject variance due to event characteristics as noise (Borsboom et al., 2009). Individual differences research generally seeks to generalize findings to the population of individuals. Typically, a large sample of individuals provides a small number of memories (e.g., Gehrt et al., 2022; Rubin, 2021). It is also crucial to recognize that findings from one level of analysis cannot be directly extrapolated to the other. For instance, a significant within-person effect of an event characteristic (e.g., valence) on memory quality does not necessarily imply a similar effect between valence and memory quality across individuals (Hamaker, 2012).

In trait-AMR research, the variation among memories reported by the same participants is typically treated by creating aggregate scores. For example, participants are typically asked to report and rate several memories in response to a set of cue words (e.g., Gehrt et al., 2022, 2023), or those that belong to certain event characteristics (e.g., memory during a vacation, and with a family member; Rubin, 2021), or several positive and negative memories (e.g., Gehrt et al., 2022). The average rating for positive memories is then calculated and used as an individual difference measure of that individual for that particular event type. Similarly, averages of memories for negative events would be used as individual difference scores (e.g., Gehrt et al., 2022). This type of by-person analysis is known as the "means-as-outcomes" analysis (Wright, 1998). Another type of by-person analysis used in trait-AMR research involves estimating a correlation or regression coefficient among the ratings of multiple memories by each participant, individually, and treating this value as an individual

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

difference measure (e.g., Gehrt et al., 2022, 2023; Rubin, 2021). This kind of by-person analysis is known as "slopes-as-outcomes" analysis.

In both the means-as-outcomes and slopes-as-outcomes approaches, an often-overlooked problem arises when sampling variability of observations within individuals are ignored. This is called the *fixed-effect fallacy* (also known as the language-as-fixed-effect fallacy; Clark, 1973; Coleman, 1964; Yarkoni, 2022). Ignoring sampling error means that by-person means and slopes are treated as *perfect* estimates of the means or slopes of the population of memories of individuals (Nezlek & Plesko, 2001). When they are not, the findings apply only to the few memories that happened to be observed in a particular study, and no generalization can be made for the population of memories.

To illustrate the fixed-effect-fallacy, let us consider a hypothetical study where the researcher would like to explore the effect of neighborhood affluence on the math performance of elementary school students living in those districts. The researcher randomly samples 100 kids from five randomly picked districts of the same size, with varying levels of affluence. The mean scores on a standardized math exam are different, say, 104, 101, 100, 98, and 95, for the sample of kids from the five districts. In other words, the researcher observes a difference in math performance between the districts on the sample of kids who participated in his study. However, wanting to generalize to the population of elementary school students in these five districts, the researcher considers the sampling error, and draws a confidence interval (CI) around the sample means. For example, the 95% CIs of the district means are [90, 118], [91, 111], [90, 110], [95, 101], and [95, 105]. The 95% CIs of the district means overlap with each other; therefore, the researcher does not find enough evidence to reject the null hypothesis that there is no district effect on math ability for this set of five districts. The researcher would have made a fixed-effects fallacy if they claimed that the differences they observed in their sample estimates (means, correlations) generalize to the population of students in the five districts. If the researcher would also like to generalize the findings to some population of districts, then they need to account for the uncertainty in the sampling of the districts, and make inferences accordingly. In this case, the fixed-effects fallacy would have occurred if the researcher claimed that the findings generalize to the target population of districts without acknowledging the sampling error related to the districts included in the study.

By-participant summaries of multiple memory ratings reported by the same participant in correlational or experimental studies are considered a type of fixed-effects fallacy for similar reasons. In by-participant analysis, within-person means or correlations, or means of correlations, are treated as individual difference measures that are free from any sampling error. Consider three research participants, Alina, Leah, and David, whose average vividness ratings across seven cue-worded memories are 7, 6, and 5.5, respectively. These three respondents differ from each other in the seven memories they report in the study, but does this mean they differ from each other in their overall population of memories from which the seven are sampled? If the goal is to generalize to their target population of memories, as is the case for individual differences studies, then the level of uncertainty (due to sampling error) in sample means or correlations must be considered. These sample means or correlations represent random observations from their respective sampling distributions, such as the distribution of means of correlations among seven memory ratings. When the sample size is large, and the number of reported memories is small, as is common in recent trait-AM studies, the fixed-effects fallacy can lead to Type I error rates approaching 100%, and the parameter estimates biased by up to 300% (Yarkoni, 2022). This is because, by-person analysis assumes no effect of study design characteristics on observations (situational factors such as the cue words presented, particular rating scales used, time of the day, etc.), treating any observed variation as entirely attributable to the effect of primary interest (i.e., individual differences).

One historically recommended remedy to account for sampling error of observations within study participants was conducting *by-observation* analysis (e.g., Clark, 1973). This would be analogous, for instance, to comparing the mean vividness ratings of memories reported by Alina, Leah, and David in response to each of the seven specific cue words. Alternatively, one could compare the mean correlation between vividness and event age for memories reported for each of the seven cue words separately. Comparable means in the former case and comparable mean correlations in the latter would inform the extent to which the results can be generalized to the population of memories *for that particular group of respondents*. That is, the limitation of this approach is that it ignores the sampling variability of the respondents, and therefore, the findings cannot be generalized to the target population of individuals (Lorch & Myers, 1990). Any generalization claim would be considered a "participants-as-fixed-effect-fallacy."

Now, what type of evidence is needed to claim trait-like variance in AMR ratings that is generalizable to the target population of individuals and the target population of the memories within individuals? The type of empirical evidence required to suggest a strong trait-AMR is, first, that by-person means (or correlations) should be as distant from each other as possible (large between-subject variance), and/or by-person variances should be as low as possible (small within-person variance). The multilevel random coefficients modeling (e.g., Nezlek, 2007; Snijders & Bosker, 2011) accounts for the sampling error at different levels of the analysis simultaneously: (1) error related to sampling of the memories within individuals; and (2) error related to sampling of the subjects within a larger population of subjects to which the findings are to be generalized. Additionally, simultaneous analysis of memories and participants is more favorable than multi-step analysis inherent in by-person analysis, which requires multiple analytical decisions at each step (i.e., "researcher degrees of freedom", see Gelman & Loken, 2013; Silberzahn & Uhlmann, 2015), hindering the generalizability of the findings further.

Multilevel modeling has been used in the study of psychological processes which are characterized by within-person (state), and between-person (trait) variance components such as personality (e.g., Churchyard et al., 2014), positive and negative affect (e.g., Merz & Roesch, 2011), coping behavior in everyday life (e.g., Roesch et al., 2010), the experience of flow (Schmidt et al., 2014)

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and academic emotions (Nett et al., 2017). For AMR, Wright (1998) illustrated the advantage of acknowledging the random error related to memories within subjects, and the separate error related to sampling of the study participants.

3 | CURRENT STUDY

The aim of the present study was to investigate the extent to which the characteristics of how individuals remember autobiographical events are stable across different memories. We also explored how well Brief ART scores are related to the average of the actual characteristics individuals report in response to specific events remembered. For that aim, we collected participants' ratings of several memories on the seven dimensions captured in ART, such as vividness, coherence, and rehearsal. We also had them complete the Turkish version of the Brief ART. In other words, we gathered memory-specific ratings in addition to the ratings on Brief ART as a measure of autobiographical recollection in general.

We used multilevel random coefficients modeling to examine ratings for multiple memories and individual-difference scores at their proper levels of analysis. This allowed us to account for both the random error associated with the sampling of the participants (Level 2 error), and the random error associated with the sampling of specific memories (Level 1 error). Using multilevel modeling, we examined two research questions: (1) How much of the variability in the individuals' ratings on autobiographical memories is attributable to stable individual differences, and how much is due to the idiosyncratic effects of specific memories? (2) What proportion of stable individual differences in memory ratings is related to the Brief ART scores?

4 | METHOD

4.1 | Participants

A total of 398 participants took part in the study. One participant was excluded due to reporting only one valid memory. The final sample consisted of 397 adults aged between 18 and 62 years old (148 males; $M_{age} = 22.93$ years, $SD_{age} = 6.53$). Participants were either recruited from the subject pool of the first author's institution or contacted via email or social media platforms (i.e., Facebook and Instagram). Students recruited from the university's subject pool received course credit for their participation. Participants contacted via email or social media platforms, on the other hand, took part in the study voluntarily and were not compensated. A large sample size at the higher level of the hierarchy (Level 2) is important for multilevel model analyses (Snijders, 2005). Our sample size is well above the rule of thumb of 50 for unbiased estimates of regression coefficients and variance components (Maas & Hox, 2005).

4.2 | Materials

4.2.1 | Brief version of the autobiographical recollection test (brief ART)

The Autobiographical Recollection Test (ART; Berntsen et al., 2019) is a self-report measure that assesses individual differences in the experience of autobiographical recollection. It has seven dimensions: *reliving*, *rehearsal*, *vividness*, *visual imagery*, *scene*, *life-story relevance*, and *narrative coherence*. It consists of 21 items, three items for each dimension, and is rated on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). As individuals rate their own recollection experiences, ART scores also represent metacognitive processes. The short form of the ART, called the Brief ART, is a 7-item scale consisting of one item for each dimension. Brief ART scores can range from 7 to 49, with higher scores indicating stronger autobiographical memories. In the present study, we used the Turkish version of the Brief ART (Ece et al., 2021) to collect data on individual differences in autobiographical memory to determine the degree to which scores on ART, do indeed represent how well individuals actually experience AMR.

4.2.2 | Cue word memory task

Participants were presented with eight cue words. They were instructed to report the first memory that came to mind in response to each cue word. Once they reported autobiographical memories in response to all cue words, they were presented with each reported memory. They were then instructed to rate each reported memory on the seven characteristics assessed by the Brief ART (i.e., reliving, rehearsal, vividness, visual imagery, scene, life-story relevance, and narrative coherence). For this, we used a modified version of Brief ART in which statements refer to specific memories (e.g., Gehrt et al., 2022, 2023; Nielsen et al., 2022). For example, the item assessing the *reliving* dimension in the Brief ART (i.e., While remembering past events, it is as if I am reliving them.) was rephrased such that participants rated the reliving dimension specific to the memory they reported (i.e., While remembering *this event*, it is as if I am reliving it.). In the remainder of this paper, we will refer to scores from this version of Brief ART as cued-memory ratings. Like the original Brief ART, a cued-memory rating can range from 7 to 49, with higher values indicating better recollection *of a given memory elicited with a cue word*. In other words, the strength of cued-memory recollection was operationalized as the total score across the seven questions in Brief ART about each memory.

The cue words used in the study were selected from a pool of 40 words that were similar in imagery, concreteness, frequency, and association dimensions according to Turkish Word Norms (Tekcan & Göz, 2005). The selected cue words were *mattress*, *ferry*, *rose*, *document*, *soup*, *box*, *bell*, and *jacket*.

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.3 | Procedure

The present study was conducted online (Qualtrics, Provo, UT). First, participants were presented with information about the study and asked to provide informed consent. Once they provided informed consent, they responded to demographic questions (e.g., age and gender) followed by the Turkish Brief ART to assess their individual differences in autobiographical recollection. Then, they completed the cue word memory task by reporting a memory in response to each cue word they were presented with. Finally, they rated their memories using the modified version of Brief ART (i.e., cued-memory ratings) as described in the materials. Completion of the whole session took approximately 30 min.

4.4 | Data analysis

We first calculated descriptive statistics using the *psych* package (version 2.3.6; Revelle, 2023) in R (version 4.2.0; R Core Team, 2022). As discussed above, cued-memory ratings provided by participants for multiple autobiographical memories are multilevel or hierarchical in nature (i.e., memory ratings nested within individuals). Therefore, we analyzed these data using multi-level random coefficient models with the *lme4* package (version = 1.1.21; Bates et al., 2015) in R (R Core Team, 2022).

5 | RESULTS

5.1 | Descriptive statistics

A total of 397 participants reported and rated a total of 2837 memories. Each participant reported an average of 7.15 ($SD = 1.07$) autobiographical memories in response to eight cue words. The number of memories reported by each participant ranged from 3 to 8, and 78.8% of the participants reported at least seven cued memories and corresponding cued-memory ratings. Cued-memory composite ratings for individual memories across participants ranged from 7 to 49 ($M = 29.21$, $SD = 9.02$). Mean cued-memory ratings across memories ranged from 11.4 to 47.1 ($M = 29.20$, $SD = 6.48$).

Table 1 presents descriptive statistics for by-person averages and standard deviations of individual dimensions across memories. By-person means were largest for *Visual imagery* (4.87), and generally large for *Vividness*, *Coherence*, *Reliving*, and *Scene* (all above 4.30) compared to the dimensions of *Rehearsal* (2.86) and *Life story relevance* ($M = 2.73$). By-person means varied most for the *Vividness* dimension (SD = 0.51), and the least for *Rehearsal* (SD = 0.24).

By-person variation across memories was larger for *Vividness*, *Coherence*, *Reliving*, and *Scene* (average by-person *SD*s ≥1.40) than for *Rehearsal*, *Visual imagery*, and *Life story relevance* (average by-person *SD*s <1.30). By-person standard deviations varied most for the Life-story relevance (*SD* of by-person *SD*s = 0.61) and least for Reliving (0.53).

5.2 | The amount of trait-variance in autobiographical memory ratings

To examine the degree of within-person stability in memory ratings for each dimension, we fit a "totally unconditional" model:

$$y_{ij} = \beta_{0j} + r_{ij},$$

where β_{0j} is the random coefficient representing the mean of cued-memory ratings for persons *j* across the *i* cue-words for which each person provided a valid memory and memory rating, r_{ij} represents the within-participant random deviation (i.e., error) from β_{0j} for each cued-memory rating, and variance of r_{ij} represents the memory-level

TABLE 1 By-person descriptive statistics of aggregate ratings, and the variance components related to memory properties.

	By-person descriptive statistics							
	Within-person means across memories		Within-person *SD*s across memories		Parameters from the totally unconditional models			
	Mean	*SD*	Mean	*SD*	Grand mean (γ_{00})	Variance of within-person error (*r0j*)	Variance of between-person error (*u0j*)	Intraclass correlation (ICC)
Composite					29.20	46.67	34.93	0.43
Vividness	4.63	0.51	1.45	0.58	4.81	1.88	0.87	0.32
Coherence	4.31	0.44	1.43	0.56	4.47	1.91	0.95	0.33
Reliving	4.38	0.42	1.40	0.53	4.55	1.83	1.05	0.37
Rehearsal	2.86	0.24	1.29	0.60	2.93	1.90	1.14	0.38
Scene	4.46	0.35	1.42	0.57	4.60	1.87	1.07	0.37
Visual	4.87	0.42	1.32	0.60	5.05	1.49	0.87	0.37
Life story	2.73	0.25	1.25	0.61	2.79	1.86	1.14	0.38

Abbreviations: Composite, total rating; *SD*, standard deviation.

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

residual (or error) variance. The basic Level 2 (person-level) model was as follows:

$$\beta_{0j} = \gamma_{00} + u_{0j},$$

where γ_{00} represents the mean of the person-level means (i.e., grand mean of the cued-memory ratings), u_{0j} represents the deviation from γ_{00}, (i.e., error of β_{0j}), and the variance of u_{0j} represents the person-level (Level 2) residual variance.

In this totally unconditional model, we estimated model parameters as well as the intra-class correlation (ICC) coefficient. ICC informed the degree to which cued-memory ratings (Level 1 units) are dependent within each individual (Level 2 unit). ICC values can also be interpreted as the expected correlation between two randomly selected cued-memory ratings coming from the same participants (Hox, 2010; p. 15). ICC values can range from 0 to 1. A value of 1 indicates perfect dependence of Level 1 units within Level 2 units (Gelman & Hill, 2007, p. 258). In this context, it represents a situation where the cued-memory ratings of participants are identical across different events, and that, between-person variance in memory ratings is fully explained by individual differences. An ICC value of 0 indicates independence of Level 1 units within Level 2 units. In this context, it represents a situation where cued-memory ratings for different events are not correlated within individual participants and therefore there is no trait-like variance in memory ratings, and between-subject variance reflects only situation effects. Higher ICC values in this context suggest that cued-memory ratings reported for different events by the same participant are more similar compared to cued-memory ratings coming from different participants.

The model-estimated mean cued-memory recollection across participants (γ_{00}) was 29.20. The between-memory variance of cued-memory ratings (the variance of r_{ij}) was 46.67, and the between-person variance (the variance of u_{0j}) was 34.93. The ICC is calculated as $34.93/(34.93 + 46.67) = 0.43$, indicating that 43% of the variability in multiple cued-memory ratings is explained by stable individual differences, whereas the remaining 57% is attributable to the idiosyncratic effects of the cued memories (i.e., situation effects). In other words, even though cued-memory ratings coming from the same individual were more similar to one another compared to those coming from different individuals, there still existed a substantial memory-to-memory variability (i.e., situation effects) within individuals.

Table 1 presents the variance components as well as the ICCs for each dimension of autobiographical recollection. ICCs were similar across the memory dimensions, with Vividness having the lowest (0.32) and *Rehearsal* and *Life story* having the largest (0.38). These findings suggest that, at least 60% of variance in memory ratings on specific dimensions are due to the situation factors, and the rest is due to between-person trait-variance.

5.2.1 | Interim discussion

Our first research question was whether, and to what degree, autobiographical recollection has trait-like properties. We addressed this question by comparing within-subject to between-subject variability of memory characteristics by incorporating the sampling error related both the sample of individuals (Level 2 sampling error), and the set of memories sampled from among the population of memories for each individual (Level 1 sampling error). We found evidence for a stable trait-like experience of autobiographical recollection. Specifically, when composite scores were used as memory ratings, we found that 43% of the variability in how people rate their memories may be attributed to individual differences in autobiographical recollection. However, the remaining 57% was related to non-stable, situational factors (memory effects). The amount of trait-like variance was lower for individual dimensions, ranging from 32% (Vividness) to 38% (Rehearsal and Life story relevance). From these we concluded that, although individuals tend to rate their memories in a consistent manner, there is considerable variability that cannot be attributed to individual differences. This suggests that both individual differences and the characteristics of the specific memories play an important role in how memories are rated.

We used cue words to elicit random autobiographical memories which generally exclude memories that can be considered outliers (e.g., trauma memories), minimizing within-person variability in memory ratings. Therefore, the amount of variance we attributed to stable, trait-like tendencies in this study is perhaps an overestimation in the context of the population of all autobiographical events of the participants. Individuals would have varied more widely across their recollection of events had we sampled a more heterogeneous set of memories.

5.3 | How much of the trait variance is related to scores from the individual differences measure?

Next, we examined the relationship between an individual's Brief ART scores[1] (i.e., the individual differences measure) and that person's mean cued-memory ratings, accounting for the sampling error at both levels of the analysis. More specifically, we examined the extent to which individuals' scores on the individual differences questionnaire correspond to the trait-like variation that is captured in the sample of memory ratings they provide.

Brief ART scores of participants ranged from 11 to 49 ($M = 32.65$, $SD = 7.0$). We used a model in which mean cued-memory rating (β_{0j}) was modeled as a function of Brief ART scores. In other words, we included the Brief ART scores as a Level 2 predictor in the model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{Brief ART}) + u_{0j},$$

where γ_{01} is the regression coefficient for Brief ART scores. Using maximum likelihood estimation, γ_{01} and its significance test showed whether the mean cued-memory ratings and the Brief ART scores were related. The strength of this relationship was examined by comparing the random parameter estimates (i.e., residual variances) in the totally unconditional model and the model in which Brief ART scores are included as Level-2 predictors. The strength was operationalized

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

as the between-person variance in the mean cued-memory ratings that is accounted for by the Brief ART scores.

Next, we examined the second multilevel model where Brief ART scores were the Level 2 (person-level) predictor. This analysis informed whether and to what degree individuals' Brief ART scores are related to their mean cued-memory ratings. The intercept was 13.16 ($p < .01$) and the regression coefficient of the Brief ART scores (γ_{01}) was estimated to be 0.49, which was significantly different from 0 ($p < .001$). This means that for every 1.0-point increase in Brief ART scores, cued-memory ratings increased by 0.49 points. In other words, higher Brief ART scores were associated with higher average ratings of specific cued memories.

As for the strength of the relationship, we compared the residual variances of the two models. The residual variance of β_{0j} from the first analysis in which Brief ART scores (i.e., the person-level individual differences measure) were not included at the person level, was 34.93, and the residual person-level variance from the second analysis in which the Brief ART scores were included was 23.05, a reduction of 34%. This corresponds to a person-level correlation of 0.58 (the square root of 0.34) between *mean* cued-memory ratings and Brief ART scores. Put another way, Brief ART scores explained 34% of the stable individual differences in cued-memory ratings (which was 43% of the overall cued-memory rating variance). That is, 66% of the stable individual differences in cued-memory ratings were related to factors other than Brief ART scores. This means that the Brief ART score of a person was related to 15% (43% × 34%) of the variability in a given cued-memory rating of that individual (corresponds to a correlation of 0.38), where 28% (43% × 66%) of the variability was due to other *stable* individual differences. The rest of the 57% variability was due to factors specific to the reported event.[2]

Table 2 presents the intercept, the regression coefficient, and the amount of variance explained in by-person means and individual memories for each dimension. *Vividness* scores on the Brief ART were much stronger than other dimensions in predicting by-person means on vividness (variance explained: 41%), and individual memories (variance explained = 13%). Rehearsal scores on Brief ART predicted 20% variance in by-person means, and 7% variance in individual memories. Although all of them were statistically significant predictors of by-person mean ratings in their respective dimensions, Brief ART item scores explained <17% in by-person means, and <7% in individual memories.

6 | GENERAL DISCUSSION

The aim of the present study was twofold: First, we examined whether and to what degree the characteristics of autobiographical recollection represented stable individual differences across different memories. Second, we examined the degree to which Brief ART scores were related to these stable differences in autobiographical recollection. Although there is a number of recently published studies examining the reliability and stability of AMR as an individual differences construct (e.g., Gehrt et al., 2022; Rubin, 2021), we argue that these studies did not fully account for the nested structure of the data (i.e., ratings on several memories are nested within individuals).

Aggregates of ratings across memories for each participant were treated as individual-level variables in further analyses. It is not uncommon to use aggregates of lower-level data (memory level) to go one level up in the unit of analysis (individual level), the replicability and generalizability of this type of step-wise, by-person analysis may be severely compromised due to two reasons: first, there is an inherent flexibility, "researcher degrees of freedom" involved in the procedure of by-person analysis (e.g., estimating the correlation between the averages of the pairwise correlations among the ratings of memories across two testing occasions, sometime apart), second, the by-person summaries of sample of memory ratings are treated as fixed-effects, rather than random. Therefore, in this study, we collected multiple memory ratings from 397 individuals, and used the

TABLE 2 Coefficients and effect sizes of Level 2 (individual differences) predictors for each memory property.

	Model estimated parameters			Variance explained by Level 2 predictor		Variance components of individual memories		
Memory rating	Intercept	Regression coefficient of the Level 2 predictor[a]	Between-person error variance	Mean ratings across memories	Rating of a specific memory	*Stable* variance due to Brief ART scores	*Stable* variance due to other trait-factors	Situational factors
Composite	13.16	0.49	23.05	0.34	0.15	0.15	0.28	0.57
Vividness	3.09	0.37	0.61	0.41	0.13	0.13	0.19	0.68
Coherence	3.80	0.13	0.91	0.05	0.02	0.02	0.31	0.67
Reliving	3.62	0.20	0.96	0.10	0.04	0.04	0.34	0.62
Rehearsal	1.77	0.24	1.04	0.11	0.04	0.04	0.34	0.62
Scene	3.16	0.29	0.93	0.16	0.06	0.06	0.31	0.63
Visual	3.98	0.25	0.73	0.20	0.07	0.07	0.30	0.63
Life story	1.96	0.20	1.06	0.08	0.03	0.03	0.35	0.62

[a]Level 2 predictor of each memory rating dimension is their corresponding composite or item score in the Brief ART. Regression coefficients are all statistically significant at the alpha level of 0.001.

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

multilevel random coefficients modeling to address the trait-variance and situation-variance in AMR.

In the first set of analyses, we investigated the extent to which autobiographical recollection reflects stable, trait-like properties. By comparing within-subject to between-subject variability of memory characteristics, we found that 43% of the variability in composite memory ratings can be attributed to individual differences, suggesting a trait-like component in autobiographical recollection. Indeed, this magnitude is comparable to the amount of shared variance in scores of various cognitive tasks captured by the *g* factor, an individual difference factor of general intelligence (Kamphaus et al., 1997). However, the remaining 57% of the variability in memory ratings is related to situational factors, highlighting the influence of specific memories on memory ratings. In this study, we employed neutral cue words to elicit random autobiographical memories, potentially minimizing within-person variability and overestimating the trait-like variance. A more heterogeneous sample of memories might reveal wider variations across individuals' recollections. We also utilized a multilevel modeling approach to examine whether individual dimensions show different within- to between-subject variability ratios in memory ratings or whether there are different levels of concordance between the subscale scores of ART (e.g., *Vividness* subscale scores) and memory ratings (e.g., average *vividness* ratings across memories). We observed different levels of trait-like variance in the individual dimensions of AMR.

In a second set of analyses, we addressed our second research question of to what extent the scores from the Brief ART, a measure of metacognitive evaluations of the characteristics of autobiographical memory in general explain variance in the by-person means, as well as those in ratings of individual memories. The strength of the current study was that by modeling the random error associated with the selection of the memories (Level 1 residual), we accounted for any unreliability of the mean of the cued-memory ratings in representing individuals' true means. Results suggested that Brief ART scores were significantly related to the mean cued-memory ratings. This finding is partly consistent with earlier research that correlated average memory ratings with ART scores. The estimated correlation of 0.58 between individuals' Brief ART scores and within-person means of cued-memory ratings corresponds to a 34% overlap between their variances. As for the predictive power, we found that Brief ART scores can account for 15% of the variance of the ratings of individual memories. As expected, amount of variance explained by items tapping individual dimensions of AMR was generally lower.

6.1 | Limitations and future directions

Despite its strengths, the current study is not without its limitations. One limitation is the small number of autobiographical memories sampled from participants. Although eight memories are within the range that is typical in autobiographical memory research, the properties of eight memories may be insufficient to generalize to the population of memories of individuals. Further support is needed for the finding that 43% of recollection ratings is due to stable individual differences. The memories that were recalled in response to neutral cue words perhaps represent a rather homogeneous subgroup of autobiographical memories and therefore, Brief ART scores are related to an even smaller portion of the variance of a given autobiographical memory.

Another limitation is that the scale used for cued-memory ratings was a modified version of the Brief ART (i.e., statements in the Brief ART reworded for specific memories). Consequently, cued-memory ratings and Brief ART scores were based on almost identical statements. This language similarity may have caused response bias, which may account for some of the 34% shared variance. Had we used a different memory rating scale for cued memories, such as a subset of items from the AMQ (Rubin et al., 2003) capturing the same dimensions using different wording, the shared variance between the stable parts of the cued-memory ratings and the Brief ART scores might have been lower than 34%, and Brief ART scores might have explained less than 15% of a given cued-memory rating. An interesting finding was the 66% variance in stable parts of cued-memory ratings not related to the Brief ART scores. Future research may identify the factors explaining the individual differences in autobiographical recollection independent of individuals' beliefs about the characteristics of their autobiographical recollection in general (the construct tapped by the Brief ART).

Another limitation is that although we treated participants and means of memory ratings as random effects, we treated the stimuli (i.e., cue words) as fixed. Treating stimuli as fixed implies that cues are treated as replicates of each other, or that cues produce different groups of memories, but the single memory provided for each cue by each person is perfectly representative of the population of memories that would have been recalled in response to that cue in different occasions. We could have included a set of dummy variables in the multilevel models to estimate the effect of the cue words (Nezlek, 2007), but this would have increased estimated parameters which a larger sample size is required for. Further research is needed to explore the amount of variance in memory ratings that are due to the sampling of cue words. The generalizability of these findings should be examined by testing them in other cultures and other samples.

This study contributed to the AMR literature in two important ways. First, using multilevel random coefficient modeling, we demonstrated that less than half the variance in autobiographical recollection characteristics is due to stable individual differences, and these individual differences were only partially captured by the Brief ART. We conclude that individuals' metacognitive evaluations of their autobiographical recollections do not fully coincide with how they actually experience their memories. Instead of using ART scores as a proxy for individual differences in autobiographical memory, researchers interested in ruling out the stable individual differences in actual autobiographical recollection may opt for multilevel modeling.

To conclude, a comprehensive understanding of AMR requires careful consideration of the various sources of variance that contribute to the experience of AMR. Researchers should be aware of the limitations of existing methods and potential biases, such as the fixed-

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

effect fallacy, when analyzing hierarchically structured AMR data (i.e., memories nested within individuals). By employing rigorous methodological approaches that coincide with the complexity of AMR, a more nuanced understanding of this cognitive process can be achieved.

AUTHOR CONTRIBUTIONS

Ezgi Aytürk: Conceptualization; data curation; writing – review and editing; writing – original draft; methodology; formal analysis. **Berivan Ece:** Project administration; conceptualization; writing – original draft; writing – review and editing; methodology. **Nilüfer Göktaş:** Data curation; writing – review and editing; methodology; conceptualization. **Sami Gülgöz:** Conceptualization; writing – review and editing; supervision.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ezgi Aytürk https://orcid.org/0000-0001-6509-4288
Berivan Ece https://orcid.org/0000-0003-1756-9685
Nilüfer Göktaş https://orcid.org/0000-0002-3702-5370
Sami Gülgöz https://orcid.org/0000-0002-1262-2347

ENDNOTES

[1] The correlation between Brief ART and ART scores were 0.96. We opted for the Brief ART scores for this analysis.

[2] We run the same analysis with the Full ART scores as the Level-2 predictor. It explained 39% of the variance in by-person means. For individual memories, ART scores were related to 17% of the variance.

REFERENCES

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). RAPA. https://doi.org/10.1037/0000165-000

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berntsen, D., Hoyle, R. H., & 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*(3), 305–318. https://doi.org/10.1037/h0101839

Berntsen, D., & Rubin, D. C. (2002). Emotionally charged autobiographical memories across the life span: The recall of happy, sad, traumatic, and involuntary memories. *Psychology and Aging*, *17*(4), 636–652. https://doi.org/10.1037//0882-7974.17.4.636

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 67–97). Springer US. https://doi.org/10.1007/978-0-387-95922-1_4

Churchyard, J. S., Pine, K. J., Sharma, S., & Fletcher, B. (2014). Same traits, different variance: Item-level variation within personality measures. *SAGE Open*, *4*(1), 2158244014522634. https://doi.org/10.1177/2158244014522634

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359. https://doi.org/10.1016/S0022-5371(73)80014-3

Coleman, E. B. (1964). Generalizing to a language population. *Psychological Reports*, *14*(1), 219–226. https://doi.org/10.2466/pr0.1964.14.1.219

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261.

Ece, B., Aytürk, E., Göktaş, N., & Gülgöz, S. (2021). Factorial structure of autobiographical recollection assessed by a Turkish version of autobiographical recollection test (ART). *Current Psychology*, *42*(13), 10894–10909. https://doi.org/10.1007/s12144-021-02388-x

Gehrt, T. B., Nielsen, N. P., & Berntsen, D. (2023). Individual differences in autobiographical memory: A Danish version of the autobiographical recollection test (ART). *Nordic Psychology*, *1–18*, 116–133. https://doi.org/10.1080/19012276.2022.2162568

Gehrt, T. B., Nielsen, N. P., Hoyle, R. H., Rubin, D. C., & Berntsen, D. (2022). Individual differences in autobiographical memory: The autobiographical recollection test predicts ratings of specific memories across cueing conditions. *Journal of Applied Research in Memory and Cognition*, *11*(1), 85–96. https://doi.org/10.1037/h0101869

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*. Department of Statistics, Columbia University.

Guler, O. E., & Mackovichova, S. (2019). The role of executive function in autobiographical memory retrieval: Does the type of cue word matter? *Memory*, *27*(4), 423–430. https://doi.org/10.1080/09658211.2018.1513040

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In *Handbook of research methods for studying daily life* (pp. 43–61). The Guilford Press.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge.

Janssen, S. M. J., Kristo, G., Rouw, R., & Murre, J. M. J. (2015). The relation between verbal and visuospatial memory and autobiographical memory. *Consciousness and Cognition: An International Journal*, *31*, 12–23. https://doi.org/10.1016/j.concog.2014.10.001

Kamphaus, R. W., Winsor, A. P., Rowe, E. W., Kim, S., Flanagan, D. P., & Harrison, P. L. (1997). A history of intelligence test interpretation. In *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 32–47). The Guilford Press.

Kopelman, M. D., Wilson, B. A., & Baddeley, A. D. (1990). *The autobiographical memory interview*. Thames Valley Test Company.

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

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 149–157.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Merz, E. L., & Roesch, S. C. (2011). Modeling trait and state variation using multilevel factor analysis with PANAS daily diary data. *Journal of Research in Personality*, *45*(1), 2–9. https://doi.org/10.1016/j.jrp.2010.11.003

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1287–1306. https://doi.org/10.1037/a0036914

Nett, U. E., Bieg, M., & Keller, M. M. (2017). How much trait variance is captured by measures of academic state emotions? *European Journal of Psychological Assessment*, *33*(4), 239–255. https://doi.org/10.1027/1015-5759/a000416

Nezlek, J. B. (2007). A multilevel framework for understanding relationships among traits, states, situations and behaviours. *European Journal of Personality*, *21*(6), 789–810. https://doi.org/10.1002/per.640

Nezlek, J. B., & Plesko, R. M. (2001). Day-to-day relationships among self-concept clarity, self-esteem, daily events, and mood. *Personality and Social Psychology Bulletin*, *27*(2), 201–211. https://doi.org/10.1177/0146167201272006

Nielsen, N. P., Gehrt, T. B., & Berntsen, D. (2022). Individual differences in autobiographical memory predict memory confidence but not memory accuracy. *Journal of Applied Research in Memory and Cognition*, *12*, 542–551. https://doi.org/10.1037/mac0000082

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raes, F., Hermans, D., de Decker, A., Eelen, P., & Williams, J. M. G. (2003). Autobiographical memory specificity and affect regulation: An experimental approach. *Emotion*, *3*(2), 201–206. https://doi.org/10.1037/1528-3542.3.2.201

Revelle, W. (2023). Psych: Procedures for psychological, psychometric and personality research (R package version 2.3.6). https://CRAN.R-project.org/package=psych

Roesch, S. C., Aldridge, A. A., Stocking, S. N., Villodas, F., Leung, Q., Bartley, C. E., & Black, L. J. (2010). Multilevel factor analysis and structural equation modeling of daily diary coping data: Modeling trait and state variation. *Multivariate Behavioral Research*, *45*(5), 767–789. https://doi.org/10.1080/00273171.2010.519276

Rubin, D., Schrauf, R., & Greenberg, D. (2004). Stability in autobiographical memories. *Memory*, *12*(6), 715–721. https://doi.org/10.1080/09658210344000512

Rubin, D. C. (2021). Properties of autobiographical memories are reliable and stable individual differences. *Cognition*, *210*, 104583. https://doi.org/10.1016/j.cognition.2021.104583

Rubin, D. C., & Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the life span. *Memory & Cognition*, *37*(5), 679–688. https://doi.org/10.3758/37.5.679

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition*, *31*(6), 887–901. https://doi.org/10.3758/bf03196443

Schmidt, J. A., Shernoff, D. J., & Csikszentmihalyi, M. (2014). Individual and situational factors related to the experience of flow in adolescence. In M. Csikszentmihalyi (Ed.), *Applications of flow in human development and education: The collected works of Mihaly Csikszentmihalyi* (pp. 379–405). Springer. https://doi.org/10.1007/978-94-017-9094-9_20

Silberzahn, R., & Uhlmann, E. L. (2015). Many hands make tight work. *Nature*, *526*(7572), 189–191.

Snijders, T. (2005). Power and sample size in multilevel modeling. *Encyclopedia of Statistics in Behavioral Science*, *3*, 1570–1573. https://doi.org/10.1002/0470013192.bsa492

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis* (2nd ed.). Sage.

Tekcan, A. İ., & Göz, İ. (2005). *Türkçe kelime normları*. Boğaziçi Üniversitesi Yayınevi.

Vannucci, M., Pelagatti, C., Chiorri, C., & Mazzoni, G. (2015). Visual object imagery and autobiographical memory: Object imagers are better at remembering their personal past. *Memory*, *24*(4), 455–470. https://doi.org/10.1080/09658211.2015.1018277

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

Wright, D. B. (1998). Modelling clustered data in autobiographical memory research: The multilevel approach. *Applied Cognitive Psychology*, *12*(4), 339–357. https://doi.org/10.1002/(SICI)1099-0720(199808)12:4<339::AID-ACP571>3.0.CO;2-D

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, e1. https://doi.org/10.1017/S0140525X20001685

How to cite this article: Aytürk, E., Ece, B., Göktaş, N., & Gülgöz, S. (2024). How much trait variance is captured in autobiographical memory ratings? *Applied Cognitive Psychology*, *38*(5), e4240. https://doi.org/10.1002/acp.4240

10990720, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/acp.4240 by Koc University, Wiley Online Library on [27/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License