Differences in autobiographical memories reported using text and voice during everyday life

Emily Pearson[1], Jenna Graff[1], Elena Bai[1], Kelly Jakubowski[2], Amy M. Belfi[1]

[1]Department of Psychological Science, Missouri University of Science and Technology
[2]Department of Music, Durham University

Correspondence concerning this article should be addressed to Amy M. Belfi, Department of Psychological Science, 135 H-SS, 500 W. 14th St., Rolla, MO 65409, amybelfi@mst.edu

Abstract

Autobiographical memories frequently occur during everyday life. One of the most common approaches to measuring memories in everyday life is a diary method: Participants record memories as they occur by writing down these memories in a paper diary or typing them on a smartphone. Conversely, many laboratory-based studies of autobiographical memory require participants to describe their memories out loud in a spoken manner. Here, we sought to directly compare memories recorded via typing to those spoken out loud in a smartphone diary study. Participants reported everyday autobiographical memories that occurred over a period of four days either by typing (n=43) or recording themselves orally describing memories (n=39) using a smartphone app. Results indicated that the audio recording group reported memories more frequently and these memories contained a greater number of words, while the text group reported memories more promptly after they occurred. Additionally, the typing group reported memories that were episodically richer and contained a greater proportion of perceptual details. This work has important implications for future autobiographical memory studies in the lab, online, and using diary methods, and suggests that certain reporting modalities may be advantageous depending on the specific research focus.

Keywords: autobiographical memory; naturalistic; everyday life; involuntary

Autobiographical memories, memories for events in one's personal past, often occur during everyday life. For example, while shopping in the grocery store you might hear a strain from your wedding song, prompting a vivid image of dancing with your spouse, causing you to relive the sights, sounds, and emotions of that night. Such memories may be evoked by external sensory cues such as activities, people, sights, sounds, or smells, or internal cues such as thoughts, feelings, or emotions (for review, see Berntsen, 2021). Additionally, autobiographical memory retrieval happens quite frequently in everyday life (e.g., at least once per day, Schlagman & Kvavilashvili, 2008) and the memories retrieved tend to be highly accurate (Mace et al., 2011). Often, these memories occur in an involuntary manner, such that they are evoked with relatively little cognitive effort or intention to retrieve a memory (Berntsen, 1996, 1998).

One of the most common approaches to measuring autobiographical memories cued in everyday life is a self-directed diary approach: Participants record memories as they occur by writing down these memories in a paper diary or recording them on a smartphone (Berntsen, 1998; Jakubowski & Ghosh, 2021; Mace, 2004; Mace et al., 2011, 2015; Rasmussen et al., 2014; Schlagman & Kvavilashvili, 2008). Other methods include experimenter-prompted or experience sampling tasks, where participants are periodically prompted or cued, often via a text message, phone call, or a push notification, to record autobiographical memories at certain intervals during a period of time (Beaty et al., 2019; Rasmussen et al., 2014; Shiffman et al., 1999). In any studies of autobiographical memories occurring during everyday life, experimenters must make many decisions regarding the reporting modality, including the duration of the task (e.g., one day versus several days), the type of reporting device (e.g., smartphone vs. paper diary), and the frequency of prompting, if any.

Prior work has indicated that these variations in reporting methods can influence the data obtained in such tasks. Much of this work has focused on the relationship between reporting method and *frequency* of autobiographical memories cued in everyday life. For example, when participants are simply asked to indicate *when* a memory occurred, they report memories more frequently than when asked to provide additional information about each memory (Rasmussen et al., 2015). Frequency estimates of autobiographical memories cued in everyday life have ranged from 1-2 memories per day to over 20 per day, depending on the reporting method (Rasmussen & Berntsen, 2011; Schlagman & Kvavilashvili, 2008). One recent study compared autobiographical memories reported using a paper diary to those reported using a smartphone app (Laughland & Kvavilashvili, 2018). The authors sought to investigate differences in the frequency of memories reported and participant compliance. Perhaps counterintuitively, they found that participants using smartphones reported fewer memories than those using paper diaries, but were more compliant – that is, participants in the smartphone group made entries sooner after the memories occurred, kept the diary with them more frequently, and reported that it was easier to record memories than the paper diary group.

Although many components of autobiographical memory studies in everyday life can differ from one study to the next, one relative consistency is the fact that participants are typically asked to write (or, type) a description of the memory that occurred. However, it is unclear whether these written (or typed) memories reflect the actual recalled memory itself. That is, given the added burden of writing/typing out a memory description, participants might write a briefer description of the memory, omitting some details, rather than the full memory they experienced. In contrast, in laboratory-based autobiographical memory studies, participants often orally describe the memories that occur out loud, and the spoken descriptions are audio recorded

and subsequently transcribed (e.g., Belfi et al., 2016; Belfi et al., 2018; Sheldon et al., 2020; Wardell et al., 2021), although in some cases laboratory-based studies of involuntary autobiographical memories have asked participants to type a description of the memories into a computer (e.g., Barzykowski & Niedźwieńska, 2018; Vannucci et al., 2014). It is plausible that spoken descriptions may be a more accurate reflection of the true memory retrieval process rather than brief descriptions written down in a diary or typed into a computer or smartphone app. For example, individuals may engage in more post-retrieval monitoring (e.g., elaboration or monitoring of the retrieval process to ensure accuracy of the memories) when they are asked to report memories via speech versus text (Barzykowski et al., 2021). Alternatively, individuals may be better able to organize their description of the memory in a written format. In sum, written versus spoken narrative descriptions may differ in their content and/or phenomenological characteristics of the memory experience. This potential difference also has important implications as increasing numbers of psychological studies are being conducted online: Online studies nearly always require participants to type their memory descriptions rather than describe them using speech (Jakubowski et al., 2021; Janata et al., 2007; Janssen et al., 2011).

While prior studies have not directly compared written versus spoken descriptions of autobiographical memories, there are some hints that these two reporting modalities may result in memories that differ in various ways. For example, in an online study of music-evoked autobiographical memories, where participants typed their memory descriptions, the word count of the memory descriptions was relatively low (average word count per memory: 15.46; Jakubowski et al., 2021). Word counts were slightly higher in a paper diary study of music-evoked autobiographical memories, where participants hand wrote memory descriptions (average word count per memory: 30.90; Jakubowski & Ghosh, 2021). In comparison, autobiographical

memories evoked by music in a laboratory study, where participants orally described their memories, contained even more words (average word count per memory: 82.68, Belfi et al., 2020). It therefore might be the case that describing a memory using speech leads an individual to further elaborate and retrieve additional details that may not occur when typing or writing a response by hand. The fact that spoken memories contained more words might reflect further elaboration, which may then be associated with differences in the phenomenological experience of the memory retrieval process; for example, the elaboration phase of autobiographical memory retrieval has been associated with visual imagery and a sense of reliving (Daselaar et al., 2008). In sum, there is currently tentative evidence to suggest that the modality of reporting autobiographical memories may affect both the memory content and the experience of recall.

To address the question of whether spoken memories differ in their qualities from typed memories, we conducted the present study, in which we directly compare memories recorded via speech versus typing. In this study, participants reported autobiographical memories that occurred during everyday life using a smartphone app. The key manipulation in this study was that participants were assigned to either 1) record their memories by *typing* the memory descriptions into an app or 2) audio recording themselves describing the memory by speaking into the app. Importantly, recent research has indicated that there may be differences between "objective" analyses of memory content and subjective ratings of memory experience (Barzykowski et al., 2021), as well as differences between automated and manual methods of memory content analysis (Belfi et al., 2020). In the present study we analyzed the content of the memories using both automated (i.e., the Linguistic Inquiry and Word Count) and manual methods (i.e., the Autobiographical Interview coding method) and assessed subjective experience of memory recall using phenomenological ratings. In sum, the goals of this study

were to assess whether typed versus spoken descriptions of memories occurring in everyday life differed in terms of their reporting characteristics (e.g., number of words, frequency of memories recorded), content (e.g., types of words used, episodic richness), and phenomenological experience (e.g., ratings of emotional valence, perceived vividness, etc.). We predicted that memories recorded via speech would contain a greater number of words, would be more episodically detailed, and would be associated with a richer phenomenological experience, including increased ratings of emotional valence and vividness.

Methods

Participants

Our target sample size was determined by an a priori power analysis conducted using G*Power software (Faul et al., 2007). We based our estimated effect size on a prior study which compared the frequency of autobiographical memories recorded using a paper diary to those recorded using a smartphone app (Laughland & Kvavilashvili, 2018). In this study (which included roughly 30 participants per group), the authors report a large effect size of η^2_p=0.22 when comparing the number of autobiographical memories evoked between these two conditions. Using this reported effect size, we conducted an a priori power analysis, which resulted in a total sample size of 50 (i.e., 25 per group). Given that we expected participant attrition due to the somewhat longitudinal nature of the study, we aimed to recruit slightly more participants than this - our target was 40 participants per group.

Participants consisted of 88 undergraduate students from Missouri University of Science and Technology who participated in the study for research credit as a part of the requirements for their coursework. Participants were assigned to one of two groups: In the "text" group (n=44), participants were instructed to type their memory responses into the app. In the "audio" group

(n=44), participants were instructed to audio record themselves describing the memories that occurred. Data from one participant in the text group and five participants in the audio group were excluded due to not following instructions properly (i.e., participants in the audio group who were excluded did not audio record memories but typed them out, as there was still a text box on the app), leaving a total of 43 participants in the text group and 39 participants in the audio group. Participants in the text group (M=28, F=15) were an average of 19.83 years old (SD=2.49), and participants in the audio group (M=23, F=16) were an average of 20.17 years old (SD=3.23); there was no significant difference in age between the two groups [t(80)=0.53, p=0.59].

Procedure

All procedures were conducted in compliance with the American Psychological Association Ethical Principles and were approved by the Institutional Review Board at the University of Missouri (IRB project number: 2047763). Before participating, participants read a brief general description of the study and were instructed that they must have an Android phone or iPhone and be willing to download the ExpiWell app in order to complete the study. The app used here, ExpiWell, was created for the purpose of conducting this type of experiment (expiwell.com). Upon agreeing to participate, participants then gave informed consent by typing their name into a text box in the app. Next, participants were provided a link to an online slideshow that explained how to download, log into, and use the Expiwell app. Participants first saw the following instructions:

This is a study about autobiographical memories and what cues them. An autobiographical memory is:

A memory of a personal experience from your past, which may include details about events, people, places, and time periods from your life. (i.e. Senior prom, first day of school).

In this study we will ask you to keep track of all autobiographical memories as they occur during your daily life, and record information about these memories and what cued them on an app.

Next, they were shown a series of slides indicating how to log in to the app and complete the informed consent. Next, they were given the following instructions:

*Remember, an **autobiographical memory** is a memory of a personal experience (i.e. getting your driver's license, Christmas morning). A **cue** is what triggers the memory to happen (i.e. a song, smell, taste, conversation, etc.). Memories can happen at any time. When a memory occurs, log it in the ExpiWell app.*

*Please record your memories for **four days**. You may begin on any day, but only record memories for **four days** once you begin. For example, if you begin logging on 10/1 you will record memories until 10/4.*

Following this, participants were shown a series of slides indicating how to log their memories and complete the questionnaire for each memory. Participants in the text group were instructed to type their memories into a text box. Participants in the audio group were instructed to select the "audio recording" option and describe their memories while the app was recording their speech.

To summarize: Participants recorded any autobiographical memories that occurred during a period of four days, using the Expiwell app on their phones. Participants were not prompted to respond at any specific times; instead, they were asked to record memories on the app as they happened in everyday life.

Memory Questionnaire.

For each memory, participants were asked twelve questions. First, participants were asked for today's date (*current date*). Next, participants recorded the time in which their entry was completed (*current time*). Participants were then asked to record the time that the memory retrieval occurred (*memory time*). Since participants could not always fill out the survey at the

moment the memory retrieval occurred, we expected the current time to sometimes differ from the time the memory retrieval occurred.

Next, participants were asked to describe what cued the memory (*cue type*). This question was a free response and participants could give any answer they wished with as much or as little detail as they wished (see **Cue Categorization** below for details on how the cues were categorized). Participants were then asked to categorize the level of effort/intention required to retrieve the memory (*retrieval effort*). Participants were then asked to describe the memory that was recalled in as much detail as possible *(memory description).* In the text group, participants were shown a large text box in which they could type their answers. Text box entries were not submitted until participants tapped a "submit" button; thus, participants had the option to type out their descriptions, read them back and revise, and submit when they were satisfied with their descriptions. Similarly, in the audio group, participants were able to record themselves talking, listen to the recording, and re-record as many times as they wished. There were no time limits imposed on the audio recordings or character limits imposed on the text entries. Due to the app, data on the number of revisions (both text and audio recording) were not able to be collected. After describing the memory, participants were asked to report the age at which the original event happened. This was a free response question (*memory age*). Participants were asked to rate the specificity of the memory, that is, whether it was a single event (lasting more or less than 24 hours), a repeated event, or a lifetime period *(memory specificity).* Participants then reported how frequently they have thought of this memory before on a 5-point Likert scale *(rehearsal frequency).* Finally, participants rated the memory vividness *(memory vividness)*, emotional valence (*memory valence*), and importance of the memory to their life story (*memory importance*), each on 5-point Likert scales. For the full questionnaire, please see **Appendix A.**

Data Quantification

We sought to quantify the content of the memories reported using two methods: the Linguistic Inquiry and Word Count (LIWC) and the Autobiographical Interview. Spoken memory descriptions from the audio group were first transcribed, word-for-word, by the researchers before being submitted to the following quantification methods.

Linguistic Inquiry and Word Count.

The LIWC categorizes words into various groups, ranging from parts of speech to affective dimensions of a text, and counts the number of words present in a text and calculates a percentage of words in each category (number of words in a category divided by the total number of words; Pennebaker et al., 2015; Tausczik & Pennebaker, 2010). To limit the number of statistical tests (the LIWC outputs hundreds of variables), we selected a subset of these for our analysis. First, we included word count (WC) to identify whether participants in the audio or text groups produced more words overall. Next, we selected the most 'higher order' categories reflecting the *content* of the text in terms of psychological processes. These included the variables of affective processes (e.g., "happy, cried"), social processes (e.g., "talk", "they"), cognitive processes (e.g., "cause", "know"), and perceptual processes (e.g., "see", "listen").

Autobiographical Interview.

Memory descriptions were also coded using the Autobiographical Interview method (Levine et al., 2002). Each memory was segmented into details (single pieces of information) that were coded as either internal or external. Internal details pertain to the central memory and reflect episodic reexperiencing, and include details about the event (e.g., actions, happenings), time (e.g., year, season, day), place (e.g., city, building, room), perceptions (e.g., auditory, tactile, visual details), and thoughts or emotions (e.g., emotional states). External details do not

directly pertain to the memory and primarily reflect semantic content (e.g., general knowledge or facts), but also can include external events (e.g., details from other unrelated incidents), repetitions (e.g., repeating details already stated), or metacognitive statements. Memories were coded by a single trained rater who was blind to the nature of the research hypothesis. This rater was trained using the training materials provided by the authors of the Autobiographical Interview (Levine et al., 2002). Briefly, this training consists of coding 25 training memories and comparing one's ratings to a set of expert coders in an iterative fashion, until the new rater establishes high reliability with the set of expert raters (e.g., correlations above 0.75).

After coding each detail, internal and external composite scores were created by counting the total number of internal and external details for each memory. The internal and external composite scores were then used to calculate a ratio of internal/total details. This ratio provides a measure of episodic detail that is unbiased by the total number of details (Levine et al., 2002). We included both the composite internal and external scores, as well as the ratio of internal/total details, in our subsequent statistical analyses. At the conclusion of the study, the rater re-rated 5% of the memories in the study to assess test-retest reliability of their ratings. Intraclass correlation coefficients (ICCs) were calculated using a two-way mixed model, once for the internal composite score and once for the external composite score. The ICC value for external details was 0.87 and for internal details was 0.92; these values are very similar to those in our prior work which used multiple raters (Belfi et al., 2016).

Cue Categorization.

In addition to reporting autobiographical memories that occurred, participants were also asked to describe what cued their memories. We sought to identify the most frequent memory cues and whether cues differed between the audio and text groups. To do this, we developed a set

of rules for coding each description as one of several cue types. While prior research using similar methods has had participants select the cue type from a list (e.g., Berntsen, 1996; Mace, 2004), here, we sought to use a data-driven approach. Therefore, we chose our cue categories based on the data itself. To this end, two raters read through all the written descriptions of the cues (for examples, see the table in Appendix B – e.g., "Talking with friends."). They first went through and separately categorized each memory into one of a broad set of categories identified in prior work: The initial categories consisted of "action", "interaction", "date", "music", "words", "audiovisual", "visual", "smell/taste", and "other" if the memory did not fit any of the above categories. If there were any times in which the raters did not agree, they discussed the memory and came to a consensus as to which category it belonged. If the rater designated a memory as "other," then they provided a more specific sub-category based on their reading of the description and alerted the other rater of this new sub-category. If there were three or more memories in any of the subcategories within "other," a new category was created. This process resulted in the final categories: *Activity/Action*, *Interaction*, *Sensory Stimuli*, and *Other*. Within the broad categories of *Sensory Stimuli* and *Other*, there were several subcategories. For *Sensory Stimuli*, the subcategories are *Music, Audiovisual, Visual, Smell/Taste, and Sensory Stimuli - Other*. We also created a general *Other* category for cues that were not sensory-related but did not fall into any other categories (i.e., categories for which there were not at least 3 examples were included in "other"). The categories are described in **Appendix B**.

Analysis

First, we sought to assess whether one group reported a greater number of memories, and whether these memories contained a greater number of words, than the other group. Additionally, we sought to assess whether one group took more time between the memory

occurring and recording the memory than the other group. To do this, we conducted three

independent-samples t-tests to evaluate the effect of group on the number of memories reported

(per participant), the word count of the memories (averaged across all memories within a

participant), and the time difference between memory retrieval and when the memory was

logged (also averaged across all memories within a participant). T-tests were calculated using the

t_test function of the rstatix package in R (Kassambara, 2020). Next, we sought to analyze the

content of the memories. We first did this by conducting a MANOVA on the LIWC data, with

four dependent variables from the LIWC output (affective, social, cognitive, and perceptual

processes), and group (audio, text) as the independent variable. Next, we conducted a similar

MANOVA on the AI coding, with three dependent variables (internal details, external details,

ratio of internal/total details) and group as the independent variable. All variables were first

averaged across all memories within each participant. We averaged within participants to control

for potential differences in the number of memories between groups. Additionally, we analyzed

all data a second time using linear mixed-effects models, which do not require averaging within

participants and can account for participant-by-participant variability using random slopes. The

outcomes of both analyses (MANOVA and mixed effects models) were nearly identical. We

have included details about the mixed effects model analyses and results of these analyses in the

Supplemental Materials.

To assess differences in the subjective experience of the memories reported using voice

versus text, we conducted a similar MANOVA on the subjective ratings, with six dependent

variables (retrieval effort, memory specificity, rehearsal frequency, memory vividness, memory

valence, and memory importance) and group as the independent variable. All MANOVAs were

conducted using the manova function in R. Additionally, we sought to investigate whether these

subjective ratings were at all related to the episodic richness of the memory descriptions, and whether this varied by group. To this end, we conducted a binomial regression using the six subjective ratings variables, as well as group (audio, text), and the interaction between each rating variable and group, as predictors in the model. The ratio of internal/total details from the AI coding was used as the outcome variable. The data were not averaged for each participant in this model. Therefore, the goal of this model was to see whether subjective experience of recalling these memories was related to the episodic richness of the memories, and whether this interacted with the reporting modality.

Finally, we sought to investigate whether certain types of memory cues were more frequently reported overall, and whether the distribution of cue types differed between the audio versus text groups. To this end, we conducted two z-tests for the equality of proportions (using the prop.test function in R) to assess whether 1) there were differences in the cue types on the whole and 2) the proportions of the eight cue types differed between the text and audio groups.

<div align="center">

Results

</div>

Reporting Characteristics: Memory Frequency, Length, and Timing

When comparing the number of memories reported by each participant, there was a significant difference between groups, $t(80)=2.66$, $p=0.009$, $d=0.58$, 95% CI:[0.29, 2.08]. The audio group reported significantly more autobiographical memories (M=4.21, SD=2.33) than the text group (M=3.12, SD=1.71). When comparing the number of words used in their memory descriptions, there was also a significant difference between groups, $t(80)=6.98$, $p<0.001$, $d=1.54$, 95% CI:[41.99, 75.47]. The memories from the audio group contained a significantly larger number of words (M=97.8, SD=50.3) than the memories from the text group (M=39.0, SD=21.6). We removed two apparent outliers (see Figure 1B) and re-ran our analyses, and the

results were still statistically significant. When comparing the time between when the memories were retrieved and when they were logged, there was no difference between groups, $t(80)=1.26$, $p=0.21$, 95% CI:[-15.5, 69.4]. See **Figure 1** for a graphical depiction of the data. Given the skewed nature of the time data (i.e., many datapoints clustered at 0), we investigated group differences in timing a second, complementary manner, following methods from prior research. As in previous work (Laughland & Kvavilashvili, 2018), we calculated the proportion of memories that were logged within 10 minutes of the memory retrieval, for each group. We then conducted a chi-square test, which indicated that there was a significantly greater proportion of memories logged within 10 minutes of the memory occurring for the text group (93/134 total memories, 69%) than the audio group (84/168 total memories, 50%), $\chi^2=10.78$, $p=0.001$.
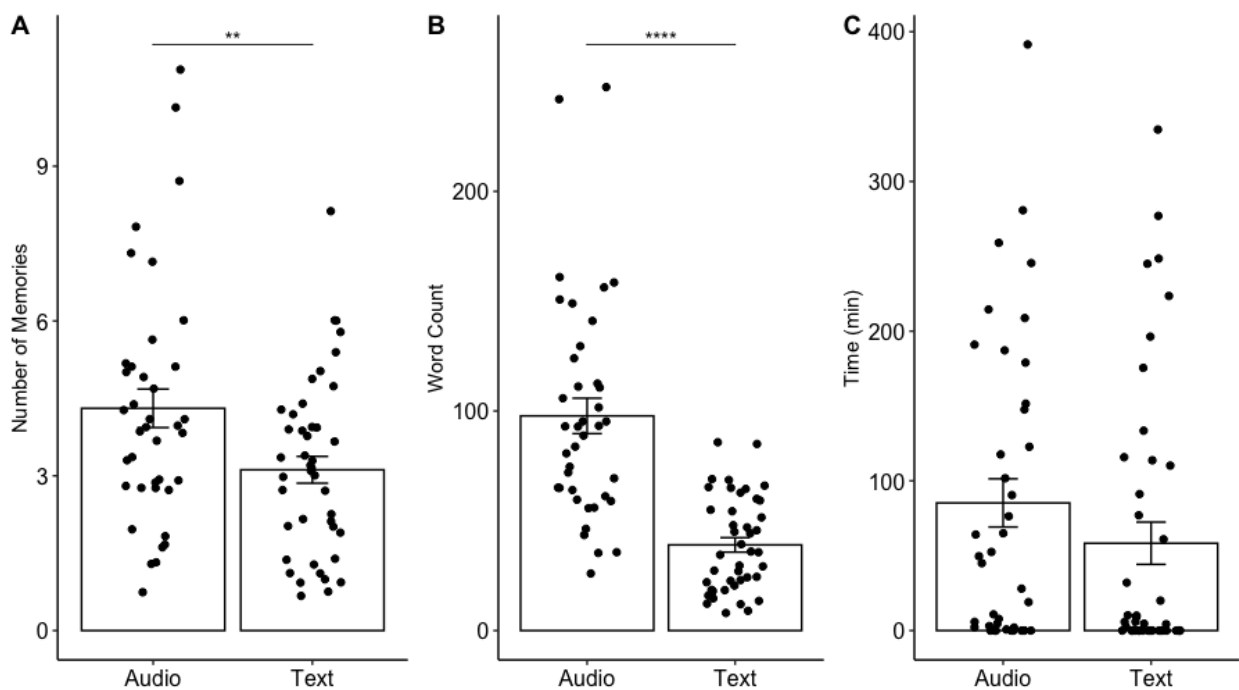

Figure 1. A) Number of memories **B)** Word count **C)** Time between memory retrieval and when the memory was logged. Bar graphs indicate mean, error bars indicate standard error of the mean, individual points indicate individual subjects. ** $p<0.01$, ****$p<0.001$

Memory Content

The MANOVA assessing the effect of group (audio, text) on the LIWC variables (affective, social, cognitive, and perceptual processes) was significant, $F(4,77)=3.96$, $p=0.005$, $\eta^2_p=0.17$. Post-hoc univariate ANOVAs for each dependent variable revealed significant effects of group on cognitive ($F(1,80)=7.52$, $p=0.007$) and perceptual processes ($F(1,80)=7.39$, $p=0.008$), but not affective ($F(1,80)=2.70$, $p=0.10$) or social processes ($F(1,80)=1.21$, $p=0.27$). Post-hoc pairwise comparisons, Bonferroni corrected for multiple comparisons, conducted following the significant univariate ANOVAs revealed that the audio group produced memories that contained a significantly larger proportion of words reflecting cognitive processes than the text group, $t(80)=2.74$, $p=0.007$, 95% CI: [0.68, 4.29]. In contrast, the text group produced memories that contained a significantly larger proportion of words reflecting perceptual processes than the audio group, $t(80)=-2.71$, $p=0.008$, 95% CI:[-3.62, -0.56]. We removed one apparent outlier from the perceptual processes LIWC scores (see Figure 2D) and re-ran our analyses, and the results were unchanged. See **Figure 2** for a graphical depiction of the data.
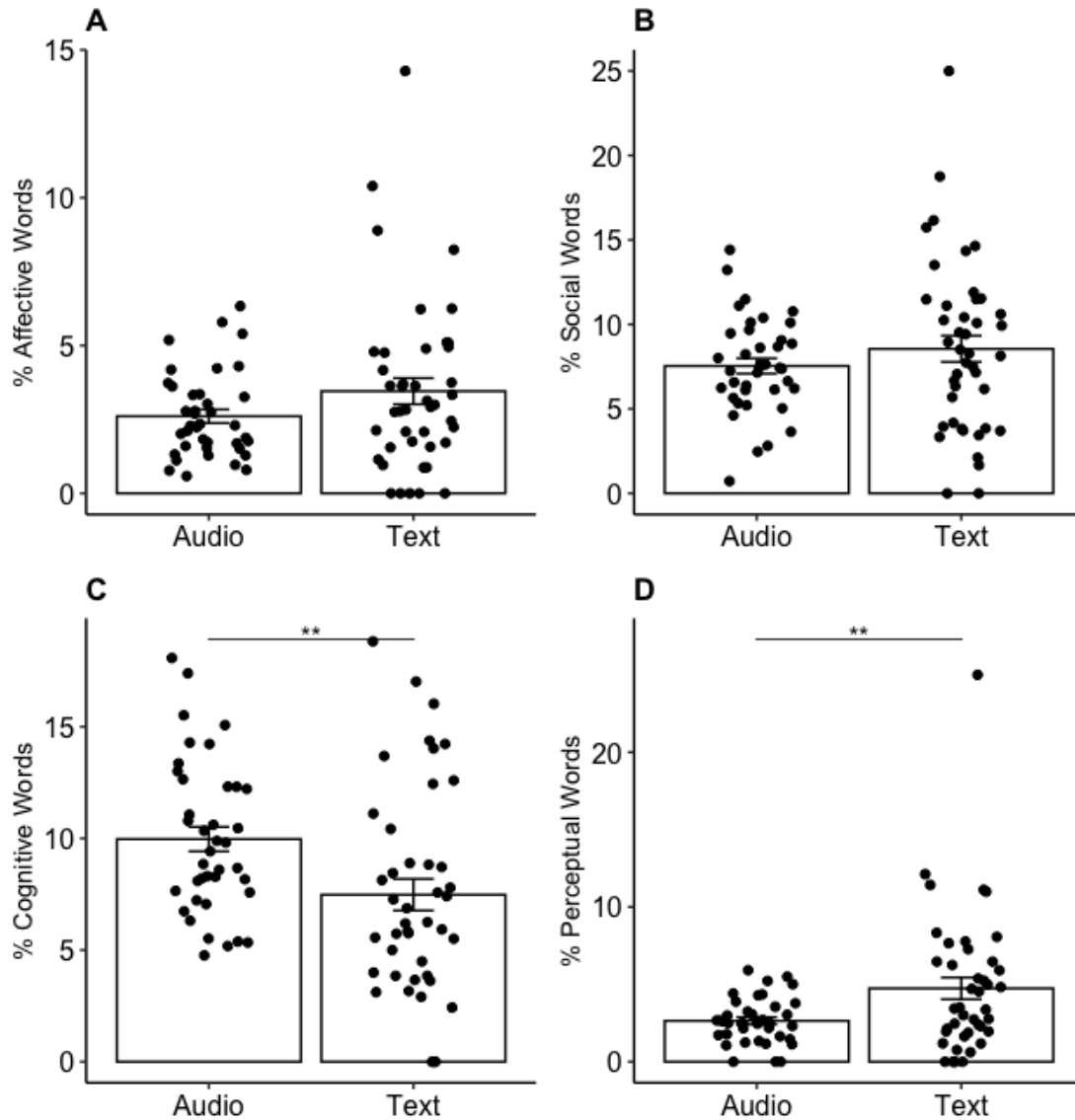
Figure 2. LIWC Data. **A)** Affective words **B)** Social words **C)** Cognitive words **D)** Perceptual words. Bar graphs indicate mean, error bars indicate standard error of the mean, individual points indicate individual subjects. Y-axes depict the percentage of words that fall into each category (averaged across all memories for each participant). ** *p<0.01*

The MANOVA assessing the effect of group (audio, text) on the AI variables (internal details, external details, and ratio of internal/total details) was significant, $F(3,78)=19.88$, $p<0.001$, $\eta^2_p=0.43$. Post-hoc univariate ANOVAs for each dependent variable revealed significant effects of group on the number of internal details ($F(1,80)=5.46$, $p=0.02$), number of external details ($F(1,80)=33.93$, $p<0.001$), and the ratio of internal/total details ($F(1,80)=34.39$,

p<0.0001). Post-hoc pairwise comparisons, Bonferroni corrected for multiple comparisons, conducted following the significant univariate ANOVAs revealed that the audio group produced memories that contained a significantly greater number of internal details than the text group, *t*(80)=2.33, *p*=0.02, 95% CI: [0.35, 4.35]. The audio group also produced memories that contained a significantly greater number of external details than the text group, *t*(80)=5.82, *p*<0.001, 95% CI:[3.43, 6.99]. In contrast, the text group produced memories that had a significantly higher ratio of internal/total details than the audio group, *t*(80)=-5.86, *p*<0.001, 95% CI:[-0.30, -0.15]. See **Figure 3** for a graphical depiction of the data.

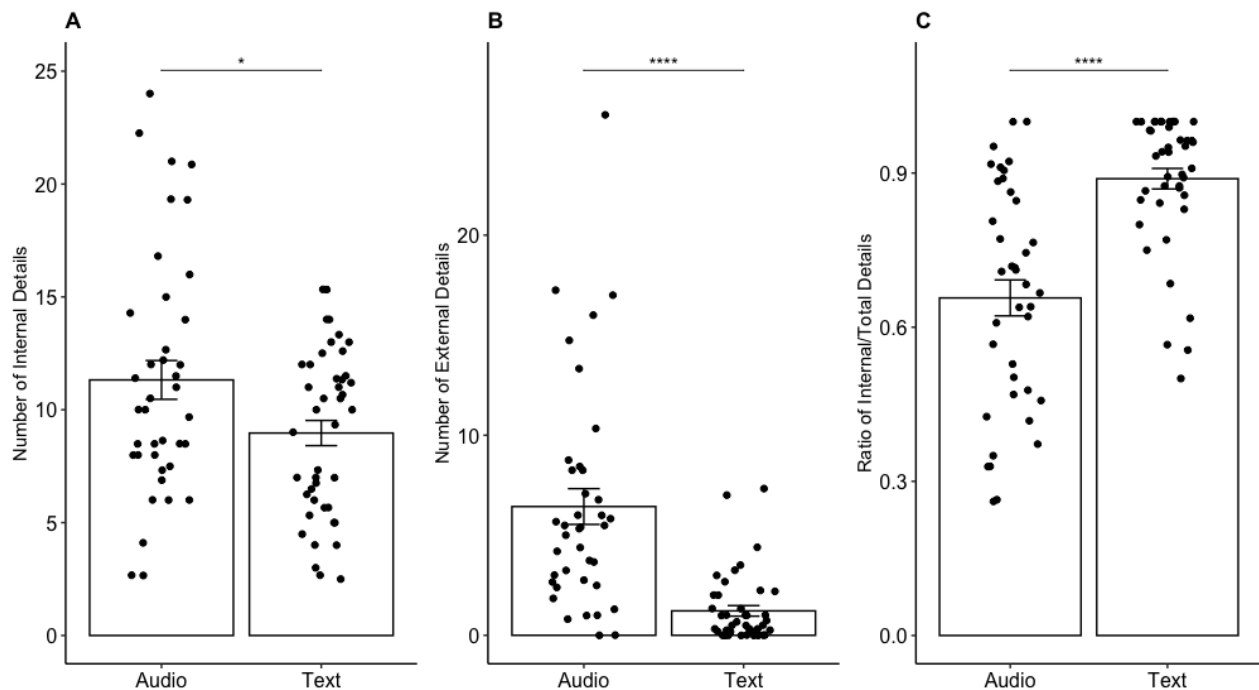**Figure 3.** AI Data. **A)** Internal details, **B)** External details, **C)** Internal/total ratio. Bar graphs indicate mean, error bars indicate standard error of the mean, individual points indicate individual subjects. Y-axis for internal and external details indicate composite scores, while the y-axis for the ratio of internal/total indicates the ratio (averaged across all memories within each participant). * *p*<0.05, ****p*<0.001

Subjective Ratings of Memory Experience

The MANOVA assessing the effect of group (audio, text) on the subjective rating variables (memory specificity, rehearsal frequency, memory vividness, memory valence, and memory importance) was not significant, $F(5,76)=0.08$, $p=0.99$, $\eta^2_p=0.007$. Post-hoc univariate ANOVAs also indicated no significant differences between groups for any of the five variables (all ps > 0.60). To investigate whether the two groups differed in retrieval effort, we calculated the number of responses to each of the three points on the retrieval effort scale (this was not included in the MANOVA as it is a categorical variable). Most participants reported that the memories were evoked in a spontaneous manner (76% of the audio group; 78% of the text group), as opposed to requiring deliberate effort (21% of the audio group, 16% of the text group). A small portion of the memories were rated as "not sure" on this scale (3% of the audio group, 6% of the text group). We then computed a chi-squared test for differences in the proportion of responses to these scale items between the two groups, which revealed no significant differences ($X^2(2)=1.67$, $p=0.43$).

We then conducted a binomial regression with the six subjective rating variables (including retrieval effort as a categorical predictor), group (audio vs. text), and their interactions as predictors, with the ratio of internal/total details from the AI coding as the outcome variable. The purpose of this analysis was to assess whether the subjective ratings had any relationship with episodic richness (as measured by the internal/total ratio), and whether this interacted with the reporting modality. The overall model was significant, $X^2(13)=426.59$, $p<0.001$. See **Table 1** for the full results of the model.

Table 1. Regression results.

	β	S.E.	z	p	Sig.
Intercept	0.63	0.05	13.6	0.002	**
Group	1.36	0.11	12.45	<0.001	***
Retrieval Effort	-0.07	0.1	-0.71	0.47	
Memory Specificity	-0.32	0.04	-8.12	<0.001	***
Rehearsal Frequency	0.12	0.05	2.62	0.008	**
Memory Vividness	0.12	0.04	2.79	0.005	**
Memory Valence	0.12	0.04	2.88	0.003	**
Memory Importance	-0.3	0.05	-6.2	<0.001	***
Group * Effort	-0.24	0.24	-1	0.32	
Group * Specificity	0.06	0.09	0.6	0.55	
Group * Frequency	-0.05	0.1	-0.49	0.62	
Group * Vividness	-0.34	0.1	-3.33	<0.001	***
Group * Valence	-0.33	0.1	-3.28	0.001	**
Group * Importance	0.3	0.11	2.66	0.007	**

*Note: * p<0.05, ** p<0.01, ** p<0.001*

Replicating our ANOVA results, there was a significant main effect of Group on the ratio of internal/total details, such that the text group had a higher ratio than the audio group. When looking at the subjective rating predictors, there were several significant effects: there was a positive relationship between *memory specificity* and internal/total details, such that more specific memories contained a higher ratio (memory specificity was coded as 1 for the most specific memories and 4 for the least specific). There were also positive relationships between *rehearsal frequency, memory vividness,* and *memory valence* with the internal/total ratio. In contrast, there was a negative relationship between *memory importance* and the internal/total ratio. Additionally, there were some significant interactions between these subjective ratings and group (audio vs. text): the significant interaction between group and *vividness* indicated that, for the audio group, memory vividness and the internal/total ratio were positively correlated, while the reverse effect was shown in the text group. A similar interaction was shown for *valence*: in the audio group, there was a positive relationship between valence and internal/total ratio, while there was a negative effect in the text group. For memory importance, the audio group showed a

negative relationship between importance and internal/total ratio, with no clear relationship in the text group. The results of all significant effects and interactions are plotted in **Figure 4** (main effects are not plotted when there was also a significant interaction involving that variable).

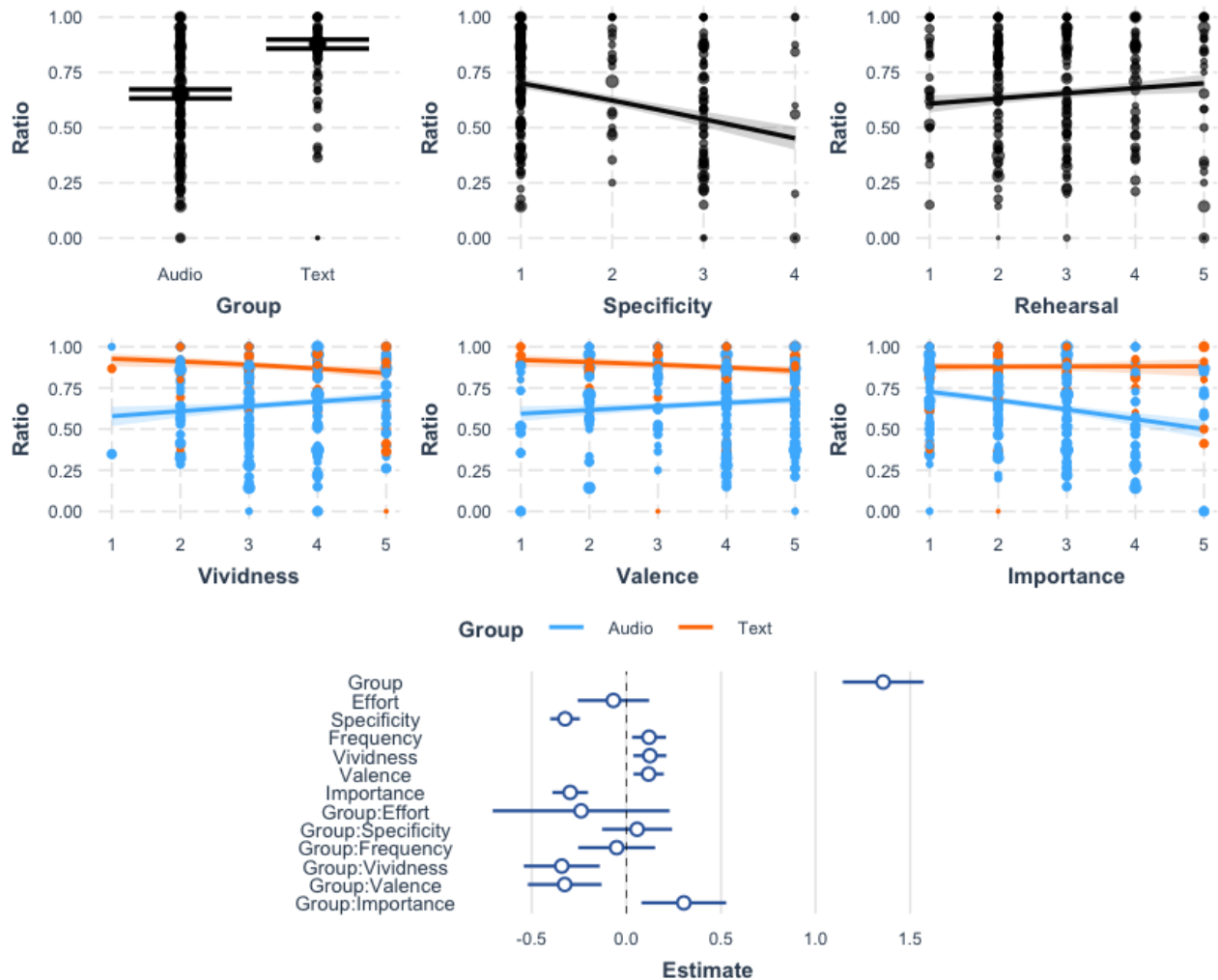Figure 4. Regression results. The top row illustrates significant main effects of group, memory specificity, and rehearsal frequency. The middle row illustrates significant interactions with group and the subjective rating variables: memory vividness, valence, and importance. The bottom row illustrates the beta estimates and 95% confidence intervals for the estimates for each of the predictors in the model.

Cue Types

First, the proportion test to assess whether the distribution of cue types (action, audiovisual, interaction, music, other non-sensory cue, other sensory cue, smell/taste, and visual) differed overall was significant, $X^2(7)=40.61$, $p<0.001$. Post-hoc pairwise proportion tests, Bonferroni corrected for multiple comparisons, indicated that there were significantly fewer "other sensory" cues than "action" ($p=0.001$), "audiovisual" ($p<0.001$), "interaction" ($p<0.001$), "music" ($p=0.03$) and "smell/taste" cues ($p=0.001$). Additionally, there were significantly fewer "visual" cues than "action" ($p=0.006$), "audiovisual" ($p<0.001$), "interaction" ($p=0.004$), and "smell/taste" ($p=0.009$). There were no other differences between cue types. That is, "other sensory" and "visual" cues were the least common, and there were no differences in the frequency of other types of cues (see bar heights in **Figure 5** for a graphical depiction of the number of trials per cue). Secondly, we sought to investigate whether the proportion of cue types reported differed between the audio and text groups. This proportion test was not significant, $X^2(7)=13.29$, $p=0.07$, indicating that the proportion of cue types was not different between groups (**Figure 5**).
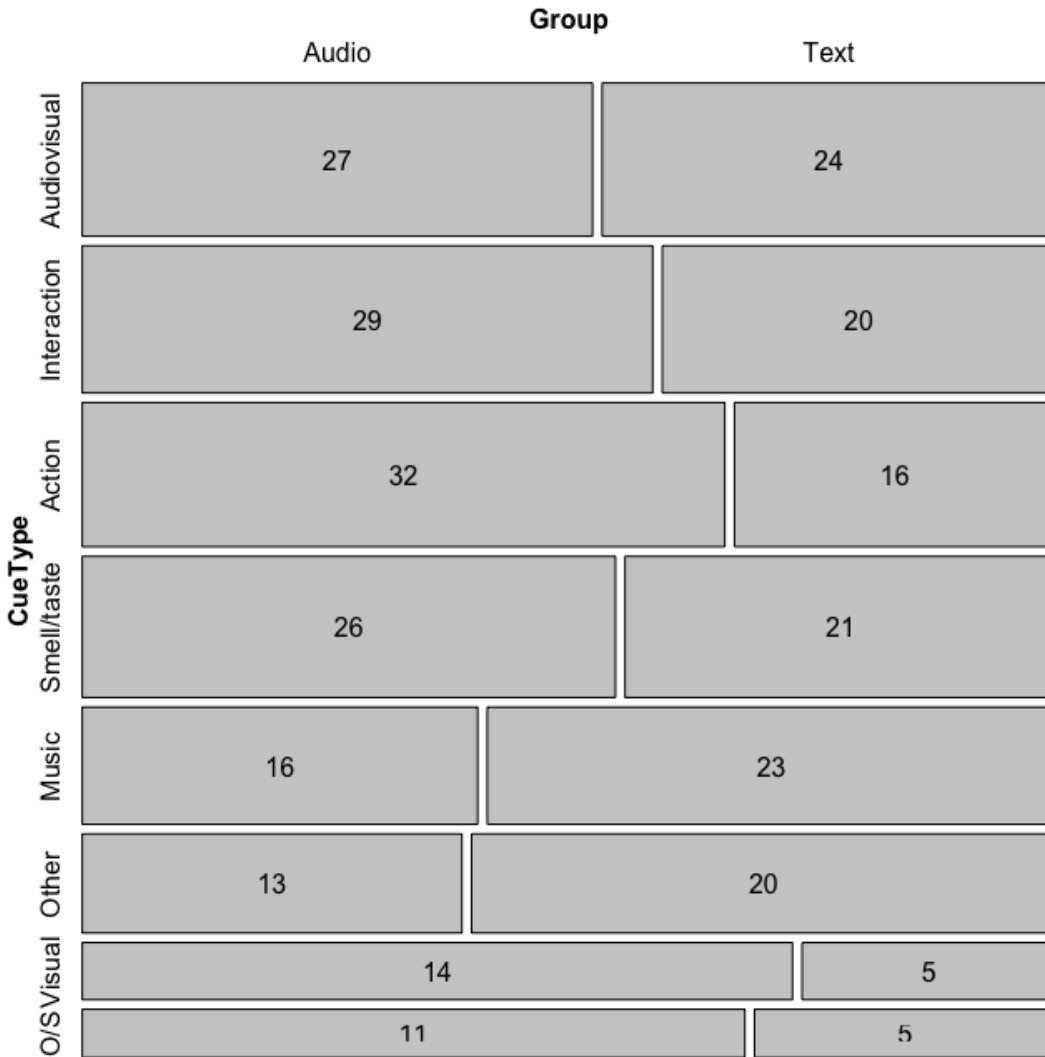
Figure 5. Proportion of cue types for audio and text groups. Numbers indicate the total number of memories reported for each cue type per group. Bar heights illustrate the proportion of total memory reports by cue type (irrespective of group). Cue categories are also ordered top to bottom from the category with the most reported memories to the least. Bar widths illustrate the proportion of memory reports by group (within cue type). O/S = "other sensory."

Discussion

In the present study, we sought to investigate differences between typed versus spoken autobiographical memories that occurred during everyday life. We hypothesized that memories recorded via speech would differ from typed memories in terms of their reporting characteristics (i.e., would contain more words), memory content (i.e., would be more episodically rich), and

phenomenological experience (i.e., would be rated as more emotional and vivid). Our first prediction was supported: Participants in the audio group reported memories that contained significantly more words than participants in the text group. Additionally, participants in the audio group reported significantly *more* memories than those in the text group. This difference in frequency of responses suggests one possible benefit of using spoken recordings as a reporting method – it is possible that describing a memory by speaking is seen as easier or less cumbersome than typing, and this may lead participants to report more memories. An additional factor that might have influenced this result is that perhaps the audio recording group might be more likely to play with the app, as it might be a more engaging way to use the app. This could also have contributed to the higher number of responses in the audio group.

These results also seem to align with prior work comparing memories recorded using a paper diary to those using a smartphone (Laughland & Kvavilashvili, 2018), which found that smartphone diaries are also accompanied by a higher compliance rate. Our results expand on this by suggesting that providing memories via speech may encourage an even greater response rate than by typing. On the other hand, we found that participants in the text group more frequently reported a memory within 10 minutes of it occurring, as compared to the audio group. It may be the case that audio recording can only be done in certain circumstances – for example, if a memory occurs while sitting in class, on public transportation, or in other scenarios where it might be difficult to make a recording. The benefits and drawbacks of a higher frequency of reporting in the audio group, but more prompt reporting in the text group, are important to consider when designing future studies.

However, in contrast to our second hypothesis, memories reported by the audio group were *not* more episodically rich than those reported by the text group. That is, despite containing

significantly more words, these additional words did not seem to add to the episodic nature of the memories. The memories reported by the audio group contained a significantly lower ratio of internal/total details than the text group, suggesting that the additional words were simply semantic statements, repetitions, or other external details, rather than providing additional episodic detail to the memory descriptions. Indeed, when observing the composite scores for internal and external details, it can be seen that participants in the text group display a floor effect for external details. That is, individuals in the text group tended to produce fewer external details than individuals in the audio group. For example, a characteristic audio-recorded memory was:

> I was playing IM flag football and it kind of brought me back in time because it was the first time that I played flag football since like 8th grade. It was the first time in a number of years that I actually played football. It's been a long time. I used to play football all the time. And I used to have really good hands but, you know, that's a memory in itself. But just playing flag football you know, it kind of brought me back to just playing it in eighth grade and how much fun I had playing and it was a pretty hard memory to just, you know, ignore like that. And it was still a pretty fun time. And it just reminded me a lot of, I guess you could say, the quote unquote glory days for me.

From the above memory, it can be seen that the participant provided some episodic content (e.g., remembering playing football in 8th grade), but that this memory description contained several repetitions and semantic statements. In contrast, an example of a characteristic text memory is below. Notably, this memory contains several perceptual details about the sights and emotions felt on that night, as well as specific content about the event itself:

I recalled the memory of my first date with my now fiancé, I vividly remembered her smiling really big at me and that she wore a blue top, had her hair up, and was wearing her glasses. She was very nervous and I asked her why and she said she wasn't even though she was, I remember I was wearing a black button up with dress shoes and suit pants and we talked about the state of the country of Palestine and our family heritage all night.

While we were unable to collect data on whether or how both groups edited their memories, it might be the case that participants in the text group were more likely to go back and edit out erroneous information in their descriptions, which may be influencing this result. This could be an interesting direction for future research (i.e., whether participants are more likely to revise descriptions using text rather than voice). A related possible explanation could be that the text group may be more likely to engage in post-retrieval monitoring, such that they more vigilantly monitor and revise their memory descriptions during the elaboration phase (Barzykowski et al., 2021). Assessing the frequency of revisions to their memory descriptions in addition to what types of changes are made could address this possible explanation for the between-group differences shown here.

In addition to differences in episodic richness, there were also differences between the audio and text groups as measured by the LIWC. The audio group reported memories that contained more details reflecting "cognitive processes" (e.g., words like "think" and "know") while the text group reported memories that contained more details reflecting "perceptual processes" (e.g., words like "see" and "feel"). When interpreting these results in the context of the Autobiographical Interview results discussed above, it seems that the audio group may have produced more "cognitive" words as a part of metacognitive statements (which are included as a

type of external detail). Similarly, words categorized as "perceptual" processes by the LIWC would likely be considered internal details using the Autobiographical Interview. Therefore, the LIWC results seem to complement the autobiographical interview results by providing additional support for the finding that memories reported via typing contain a greater proportion of vivid, episodically-rich details.

Finally, our third prediction was also unsupported by our data: An additional surprising finding was that there were no differences in the phenomenological experience of autobiographical memories reported using typing versus speech. We had predicted that the process of orally describing the memories may have caused participants to recall more vivid or emotional details and would potentially affect their subjective experience of recalling the events. However, our data did not support this prediction and suggest few differences in the experience of recalling and reporting memories in either the spoken versus typed modality. Furthermore, the proportions of different cue types were not different between the two reporting modalities. These results suggest that either method of reporting could be used without impacting the subjective experience of the participants' recall.

This finding points to the potential importance of analyzing the content of memory reports in addition to the phenomenological characteristics. Our data illustrate that while the phenomenological features of recalling a memory may not differ between the two types of reporting modalities, the memories themselves do. Content analyses may provide a more "objective" measure which can complement self-report ratings of memory experience (Barzykowski et al., 2021). For example, one could measure the proportion of perceptual words in a memory report using the LIWC and the proportion of perceptual internal details using the Autobiographical Interview as measure of memory vividness, which could complement and

extend phenomenological ratings of vividness. Looking at memory content can, for example,

also provide a more nuanced insight into what types of perceptual details might be driving

subjective ratings of vividness. Furthermore, using both an automated method (the LIWC) and a

manual method (the Autobiographical Interview) provides additional complementary measures

of the constructs of interest. That is, while in the present study, LIWC and AI methods tended to

converge, in prior work they have not always precisely aligned (Belfi et al., 2020). Therefore,

both methods can both provide additional information that a single measure alone may lack.

Of course, our study is not without limitations. For one, our sample consisted of US

undergraduate students, who are highly familiar with and likely use smartphones differently than

other groups might. For example, our findings cannot be generalized to older adults, who may

prefer a more traditional paper diary. However, it is possible that using the speech reporting

modality might be preferable for older adults, those less familiar with technology, or those less

able to type on a phone. Therefore, it may be a fruitful avenue for future research to explore the

ease of entry for various types of reporting methods in an older adult population. As there is a

large body of research on differences in autobiographical memory retrieval between younger and

older adults (Ford et al., 2021; Levine et al., 2002; Piolino et al., 2010; Wank et al., 2021), this

could be a beneficial method for assessing which reporting modality best works for capturing

such memories as they occur in everyday life.

To conclude, in the past few years a large proportion of psychological research has

shifted online, with some prior work suggesting that online data is of equal or higher quality to

that collected in the lab (Horton et al., 2011; Peer et al., 2017). When it comes to studying

autobiographical memories, there is one key difference in the methodological approaches taken

by online versus lab studies: That is, online studies typically ask participants to type descriptions

on their computer, while many lab-based studies rely on participants' spoken reports of memories. Here, we directly compared memories reported via typing to those recorded via speech. Perhaps surprisingly, our data indicate that the typed descriptions were more vivid (i.e., contained more perceptual details) and episodically rich (i.e., contained a greater proportion of episodic content) than those recorded using audio. Therefore, there may be both logistical and theoretical advantages to using a typed response modality (although spoken descriptions may be easier for certain groups of participants). However, the spoken modality may reduce burden on participants, as we saw an increase in the number of memories reported for this group. Finally, the two modalities may provide differing insights into the memory retrieval process – while spoken descriptions may be more "stream of consciousness" and illustrative of the mental process of retrieval, typed descriptions may be more concise and focused. In sum, our results add to the body of work suggesting that response modality can have substantial influences on the frequency and quality of autobiographical memories reported and suggest that response modality be taken into account when designing future studies.

References

Barzykowski, K., & Niedźwieńska, A. (2018). Priming involuntary autobiographical memories in the lab. *Memory*, *26*(2), 227–289. https://doi.org/10.1080/09658211.2017.1353102

Barzykowski, K., Skopicz-Radkiewicz, E., Kabut, R., Staugaard, S. R., & Mazzoni, G. (2021). Intention and Monitoring Influence the Content of Memory Reports. *Psychological Reports*, *0*(0), 1–27. https://doi.org/10.1177/00332941211048736

Beaty, R. E., Seli, P., & Schacter, D. L. (2019). Thinking about the past and future in daily life: an experience sampling study of individual differences in mental time travel. *Psychological Research*, *83*(4), 805–816. https://doi.org/10.1007/s00426-018-1075-7

Belfi, A. M., Bai, E., & Stroud, A. (2020). Comparing methods for analyzing music-evoked autobiographical memories. *Music Perception*, *37*(5), 392–402.

Belfi, A. M., Karlan, B., & Tranel, D. (2018). Damage to the medial prefrontal cortex impairs music-evoked autobiographical memories. *Psychomusicology: Music, Mind, and Brain, 28,* 201-208.

Belfi, A. M., Karlan, B., & Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory*, *24*, 979–989.

Berntsen, D. (1996). Involuntary Autobiographical Memories. *Applied Cognitive Psychology*, *10*, 453–454.

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*, 113–141.

Berntsen, D. (2021). Involuntary autobiographical memories and their relation to other forms of spontaneous thoughts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, 20190693. https://doi.org/10.1098/rstb.2019.0693rstb20190693

Daselaar, S. M., Rice, H. J., Greenberg, D. L., Cabeza, R., LaBar, K. S., & Rubin, D. C. (2008).
The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall,
emotional intensity, and reliving. *Cerebral Cortex*, *18*(1), 217–229.
https://doi.org/10.1093/cercor/bhm048

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical
power analysis program for the social, behavioral, and biomedical sciences. *Behavior
Research Methods*, *39*, 175–191.

Ford, J. H., Garcia, S. M., Fields, E. C., Cunningham, T. J., & Kensinger, E. A. (2021). Older
adults remember more positive aspects of the COVID-19 pandemic. *Psychology and Aging*,
36(6), 694–699. https://doi.org/10.1037/pag0000636

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting
experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425.
https://doi.org/10.1007/s10683-011-9273-9

Jakubowski, K., Belfi, A. M., & Eerola, T. (2021). Phenomenological differences in music- and
television-evoked autobiographical memories. *Music Perception*, *38*(5), 435–455.
https://doi.org/https://doi.org/10.1525/mp.2021.38.5.435

Jakubowski, K., & Ghosh, A. (2021). Music-evoked autobiographical Memories in Everyday
Life. *Psychology of Music*, *49*, 649–666. https://doi.org/10.1177/0305735619888803

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterization of music-evoked
autobiographical memories. *Memory*, *15*(8), 845–860. https://doi.org/DOI:
10.1080/09658210701734593

Janssen, S. M. J., Rubin, D. C., & St Jacques, P. L. (2011). The temporal distribution of
autobiographical memory: changes in reliving and vividness over the life span do not

explain the reminiscence bump. *Memory & Cognition*, *39*, 1–11.

Kassambara, A. (2020). *rstatix: Pipe-friendly framework for basic statistical tests.*

https://CRAN.R-project.org/package=rstatix.

Laughland, A., & Kvavilashvili, L. (2018). Should Participants be Left to their Own Devices?

Comparing Paper and Smartphone Diaries in Psychological Research. *Journal of Applied*

Research in Memory and Cognition, *7*(4), 552–563.

https://doi.org/10.1016/j.jarmac.2018.09.002

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*, 677–689.

Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract

cuing: The proustian view is incorrect. *Applied Cognitive Psychology*, *18*(7), 893–899.

https://doi.org/10.1002/acp.1020

Mace, J. H., Atkinson, E., Moeckel, C. H., & Torres, V. (2011). Accuracy and perspective in

involuntary autobiographical memory. *Applied Cognitive Psychology*, *25*(1), 20–28.

https://doi.org/10.1002/acp.1634

Mace, J. H., Bernas, R. S., & Clevinger, A. (2015). Individual differences in recognising

involuntary autobiographical memories: Impact on the reporting of abstract cues. *Memory*,

23(3), 445–452. https://doi.org/10.1080/09658211.2014.900083

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative

platforms for crowdsourcing behavioral research. *Journal of Experimental Social*

Psychology, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and*

Word Count: LIWC2015.

Piolino, P., Coste, C., Martinelli, P., Macé, A.-L., Quinette, P., Guillery-Girard, B., & Belleville, S. (2010). Reduced specificity of autobiographical memory and aging: do the executive and feature binding functions of working memory have a role? *Neuropsychologia*, *48*(2), 429–440. https://doi.org/10.1016/j.neuropsychologia.2009.09.035

Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, *20*(4), 1842–1846. https://doi.org/10.1016/j.concog.2011.07.010

Rasmussen, A. S., Johannessen, K. B., & Berntsen, D. (2014). Ways of sampling voluntary and involuntary autobiographical memories in daily life. *Consciousness and Cognition*, *30*, 156–168.

Rasmussen, A. S., Ramsgaard, S. B., & Berntsen, D. (2015). Frequency and functions of involuntary and voluntary autobiographical memories across the day. *Psychology of Consciousness: Theory, Research, and Practice*, *2*(2), 185–205. https://doi.org/10.1037/cns0000042

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition*, *36*, 920–932.

Sheldon, S., Williams, K., Harrington, S., & Otto, A. R. (2020). Emotional cue effects on accessing and elaborating upon autobiographical memories. *Cognition*, *198*, 104217. https://doi.org/10.1016/j.cognition.2020.104217

Shiffman, S., Stone, A., & Hufford, M. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-Being: Foundations of hedonic*

psychology. Russell Sage Foundation.

https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and

computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1),

24–54. https://doi.org/10.1177/0261927X09351676

Vannucci, M., Batool, I., Pelagatti, C., & Mazzoni, G. (2014). Modifying the frequency and

characteristics of involuntary autobiographical memories. *PLoS ONE*, *9*(4).

https://doi.org/10.1371/journal.pone.0089582

Wank, A. A., Andrews-Hanna, J. R., & Grilli, M. D. (2021). Searching for the past: Exploring

the dynamics of direct and generative autobiographical memory reconstruction among

young and cognitively normal older adults. *Memory and Cognition*.

https://doi.org/10.3758/s13421-020-01098-2

Wardell, V., Esposito, C. L., Madan, C. R., & Palombo, D. J. (2021). Semi-automated

transcription and scoring of autobiographical memory narratives. *Behavior Research*

Methods, *53*(2), 507–517. https://doi.org/10.3758/s13428-020-01437-w

Appendix A – Memory Questionnaire

1) Today's Date
2) Time entry completed (please include AM or PM)
3) Time of memory (please include AM or PM). Note: if you are filling this out right after a memory has occurred this time should be the same as your previous answer.
4) What cued the memory? (e.g. song, smell, tv show) Please be as specific as possible.
5) Did the memory that you recalled come to mind spontaneously, or did you try deliberately to recall it?
> *1-It came to mind spontaneously*
> *2-I made a deliberate effort to recall it*
> *3-Not sure*

6) Please describe the memory that you recalled. Please give as much detail as possible (e.g. what you were doing, who you were with, where you were in the remembered event, etc.)
7) How old were you when the event or time period you remembered occurred? If you're not sure, please estimate or give a range (e.g. if the memory occurred in high school put 15-18)
8) Is this memory of:
> *1-A single event lasting fewer than 24 hours*
> *2-A single event lasting more than 24 hours*
> *3-A repeated event that occurred more than once*
> *4-An entire lifetime period*

9) How often have you thought of this memory before?
> *1-Never before*
> *2-Once or twice*
> *3-A few times*
> *4-Several times*
> *5-Many times*

10) How vivid was the memory? (e.g., Extremely vivid=felt like you were reliving this memory; Not at all vivid=hardly remember any details)
> *1-Not at all vivid*
> *2-A little vivid*
> *3-Somewhat vivid*
> *4-Very vivid*
> *5-Extremely vivid*

11) How negatively or positively did this memory make you feel*?*
> *1-Very negative*
> *2-Somewhat negative*
> *3-Neither positive nor negative*
> *4-Somewhat positive*
> *5-Very positive*

12) How important is this memory to your life story?
> *1-Not at all important*
> *2-A little important*
> *3-Somewhat important*
> *4-Very important*
> *5-Extremely important*

Appendix B – Cue Categorization

The *Activity/Action* category includes any cue having to do with the participant's physical actions or the actions of someone else (i.e. walking to class, waiting for an interview, my friend threw a basketball). The *Interaction* category includes cues related to a conversation or interpersonal interaction that the participant took part in with another person, including face-to-face or via text, email, or another online platform (i.e. talking on the phone, talking with a roommate, texting). The *Sensory Stimuli - Music* category includes cues that specifically mention a song, regardless of the mode of presentation of the song, including hearing another person singing a song (i.e. listening to the radio, streaming music, a song in a movie). The *Sensory Stimuli - Audiovisual* category includes cues related to media that utilizes both audio and visual components (i.e. video games, movies, TV shows). The *Sensory Stimuli - Visual* category contains cues that involve only visual components (i.e. photos, paintings). The *Sensory Stimuli - Smell/Taste* category includes any cues related to a smell or taste (i.e. food, drink, aroma). If memories were cued by the action of preparing food or actively creating a smell or taste it was categorized as action. The *Sensory Stimuli - Other* category includes cues that relate to the senses but do not fit into any other subcategories (i.e. sounds that are not musical, seeing words on a paper or screen). The *Other* category includes memory cues that did not fall under the other four categories nor did it have enough other similar cues to warrant needing its own category (i.e. the month of June, another memory, feelings). See the table below for specific examples of each cue type.

Cue type	Description	Examples from participants
Activity/Action	Something physically done by the participant or something physically done by another person that the participant witnessed firsthand	"I was playing IM flag football," "Playing games," "Waiting for [an] internship interview"
Interaction	Any interaction between the participant and another person, face-to-face or otherwise	"I was having a text conversation with someone and started talking about my father," "Talking with friends"
Sensory Stimuli - Music	A song or part of a song, whether heard on the radio, a streaming service, a show or movie, or sung/hummed from another person	"Song, a Blood on The Dancefloor song," "a song by Lil' Wayne came on my Apple Music shuffle," "I was listening to music when it happened"
Sensory Stimuli - Audiovisual	Media that involves both audio and visual components, including movies, TV shows, music videos, video games, etc.	"A TV scene about going on a date," "Social media videos of snow," "A movie"
Sensory Stimuli - Visual	A still image, including photographs, paintings, etc.	"A picture I have on my wall of a scene from a movie," "Drawing of a dragon in my room"
Sensory Stimuli - Smell/Taste	A food, drink, or other taste or smell	"Smell of apple muffins," "Eating and smelling pizza"
Sensory Stimuli - Other	Any other sensory stimuli not covered by the above categories, such as written words and non-music audio	"The fire alarm," "Assignment description on canvas"
Other	Any cue that does not fall into the other categories	"Today's date," "Another memory," "Nothing"

Supplemental Analyses

We conducted a series of linear mixed-effects models to account for individual subject-level effects. For all dependent variables, computed a linear mixed effects model using the lmer function from the lmerTest package (Kuznetsova et al., 2017) in R (version 3.6.2). All continuous predictors were centered within each subject (i.e., "group mean centering") before being submitted to the model (Enders & Tofighi, 2007). For each dependent variable (word count, timing, all AI variables and all LIWC variables), a linear mixed-model analysis was run with the dependent variable as the outcome variable, and group (audio vs. text) as a fixed effect. Random effects included random intercepts and slopes for subjects. A model was not conducted for number of memories due to its nature of being a count variable and the fact that we were interested precisely in whether participants in each group reported different numbers of memories (therefore counting the memories for each participant). The only result that differs somewhat between our two analytical approaches is that the effect of cognitive words, which is significant in the ANOVA approach, is marginally significant in this mixed-effects model approach.

Outcome variable	β	SE	T	p	sig.
Word count	-29.10	4.06	-7.17	<0.001	***
Time difference	-10.15	16.06	-0.63	0.52	
Affective words	0.59	0.38	1.52	0.12	
Cognitive words	-1.37	0.52	-2.60	0.009	**
Social words	0.23	0.59	0.39	0.69	
Perceptual words	0.61	0.35	1.75	0.07	
External details	-2.69	0.42	-6.36	<0.001	***
Internal details	-1.31	0.63	-2.06	0.03	*
Internal/total ratio	0.12	0.01	6.21	<0.001	***

Table S1. Results of linear mixed effects models. Each row represents a different model, the results are the fixed effect of group (text vs. audio) on the listed outcome variable in each row. *p<0.5, **p<0.01, ***p<0.001