LUNDS UNIVERSITET

Department of Psychology

Memories Between Us: Autobiographical Memory Sharing with Friends and Strangers

Simone Mikaela Jalk Carlé

Master's Thesis (PSYP01)

August 24th, 2026

Supervisor: Mikael Johansson

Abstract

Autobiographical remembering is reconstructive and frequently occurs in social interaction, yet it remains unclear whether the relationship between narrator and listener shapes how memories are expressed and subsequently remembered. This study examined autobiographical memory sharing with friends versus strangers and evaluated the use of a large language model (GPT-5.5) for coding narrative data. Fifty-six adults (28 narrators, 28 listeners) participated in a repeated-measures memory-sharing paradigm involving both relationship conditions. Narrators shared positive autobiographical memories, with a non-shared memory serving as a baseline for pre-post comparisons. Narratives were coded for episodic richness and self-event connections (SECs), and listeners later recalled the memories they had heard. Relationship condition did not significantly affect episodic richness during narration, pre-post change, or listener recall. Pre-post change in SECs likewise did not differ across friend, stranger, and baseline conditions. During narration, narrators produced more SECs when speaking to strangers than friends, although this effect was no longer significant after accounting for pre-session SECs and should therefore be interpreted cautiously. GPT coding showed framework-dependent performance, with stronger reliability and human agreement for total episodic richness than for some other coding frameworks. Overall, the findings provide little evidence that a single socially situated recall produces broad or persistent changes in autobiographical memory, while suggesting that relationship context may selectively influence self-related framing. LLM-based narrative coding showed promise but requires framework-specific validation.

Keywords: autobiographical memory, memory sharing, reconstructive memory, interpersonal relationships, large language models.

Thank you!

I would first like to express my sincere gratitude to my husband for his unwavering support throughout my education. I am especially grateful for his willingness to move across borders with me on multiple occasions, for encouraging me in my academic pursuits, and for generously sharing his expertise in software and large language models throughout the development of this thesis. I would also like to thank my parents for always supporting my adventurous nature and encouraging me to pursue a life and education abroad. Their support has given me the confidence to explore the world while always knowing that home is never far away.

I am grateful to my supervisor, Mikael Johansson, for his guidance, thoughtful feedback, and reassurance throughout the research process. His support helped make the completion of this thesis a smooth and rewarding experience.

My sincere thanks also go to all the participants who generously gave their time over multiple days to take part in this study. I am especially grateful for their engagement, enthusiasm, and interest in the project, as well as for their willingness to invite a friend to participate alongside them.

Finally, I would like to thank my friends in Lund, who made these two years some of the most rewarding of my life. I am grateful for the many academic discussions that broadened my perspective and deepened my curiosity, but above all, for their genuine friendship. Your encouragement, companionship, and support have been a gift.

Introduction

Human episodic memory is shaped not only by how events are initially experienced and encoded, but also by how they are subsequently retrieved and, under certain conditions, modified. Rather than functioning as a passive store of past experiences, memory is increasingly understood as an ongoing, dynamic, and reconstructive process that serves adaptive functions beyond the accurate preservation of past events (Lepsius, 2026).

These reconstructive properties are particularly relevant to autobiographical memory, as such memories are closely tied to self-identity and self-continuity (Harris et al., 2014). This relationship is reciprocal: autobiographical memories are reconstructed over time in response to new experiences and current psychological needs and goals, while simultaneously contributing to the maintenance of self-identity (Jiang et al., 2020; Wilson & Ross, 2003). Autobiographical remembering also frequently occurs within social interaction (Beike et al., 2016), where memories are not only individually reconstructed but socially organized and given meaning across repeated retellings (Fivush et al., 2011). Memory sharing is therefore not simply the reporting of an already formed memory, but a socially embedded process involving both a narrator and a listener. This creates the possibility that the relationship between them influences how an event is reconstructed during narration and subsequently recalled.

The present study lies at the intersection of cognitive memory research and social psychology. From a memory perspective, it examines how autobiographical events are reconstructed during retrieval; from a social-psychological perspective, it considers how interpersonal goals, relationships, and the listener's presence may shape that process. Autobiographical memory sharing can accordingly be understood as both a cognitive act of retrieval and a socially situated interaction between narrator and listener.

Episodic Memory and Retrieval

Memory comprises multiple distinguishable systems that support different forms of learning and remembering (Tulving, 2002). Within long-term memory, a key distinction is between declarative and non-declarative memory, with declarative memory commonly divided into semantic and episodic memory (Squire, 1992). Semantic memory concerns general knowledge largely independent of its original learning context, whereas episodic memory concerns personally experienced events situated within a specific time and place (Tulving, 2002). Episodic remembering is characteristically associated with autonoetic consciousness: the self-aware capacity to mentally travel back in subjective time and experience an event as part of one's own past (Tulving, 2002). It is this capacity to recollect and subjectively re-experience contextually situated events that is relevant to the present study.

At a neural level, episodic retrieval involves the reactivation of event-specific memory representations. According to the cortical reinstatement hypothesis, retrieval reinstates patterns of neural activity associated with the original experience (Bosch et al., 2014), supported partly by hippocampal pattern completion, whereby partial retrieval cues reactivate associated elements of an event (Horner et al., 2015). These processes contribute to the recovery of event information and may help explain the subjective sense of reliving that characterizes episodic remembering (Richter et al., 2016).

Memory as a Dynamic Adaptive System

Importantly, however, the reactivation of event information during retrieval should not be understood as a literal replay of the original experience. Rather, retrieval is a reconstructive process in which stored event information is assembled in relation to current goals and demands. This flexibility can be understood as an adaptive property of memory: rather than reproducing every aspect of a past experience indiscriminately, retrieval can prioritize information relevant to the individual's present needs, goals, and context (Lepsius, 2026). The same flexibility, however, also creates opportunities for selective accessibility, updating, and distortion, such that what is retrieved may vary across occasions (St. Jacques & Schacter, 2013).

The purpose for which a memory is recalled can influence how it is retrieved. For example, participants instructed either to recall an event as fully as possible or to recount it as though sharing it with a friend produced narratives that differed in content and linguistic style (Eckardt et al., 2024). These findings demonstrate that communication goals can shape the information selected and expressed during recall. Recall biases may also serve self-protective functions. The mnemic neglect effect refers to the relatively poorer recall of negative, self-relevant information that threatens central self-beliefs and is commonly understood as a self-protective process (Liu et al., 2025; Sedikides et al., 2016). Together, these findings illustrate how current goals and psychological needs can shape what is selected, communicated, and made accessible during retrieval.

These effects demonstrate flexibility during retrieval, but a distinction is required between changes in what is selected or expressed during a particular recall episode and changes that persist into subsequent remembering. Under particular conditions, reactivating a consolidated memory may render it temporarily labile and susceptible to modification before it is restabilized through reconsolidation (Hupbach et al., 2008; Lepsius, 2026; Nader et al., 2000). This temporary plasticity may support adaptive updating by allowing information encountered in the retrieval context to influence subsequent remembering (Hupbach et al., 2008; Lepsius, 2026; Yang et al., 2026). Therefore, the flexibility that allows memory to respond adaptively to current circumstances may also provide a mechanism through which retrieval experiences influence later representations of the past.

Autobiographical Memories

These reconstructive and adaptive properties are particularly relevant to autobiographical memory. Within the Self-Memory System, Conway and Pleydell-Pearce (2000) conceptualize autobiographical memories as temporary constructions rather than fixed representations that are simply accessed during retrieval. Their construction is constrained partly by the working self, which contains currently active goals and influences which autobiographical knowledge becomes accessible. The model therefore provides a mechanistic account of how the broader flexibility of memory is expressed in autobiographical remembering: different goals and retrieval contexts can alter which aspects of the same past event are accessed and how they are organized into a particular recollection.

Because autobiographical memories concern personally experienced events, their reconstruction is also closely related to self-concept and identity. A central function of autobiographical memory is thought to be maintaining coherence and continuity in the representation of the self across time (Conway & Pleydell-Pearce, 2000; Conway et al., 2004). Consistent with this account, Jiang et al. (2020) found that recalling autobiographical memories could restore self-continuity when self-concept clarity had been disrupted. Autobiographical remembering can therefore serve current self-related goals by organizing past experiences in relation to the present self.

Current goals, however, do not arise independently of the context in which remembering occurs. Because autobiographical memories are frequently retrieved and communicated in interaction with other people, conversational remembering provides a particularly relevant context in which the adaptive and reconstructive properties of memory may be socially influenced.

Memory Sharing

In conversational remembering, the goals and demands guiding retrieval are not determined solely by the narrator but also arise from the social interaction in which remembering occurs. Communicating an event requires the narrator to reactivate, select, organize, and express information for a particular listener, making the audience itself part of the retrieval context. The reconstructive flexibility described above therefore creates the possibility that who a memory is shared with may influence both how it is narrated and how it is subsequently remembered.

References to past events occur in approximately 44% of everyday conversations (Beike et al., 2016), and autobiographical memories are shared for functions including social connection, communication of information, and support of the self (Alea & Bluck, 2003, 2007). Conversational remembering is therefore not an exceptional retrieval condition but a common context in which autobiographical reconstruction occurs.

One social motivation that may shape this reconstruction is the establishment of shared reality: the experience of sharing an interpretation, belief, or feeling about an event or topic with another person. Shared reality can serve epistemic goals by increasing confidence in one's understanding and relational goals by fostering interpersonal connection (Echterhoff et al., 2009). Consistent with this account, audience-congruent information becomes more cognitively accessible during retrieval, suggesting that the pursuit of shared understanding can influence which aspects of an event are most readily available for communication (Wagner et al., 2024). Greater shared reality during everyday memory conversations has also been associated with higher psychological well-being (Boytos & Costabile, 2023).

Memory sharing consequently involves two linked constructive processes: the narrator reconstructs a personally experienced event, while the listener constructs a representation of that event from the narration. These processes are interdependent, as the narrator determines which information is available to the listener, while the listener's responses and the interpersonal goals of the interaction may influence what the narrator retrieves and emphasizes. Consistent with the transmission of event representations between individuals, Zadbood et al. (2017) found that listeners who had not experienced an event developed neural representations resembling those communicated by the speaker. However, because listeners heard prerecorded narrations, the study could not examine reciprocal influences on the narrator.

Critically, these social demands may differ according to the relationship between narrator and listener. Friends enter an interaction with accumulated common ground, including shared history and knowledge of one another's preferences and interpretations, whereas strangers have less established common ground and may require greater contextualization and negotiation of shared understanding. Because reconstructive retrieval is sensitive to communicative goals and demands, these relationship contexts may influence which event details are selected, how extensively they are elaborated, and how the event is framed. Shared reality is also positively related to interpersonal closeness in established relationships and may facilitate connection between strangers (Rossignac-Milon et al., 2021; Rossignac-Milon & Higgins, 2018).

Friend-related information has been shown to receive processing priority relative to stranger-related information (Leng et al., 2025). Although this finding was not obtained in a conversational memory-sharing context, it suggests that information associated with close others may receive greater cognitive priority. This raises the possibility that listeners may engage more with autobiographical memories shared by friends, potentially supporting richer subsequent recollection of the narrated event.

Listener behavior provides further evidence that retrieval is sensitive to the interaction itself. Grysman et al. (2023) found that narrators paired with attentive listeners included more information than those paired with distracted listeners, although attentiveness did not produce a persistent effect on later recall. Consequently, listeners can influence retrieval during an interaction without those differences necessarily persisting at subsequent recall.

Taken together, conversational remembering involves reconstruction by the narrator and construction by the listener, with both processes potentially shaped by their interpersonal relationship. The central question for the present study is therefore whether sharing a memory with a friend versus a stranger influences these constructions during the interaction and whether any differences persist in subsequent recall.

Narrative Features of Autobiographical Memory

The recall of autobiographical memories can be examined through two complementary features of autobiographical accounts: episodic richness and self-event connections (SECs). Episodic richness reflects the amount and type of event-specific information retrieved and communicated, including spatial references, entities present, sensory descriptions, thoughts and emotions, and actions. It therefore provides an indication of how extensively an event is reconstructed and elaborated during recall (Hassabis et al., 2007; Marlatte et al., 2022). Because narrators select and organize these details for a particular audience, episodic richness may vary depending on the social context in which the memory is shared.

SECs capture explicit links between a remembered event and broader aspects of the narrator’s self, including dispositions, values, outlook, and personal growth (McLean & Fournier, 2008). These connections indicate how an event is interpreted and integrated into broader self-understanding rather than merely recalled as an isolated experience. Together, episodic richness and SECs allow the present study to examine whether relationship context shapes both the reconstruction of event information and the meaning assigned to the event in relation to the self.

Previous Research and the Current Research Gap

Several relevant narrator-listener paradigms have relied on indirect or non-reciprocal interactions, including prerecorded narratives (Wei et al., 2024; Zadbood et al., 2017) and instructions to narrate a memory "as if" speaking to a friend (Eckardt et al., 2024). Although these approaches provide experimental control, they cannot capture the reciprocal interpersonal dynamics through which a listener may influence the narrator's reconstruction during retrieval. Grysman et al. (2023) examined later recall following direct interaction between friends but manipulated listener attentiveness rather than relationship type and did not include a stranger condition. It therefore remains unclear whether the relationship between actual narrators and listeners influences both the immediate construction of autobiographical narratives and their subsequent reconstruction.

A further methodological challenge concerns the coding of autobiographical narratives. Measures such as episodic richness, memory specificity, and self-event connections typically require resource-intensive manual coding. Large language models may offer a more efficient way to apply predefined coding frameworks, but their consistency and agreement with human coding must be established before their scores can be interpreted confidently. Recent work has shown promising applications of GPT-4 and related natural language processing approaches for coding emotional content and autobiographical-memory data (Mistica et al., 2024; Oded et al., 2026; van Genugten & Schacter, 2024). The present study therefore evaluates the use of an LLM for structured coding of autobiographical narratives, extending previous work by assessing its consistency and agreement with human coding across predefined narrative-coding frameworks.

The Present Study

The present study extends reconstructive accounts of autobiographical memory by testing whether the interpersonal relationship in which retrieval occurs acts as a socially situated constraint on memory construction. Specifically, it examines whether interactions with friends versus strangers influence the narrator's immediate reconstruction of an autobiographical event, whether any differences persist in later recall, and how listeners subsequently represent the narrated event.

To address limitations of previous memory-sharing paradigms, the study uses real autobiographical memories and direct, real-time interactions between actual friends and previously unfamiliar narrator-listener pairs, rather than prerecorded narratives or imagined audiences. This design also extends prior narrator-listener work by assessing listeners' later representations of shared memories, which was not examined in Eckardt et al. (2024).

The study focuses not on memory accuracy, but on the organization and subjective experience of autobiographical remembering. Episodic richness and self-event connections are therefore considered alongside subjective ratings of vividness and emotional intensity.

A non-shared baseline memory was included to distinguish changes associated with social narration from changes related to repeated recall alone. This memory was recalled before and after the session but was not narrated to another person during the interaction.

As a secondary methodological aim, the study evaluated an LLM for coding autobiographical narratives using predefined frameworks for episodic richness, memory specificity, and self-event connections. Its run-to-run reliability and agreement with human coding were assessed to evaluate consistency of the generated scores and their correspondence with human coding.

Accordingly, the study addresses four research questions:

1. How does sharing an autobiographical memory with a friend versus a stranger affect its episodic richness and self-event connections during narration?
2. Do pre-post changes in episodic richness and self-event connections differ between memories narrated to a friend, narrated to a stranger, and not narrated during the sharing session?
3. Do listeners recall greater episodic richness from autobiographical memories narrated by a friend than from memories narrated by a stranger?
4. To what extent can a large language model consistently apply predefined narrative-coding frameworks and generate codes that agree with human coding across autobiographical narratives?

It was hypothesized that memories narrated to a stranger would contain greater episodic richness than those narrated to a friend, reflecting the greater communicative demands of establishing shared understanding with an unfamiliar listener (H1a). Memories narrated to friends and strangers were also expected to differ in the number of self-event connections produced during sharing (H1b). For pre-post change, episodic richness was expected to differ across the friend, stranger, and non-shared baseline conditions, reflected in a phase × condition interaction (H2a), with both the friend (H2b) and stranger (H2c) conditions expected to differ from baseline. The same pattern was hypothesized for self-event connections: a phase × condition interaction (H2d), with change in both the friend (H2e) and stranger (H2f) conditions expected to differ from baseline. Finally, listeners were expected to recall greater episodic richness from memories narrated by a friend than by a stranger (H3).

Method

Design

This study employed a mixed repeated-measures experimental design conducted online. The main within-subjects factor was social relationship condition, as all participants engaged in autobiographical memory sharing with both a friend and a stranger. The condition order was counterbalanced, with half of the participants starting in the friend condition and the other half in the stranger condition, to control for potential order effects. For narrator recall analyses, the design additionally included a non-narrated baseline memory. The study also included one between-subjects factor: participant role (narrator vs. listener). Within each friend dyad, participants were assigned to either the narrator or listener role, which remained fixed throughout the study.

Narrators prepared three positive autobiographical memories (A, B, and C) from the past three years. Two memories were used during the experimental sessions, while the third served as a non-narrated baseline comparison condition. To further control for potential order and memory effects, the study implemented a counterbalanced design across the three memories, resulting in six possible memory orders that were distributed evenly across participants. Outcome measures included post-session written recall tasks in which narrators recalled all three personal memories and listeners recalled the memories shared with them during the experimental sessions.

Participants

A planned sample of 24 dyads/48 participants was selected to permit equal distribution across the six possible memory orders. However, four additional narrator-listener pairs were recruited to account for participant errors and exclusions, resulting in a total sample of 56 participants contributing data to the study. Participants were adults aged 18 years or older with sufficient English proficiency to complete the study tasks.

Four narrators were excluded from analyses involving pre-session and post-session narrator data. Two exclusions were due to procedural modifications introduced after the initial data collection sessions, which made these data non-comparable with the rest of the sample. The remaining two narrators were excluded because they did not complete the pre-session memory-writing task as instructed. Their in-session narratives were retained, however, because the corresponding listeners had heard and responded to these memories during the experimental sessions.

Narrator recall analyses were therefore conducted on the final fully counterbalanced sample of 24 narrators, evenly distributed across the six memory orders. Listener analyses retained data from all 28 listeners, as they had completed the memory-sharing sessions according to protocol. Although the listener sample was not perfectly counterbalanced, all valid listener data were retained.

Recruitment and Incentive

Participants were recruited using convenience sampling via social media (Facebook and Instagram), campus posters, in-class presentation, and word-of-mouth. Recruitment additionally involved participant-recruited friend dyads, whereby each participant was required to recruit a friend to participate in the study. All participants were automatically entered into a raffle for a single 500 SEK Zalando gift card.

Ethical Considerations

Participants were informed that participation was voluntary and received information about potential risks, withdrawal rights, data handling, and Zoom audio/video recording before providing informed consent. To minimize distress associated with autobiographical recall, participants were instructed to describe positive memories.

Participant data were pseudonymized using unique identification codes. A separate document linking participant ID codes to contact information (email addresses) was created solely for the purpose of coordinating participation and scheduling online sessions. This file was stored locally on the researcher's computer and was not shared, to limit access to identifiable information.

The online OpenAI large language model GPT-5.5 Thinking was used to code autobiographical memories. Additional measures were therefore implemented to protect participants' confidentiality. Before submission for coding, the narrative data were pseudonymized by redacting names, locations, significant dates, religious practices, and other potentially identifying information. Each coding was conducted using the Temporary Chat setting, and the account-level setting allowing conversations to be used for model improvement was disabled.

Materials and Measures

The study was conducted online using the video conferencing platform Zoom and the survey platform Qualtrics.

Narrative Data Transcription

Narrative data from the memory-sharing sessions were transcribed using Sunet Scribe, a speech-to-text service operated by Sunet. Prior to transcription, the audio was normalized and converted to a standardized format using FFmpeg, with noise filtering applied as part of preprocessing. The recordings were then transcribed in English using the OpenAI Whisper

Large-v3 automatic speech-recognition model. The resulting transcripts included timestamps and were segmented into logical units based on pauses and speech rhythm. All processing and transcription were conducted within Sunet's data centers in Sweden (Sunet, n.d.).

Subjective Vividness and Emotional Intensity Scales

Subjective memory characteristics were measured using the vividness and emotional-intensity subscales of the Memory Experiences Questionnaire (MEQ; Sutin & Robins, 2007). The items were adapted according to participant role, such that narrators rated each memory based on their experience while narrating it, whereas listeners rated the corresponding memory based on how they experienced it when receiving it from the narrator. Each subscale consisted of six items, with items 4, 5, and 6 reverse scored. Participants rated their agreement with each statement on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). Mean scores were calculated across the six items for each subscale, yielding scores ranging from 1 to 5, with higher scores indicating greater vividness or emotional intensity. The adapted MEQ subscales are provided in Appendix A.

Interpersonal Closeness Scale

Perceived interpersonal closeness was measured using the single-item Inclusion of the Other in the Self Scale (Gächter et al., 2015). Participants selected one of seven pairs of increasingly overlapping circles, scored from 1 to 7, with higher scores indicating greater perceived closeness. Participants completed the scale separately in relation to the friend and stranger.

Narrative Data Coding Frameworks

All memories, across narrators and listeners, were coded using the two coding frameworks: autobiographical memory specificity and Episodic Richness. Memory specificity was classified into three mutually exclusive categories: specific, extended, and categoric. Specific memories referred to a single event lasting no longer than one day. Extended memories referred to a single event or period lasting longer than one day, whereas categoric memories referred to repeated or similar events without identifying one particular occurrence (Ford et al., 2014).

Episodic richness was measured using an adaptation of the Experiential Index (Hassabis et al., 2007; Marlatte et al., 2022). The original Experiential Index combined scores derived from narrative content with participants' subjective ratings. In the present study, only the content component was used as the measure for episodic richness. This component included scores for spatial references (SPA), entities present (EP), sensory descriptions (SD), and thought/emotion/action (TEA). For the purposes of this study, the TEA component was divided into two separate categories: thought and emotion (TE), and action (A). As the original coding system was developed to assess episodic scene construction, this modification was introduced to better capture the autobiographical memories examined in the present study. In addition, no maximum score was imposed because the study aimed to detect changes in richness over time. Total episodic richness was calculated as the sum of SPA, EP, SD, TE, and A.

Narrators' memories were additionally coded using the Self-Event Connection framework (McLean & Fournier, 2008). Each connection was classified into one of four mutually exclusive categories: dispositions, values, outlook, or personal growth. The frequencies across the four categories were summed to produce a total self-event connections (SECs) score for each memory.

Procedure

Pre-Session Questionnaire

The day before the scheduled Zoom session, participants received an email containing their personal ID code, pair identifiers for Sessions 1 and 2, assigned role (narrator or listener), and links to the Qualtrics questionnaire and Zoom session. This temporal separation was intended to reduce recency and rehearsal effects that might arise if memories were written immediately before the experimental session.

The questionnaire included informed consent and demographic questions. Narrators additionally completed a memory-preparation task in which they described three strong, positive autobiographical memories from the past three years, labeled Memories A, B, and C. They were instructed to take their time while aiming to complete the task within approximately 30-35 minutes, to select memories the listener had not previously heard, and to note the A-C labels for later reference. Listeners completed only the consent and demographic sections.

Experimental Session

Two recruited friend dyads joined the Zoom session at the scheduled time. The session started with the researcher providing instructions for the procedure. Narrators were instructed which memory to share during the session (A, B, or C), depending on their assigned counterbalancing order. Listeners were instructed to listen without interrupting, although brief verbal responses (e.g., "ah," "ok") and non-verbal behaviors (e.g., smiling or nodding) were permitted. Narrators were given two minutes for free recall of each memory; if they finished early, they were instructed to verbally indicate that they had completed their narration.

After the instructions, each dyad was assigned to a breakout room. The order of conditions was counterbalanced, such that half of the sessions began in the friend condition and the other half in the stranger condition. Within each breakout room, one participant was assigned the role of narrator and the other the role of listener; depending on condition, dyads consisted either of recruited friend dyads or of cross-dyad pairings of participants from the other friend dyad within the same session block (i.e., unfamiliar partners).

The researcher joined the first dyad, instructed the narrator on which specific memory (A, B, or C) to share, and initiated the recording for the two-minute free-recall task. Upon completion, participants were informed that the first session had ended and were instructed to wait for their second session. While this dyad waited, the second dyad remained in their assigned breakout room. The researcher then left the first dyad and joined the second dyad, where the same procedure was repeated. Once both dyads had completed this procedure, the first session was concluded.

All participants then proceeded to a second narration session in the alternate condition. Dyads were reassigned to new breakout rooms, and the procedure was repeated under the new condition. Narrators again shared an autobiographical memory (A, B, or C) according to their assigned counterbalancing order. After completing the second narration session, all participants were provided with a link for the second questionnaire, concluding the experimental session.

Post-Session Questionnaire

Following the experimental session, narrators and listeners rated the vividness and emotional intensity of the memories shared in both conditions and reported perceived closeness with their interaction partner, based on how they perceived the relationship before the session.

Narrators then completed written recall of all three memories (A, B, and C) in their assigned counterbalancing order. Listeners completed written recall of the two memories shared with them, in the order in which they had been presented.

Data Preparation and Scoring

Subjective Vividness and Emotional Intensity

Negatively keyed items were reverse-scored, after which mean vividness and emotional-intensity scores were calculated across the six items in each subscale, with a maximum possible mean score of 5.

Perceived Closeness

Each participant contributed one IOS score for the friend condition and one for the stranger condition. Scores ranged from 1 to 7, with higher scores indicating greater perceived closeness, rated as it had been before the online session.

Narrated Session Memories

All Sunet Scribe transcripts were manually reviewed, with the original audio recordings consulted where necessary to verify and correct transcription errors.

Coding of Autobiographical Narrative Data

A large language model was used as an automated rater to apply predefined coding frameworks across the narrative dataset, allowing standardized instructions to be applied consistently while reducing the burden of manual coding. Before coding, all memories were manually pseudonymized by redacting identifying information, such as names and locations, and replacing it with descriptive labels (e.g., "[Redacted: country name]"). The LLM was instructed to treat these labels as representing the original information. Punctuation was also added manually to support sentence-level coding. GPT-5.5 Thinking was used for all LLM-based coding and was accessed on June 22, 2026. The account-level "Improve the model for everyone" setting was disabled, and all coding was conducted using Temporary Chat.

Separate standardized prompts were developed for autobiographical memory specificity, episodic richness, and self-event connections (SECs). Each included operational definitions, scoring criteria, and instructions not to infer information beyond what was explicitly stated. Each eligible memory was coded independently five times for each applicable framework. The prompts were developed with reference to the framework presented by Oded et al. (2026). See Appendix B for the full prompts for specificity classification (B1), episodic richness (B2) and self-event connections (B3).

Specificity Classification

Autobiographical memory specificity was coded using a standardized prompt with three mutually exclusive categories: specific, extended, and categoric. Memories were classified according to whether they described a single event lasting no longer than one day, a continuous event lasting longer than one day, or repeated or similar experiences; mixed memories were classified according to their primary focus. Outputs were coded as 1 = specific, 2 = extended, and 3 = categoric. The final classification was determined by majority vote across the five coding runs. If no category received at least three votes, the memory was manually reviewed by the researcher to determine the final classification.

Episodic Richness

Episodic richness (ER) was coded using a standardized prompt containing definitions of the five subcomponents: spatial reference (SPA), entities present (EP), sensory description (SD), thought and emotion (TE), and action (A). Narratives were processed at the sentence level, with each explicit occurrence of SPA, SD, TE, and A counted separately. EP was coded across the full narrative, such that each unique person, object, or animal was counted only once, at its first mention. The final score for each component was the median of its five model-generated scores. Total episodic richness was then calculated by summing the median SPA, EP, SD, TE, and A scores.

Self-Event Connections

Self-event connections (SECs) were coded using a standardized sentence-level prompt. An SEC was defined as an explicit link between the remembered event and an aspect of the narrator's self, relating to dispositions, values, outlook, or personal growth. Each sentence was evaluated independently and could receive a maximum score of one. Connections were only counted when the narrator was explicitly referenced and the connection was directly stated rather than inferred. Statements describing emotions alone or referring only to other people were not counted. Sentence-level counts were summed to produce a total SEC score for each memory. The final SEC score for each memory was the median of the five runs.

Reliability

GPT-based coding reliability was assessed in two stages. First, inter-run reliability across the five independent coding runs was evaluated using two-way random-effects, absolute-agreement, single-measure intraclass correlation coefficients (ICCs) for the count-based variables episodic richness (ER) and self-event connections (SECs). Category-specific ICCs were also calculated for the ER subcomponents (SPA, EP, SD, TE, and A). ICC values were interpreted according to Koo and Li (2016), with values below .50 considered poor, .50-.75 moderate, .75-.90 good, and above .90 excellent. For the three-category nominal memory-specificity variable, inter-run reliability was assessed using Fleiss' kappa.

Second, final GPT-derived scores were compared with human coding in validation subsets. For ER and memory specificity, 28 of 256 memories (10.9%) were manually coded; for SEC, 24 of 200 eligible memories (12.0%) were manually coded. The validation subsets were selected without reference to their scores or classifications and without a formal randomization procedure. Although human inter-rater reliability is commonly assessed using approximately 20% of coded material, a smaller validation subset was used due to the time-intensive nature of manual narrative coding (Wank et al., 2020). Human-GPT agreement was assessed using absolute-agreement ICCs for ER and SEC and Cohen's kappa for memory specificity.

Statistical Analyses

All analyses were conducted in R version 4.6.1 (R Core Team, 2026) using RStudio version 2026.07.1+147 (Posit team, 2026). Mixed-effects models were fitted using *lme4*, with *lmerTest* used for inferential tests of linear mixed-effects models (LMMs), *emmeans* for estimated marginal means and planned contrasts, and *DHARMa* for simulation-based diagnostics of Poisson generalized linear mixed-effects models (GLMMs).

Linear mixed-effects models (LMMs) were used for episodic richness and other continuous outcomes, whereas Poisson generalized linear mixed-effects models (GLMMs) were used for SEC counts. Mixed-effects models accounted for the non-independence of repeated observations within participants and, where applicable, repeated assessments of the same memories. Participant-level random effects were particularly relevant given the free-recall format, as participants could differ systematically in their general tendency to provide more extensive or detailed responses.

Statistical significance was evaluated at $\alpha = .05$, with 95% confidence intervals reported for model estimates. Where multiple related planned or exploratory comparisons were conducted, p values were adjusted using the Holm procedure.

The relationship manipulation was assessed using a two-sided paired Wilcoxon signed-rank test comparing friend and stranger closeness ratings, with the Hodges-Lehmann estimate and 95% confidence interval reported for the paired difference. GPT coding reliability was evaluated using Fleiss' κ for run-to-run memory-specificity classifications and Cohen's κ for human-GPT specificity agreement. For numeric coding outcomes, run-to-run reliability was assessed using two-way random-effects, absolute-agreement, single-measure ICCs [ICC(2,1)], whereas human-GPT agreement was assessed using two-way mixed-effects, absolute-agreement, single-measure ICCs [ICC(A,1)]. For human-GPT agreement, ICCs were supplemented by descriptive measures of score differences, and Pearson's correlation was additionally reported for total episodic richness.

For H1, effects of relationship condition on session episodic richness and SECs were tested using an LMM and Poisson GLMM, respectively, with narrator as a random intercept. H2 was tested using phase × condition models examining pre-post change across the friend, stranger, and non-shared baseline conditions. The episodic richness model included participant intercepts and phase slopes and a memory intercept, whereas the SECs model included participant and memory random intercepts. Planned contrasts compared change in the friend and stranger conditions with baseline. H3 was tested using an LMM predicting listener-recall episodic richness from relationship condition, with listener as a random intercept.

For H3, a secondary model additionally adjusted for the episodic richness of the specific narration heard by each listener. Exploratory analyses examined condition differences in narrators' subjective vividness and emotional intensity and whether within-participant variation in these ratings was associated with narrator episodic richness or listener recall, adjusting for relevant baseline or narration-level episodic richness. Vividness and emotional-intensity ratings were participant-mean centered to isolate within-participant effects.

Supplementary analyses examined narrative length using log-transformed word count, including its relationship with condition and episodic richness. The H1 episodic-richness model was additionally adjusted for log-transformed word count, while the natural logarithm of word count was included as an exposure offset in the SEC model. An additional sensitivity analysis for H1b, based on the 24 participants with complete pre-session data, adjusted session SEC counts for the corresponding pre-session SEC score to assess whether the relationship-condition effect remained after accounting for pre-existing differences between memories. Model assumptions were assessed using residual and Q-Q plots and checks for convergence and singularity; Poisson models were additionally evaluated for dispersion, zero inflation, outliers, and residual-distribution deviations using simulation-based diagnostics.

Results

Sample Composition

The sample comprised 56 participants: 28 narrators and 28 listeners. Most participants were aged 25-34 years (67.9%); full demographic characteristics are presented in Table 1.

Table 1

Demographic Characteristics, n(%)

Characteristics	Total	Narrators	Listeners
Sex			
Female	33 (58.9%)	17 (60.7%)	16 (57.1%)
Male	23 (41.1%)	11 (39.3%)	12 (42.9%)
Age group			
18-24	5 (8.9%)	1 (3.6%)	4 (14.3%)
25-34	38 (67.9%)	20 (71.4%)	18 (64.3%)
35-44	8 (14.3%)	5 (17.9%)	3 (10.7%)
45-54	2 (3.6%)	0 (0%)	2 (7.1%)
55-64	2 (3.6%)	1 (3.6%)	1 (3.6%)
65 or older	1 (1.8%)	1 (3.6%)	0 (0%)

Note. Values are n (%). N = 56 (28 narrators and 28 listeners). Percentages are calculated within columns.

Manipulation Check

To confirm that the friend and stranger conditions represented meaningfully different relationship contexts, a manipulation check was conducted using participants' subjective ratings of closeness to each interaction partner on the Inclusion of Other in the Self Scale (IOS). Closeness ratings were available for 55 participants. Participants rated friends as substantially closer (*Mdn* = 6, IQR = 2) than strangers (*Mdn* = 1, IQR = 1), $V = 1537$, $p < .001$. The Hodges-Lehmann estimate indicated a difference of 4.22 IOS points, 95% CI [3.50, 5.00]. 54 of 55 participants (98.2%) rated their friend as closer than the stranger, confirming the effectiveness of the relationship manipulation.

Reliability

To assess the run-to-run reliability of narrative coding performed by GPT, agreement across five runs was evaluated for memory-specificity classification, episodic richness (ER), and self-event connections (SECs).

Run-to-run agreement for memory-specificity coding across the five runs was high, $\kappa = .867$, $z = 51.20$, $p < .001$. Run-to-run reliability of the total ER score was good, ICC(2,1) = .864, 95% CI [.754, .917], $F(255, 1020) = 56.62$, $p < .001$. As secondary diagnostics, ICC(2,1) values were computed separately for each ER subcomponent; SPA, EP, SD, TE and A. As shown in Table 2, run-to-run reliability was excellent for TE, good for SPA and A, and moderate for EP and SD. However, the confidence intervals for SPA, EP, SD, and A crossed the threshold between moderate and good reliability.

Table 2

Run-to-Run Reliability of Episodic Richness Subcomponents

Subcomponents	ICC(2,1)	95% CI
SPA	.810	[.734, .862]
EP	.707	[.602, .783]
SD	.730	[.634, .800]
TE	.948	[.924, .964]
A	.800	[.714, .857]

Note. Values represent agreement across five independent GPT coding runs. ICC(2, 1) = two-way random-effects, absolute-agreement, single-measure intraclass correlation coefficient; CI = confidence interval; SPA = spatial references; EP = entities present; SD = sensory descriptions; TE = thoughts and emotions; A = actions.

Run-to-run reliability of SEC was moderate, ICC(2,1) = .531, 95% CI [.419, .627], F(199, 796) = 8.49, p < .001. To further validate the GPT-coded narrative data, a subset of memories was independently coded by the researcher. Agreement between the GPT-generated and human-coded scores was then assessed for memory-specificity classification, ER, and SEC.

Human-GPT agreement for memory-specificity classification was assessed using Cohen's kappa for the 28 memories coded by both the human coder and GPT. Final GPT classifications were based on the modal category across the five runs. Agreement was limited, $\kappa = .276$, $z = 2.12$, $p = .034$, with 67.9% raw agreement. Disagreement was concentrated among memories classified as Extended by the human coder: GPT classified 9 of the 12 human-coded Extended memories as Specific. Neither coder classified any memories as Categoric/general.

Human-GPT agreement for total episodic richness was assessed for 28 memories using a two-way mixed-effects, absolute-agreement, single-measure ICC. The point estimate indicated good agreement, ICC(A,1) = .894, 95% CI [.446, .966], $F(27, 27) = 33.60$, $p < .001$, although the wide confidence interval indicated substantial uncertainty. As a complementary analysis, the Pearson correlation indicated a very strong linear association between human- and GPT-coded scores, $r = .953$. However, GPT scores were, on average, 5.61 points higher than human-coded scores, with a mean absolute difference of 6.25 points, suggesting that GPT closely tracked differences in episodic richness across memories but tended to assign higher absolute scores.

Human-GPT agreement for the ER subcomponents was assessed using two-way mixed-effects, absolute-agreement, single-measure intraclass correlation coefficients, ICC(A,1). As shown in Table 3, the ICC point estimates indicated good agreement for EP, TE, and A, moderate agreement for SD, and poor agreement for SPA. However, several confidence intervals were wide, indicating uncertainty in the precision of these estimates. GPT assigned higher mean scores than the human coder for SPA, EP, SD, and TE, with the largest differences observed for SPA and SD. In contrast, GPT assigned lower mean scores for A.

Table 3

Human-GPT Agreement for Episodic-Richness Subcomponents

Subcomponents	ICC(A,1)	95% CI	Human *M*	GPT *M*	Mean difference
SPA	.491	[.109, .737]	4.11	6.89	2.79
EP	.832	[.658, .920]	7.36	8.07	0.71
SD	.628	[.052, .851]	1.43	3.46	2.04
TE	.825	[.557, .925]	8.32	10.04	1.71
A	.779	[.569, .892]	13.04	11.39	-1.64

Note. ICC(A, 1) = two-way mixed-effects, absolute-agreement, single-measure intraclass correlation coefficient; CI = confidence interval; *M* = mean; GPT = GPT-5.5 Thinking; SPA = spatial references; EP = entities present; SD = sensory descriptions; TE = thoughts and emotions; A = actions. Mean difference was calculated as GPT minus human; positive values indicate higher GPT scores.

Human-GPT agreement for SEC was assessed in a 24-memory subset. Agreement was assessed using a two-way absolute-agreement single-score ICC. Agreement was moderate, ICC(A,1) = .710, 95% CI [.249, .884], $F(23, 23) = 8.66$, $p < .001$. GPT slightly undercounted

SECs relative to the human coder, with a mean difference of −0.46 and a mean absolute error of 0.46 SECs per memory.

Narrative Data Descriptives

Narrator pre- and post-session memory data were available for 24 participants, following the exclusion of four narrators. Session memory data were available for all 28 narrators and post-session memory data were available for all 28 listeners.

Memory Specificity Classification Descriptives

Memory specificity was predominantly classified as specific across roles, phases, and relationship conditions, ranging from 75.0% to 83.3%. Extended memories accounted for 16.7%-21.4% of classifications, whereas categoric memories accounted for 0%-8.3%.

Episodic Richness Descriptives

For both listeners and narrators, total episodic richness was descriptively similar across relationship conditions. Among narrators, scores differed more clearly across phases, with session narratives showing higher episodic richness than pre- and post-session memories (see Table 4).

Table 4

Descriptive Statistics for Total Episodic Richness Scores by Role, Phase and Condition (Mean and SD)

Phase	Condition	n	ER
Listener			
Post-session	Friend	28	22.68 (10.32)
	Stranger	28	20.50 (9.49)
Narrator			
Pre-session	Friend	24	27.21 (20.26)
	Stranger	24	26.38 (15.60)
	Baseline	24	29.33 (19.68)
Session	Friend	28	43.89 (18.82)
	Stranger	28	44.61 (12.70)
Post-session	Friend	24	20.96 (13.28)
	Stranger	24	22.50 (14.06)
	Baseline	24	23.50 (14.17)

Note. Values are M (SD). ER = total episodic richness; M = mean; SD = standard deviation; n = number of participants contributing data in each condition.

Subcomponents of ER

Across ER subcomponents, scores were descriptively similar between friend and stranger conditions within each role and phase. Narrator session narratives had higher TE (Ms = 15.11-16.18) and A scores (Ms = 11.75-12.14) than pre-session memories or post-session recollections. EP varied little across phases, whereas sensory descriptions were consistently least frequent and especially zero-heavy in listeners' post-session recollections (53.6% in the friend condition; 60.7% in the stranger condition).

Self-Event Connections

SEC scores were only relevant for narrators' memory data. SEC scores were generally low and frequently zero, with descriptively higher scores in the stranger than friend condition across phases (see Table 5).

Table 5

Descriptive Statistics for Self-Event Connections Scores by Phase and Condition (Mean and SD)

Phase	Condition	n	*SEC*
Pre-session	Friend	24	0.33 (0.64)
	Stranger	24	0.58 (0.83)
	Baseline	24	0.79 (1.06)
Session	Friend	28	0.50 (1.00)
	Stranger	28	1.00 (1.05)
Post-session	Friend	24	0.17 (0.48)
	Stranger	24	0.71 (0.81)
	Baseline	24	0.38 (0.65)

Note. Values are *M* (*SD*). SEC = self-event connections count; *M* = mean; *SD* = standard deviation; *n* = number of narrators contributing data in each condition.

Word Count

Because word counts were positively skewed, medians and interquartile ranges were used as the primary descriptive statistics. Narrator session narratives were descriptively longer than pre- and post-session memories, with similar medians in the friend and stranger conditions. Pre- and post-session medians were lower and similar across the friend, stranger, and baseline conditions. Listener post-session recollections were somewhat longer in the friend than stranger condition (see Table 6).

Table 6

Descriptive Statistics for Word Count by Role, Phase and Condition (Median and IQR)

Phase	Condition	n	Word Count
Listener			
Post-session	Friend	28	100.5 (70.5)
	Stranger	28	77.0 (67.5)
Narrator			
Pre-session	Friend	24	123.0 (118.5)
	Stranger	24	114.5 (52.5)
	Baseline	24	133.5 (95.3)
Session	Friend	28	253.0 (150.8)
	Stranger	28	255.0 (101.8)
Post-session	Friend	24	83.5 (89.0)
	Stranger	24	90.5 (58.5)
	Baseline	24	95 (92.5)

Note. Values are Mdn (IQR). Mdn = median; IQR = interquartile range; n = number of participants contributing data in each condition.

Subjective Ratings of Vividness and Emotional Intensity Descriptives

Session ratings were available for all 28 listeners and for 27 of the 28 narrators; one narrator had missing data for these measures. Listeners rated memories heard from friends as more emotionally intense than memories heard from strangers. Vividness ratings were similar across conditions in listeners. Narrators rated memories shared with strangers as somewhat more emotionally intense than memories shared with friends. Vividness ratings were similar across conditions in narrators. See Table 7.

Table 7

Descriptive Statistics for Ratings of Emotional Intensity and Vividness by Role and Condition (Mean and SD)

Role	Condition	n	Emotional Intensity	Vividness
Listener	Friend	28	3.62 (1.14)	4.10 (0.79)
	Stranger	28	3.10 (1.15)	4.14 (0.88)
Narrator	Friend	27	3.85 (0.93)	4.33 (0.56)
	Stranger	27	4.18 (0.77)	4.22 (0.63)

Note. Values are *M* (*SD*). MEQ = Memory Experiences Questionnaire; *M* = mean; *SD* = standard deviation; *n* = number of participants contributing data in each condition. Emotional Intensity and Vividness are MEQ subscale mean scores ranging from 1 to 5, with higher scores indicating greater emotional intensity or vividness.

Central Analyses

No substantial violations of model assumptions or diagnostic concerns were identified unless otherwise stated.

Memory Narration: Episodic Richness and Self-Event Connections

H1a predicted greater episodic richness for memories narrated to strangers than to friends. Contrary to this prediction, relationship condition did not significantly affect episodic richness, $F(1, 27) = 0.04$, $p = .834$, $B = -0.71$, 95% CI [−7.64, 6.21]. H1a was therefore not supported.

H1b predicted that SEC would differ between friend and stranger narrations. A significant relationship effect was found, $\chi^2(1) = 4.76$, $p = .029$, RR = 0.50, 95% CI [0.26, 0.95], corresponding to an estimated 50% lower rate of SEC in friend than stranger narrations. H1b was therefore supported.

Exploratory analyses examined narrators' subjective vividness and emotional intensity during narration. Vividness did not differ significantly between conditions, $F(1, 26) = 0.63$, Holm-adjusted $p = .436$. Emotional intensity was lower in the friend than stranger condition, $B = -0.33$, $SE = 0.16$, $F(1, 26) = 4.42$, but did not remain significant after Holm adjustment, adjusted $p = .091$. After controlling for pre-session episodic richness, neither within-narrator vividness, $B = 1.54$, 95% CI [−8.18, 11.26], adjusted $p = .759$, nor emotional intensity, $B = 3.30$, 95% CI [−3.32, 9.91], adjusted $p = .677$, was significantly associated with session episodic richness. Together, these analyses provided no evidence that the absence of a friend-stranger difference in episodic richness was attributable to differences in narrators' subjective vividness or emotional intensity during sharing.

Supplementary analyses examined whether narrative length accounted for the H1 findings. Narrative length did not differ significantly between friend and stranger conditions, $F(1, 27) = 0.62$, $p = .437$, although longer narratives contained greater episodic richness, $B = 35.52$, $SE = 3.51$, $F(1, 41.77) = 102.69$, $p < .001$. Adjusting for narrative length did not alter the null relationship effect on episodic richness, $F(1, 27.15) = 0.21$, $p = .650$, $B = 1.10$, 95% CI [−3.82, 6.02]. The relationship effect on SEC also remained significant after adjustment for word count, $\chi^2(1) = 4.25$, $p = .039$, RR = 0.52, 95% CI [0.27, 0.99].

Because pre-session SEC was descriptively higher for stranger than friend memories, an additional sensitivity analysis adjusted H1b for pre-session SEC. The condition effect was attenuated and no longer statistically significant, $\chi^2(1) = 1.47$, $p = .225$, RR = 0.65, 95% CI [0.32, 1.32]. Pre-session SEC was positively associated with session SEC, $B = 0.87$, $SE = 0.22$, $z = 4.06$, $p < .001$. Accordingly, H1b was robust to narrative-length adjustment but not to pre-session SEC adjustment. No analogous baseline-adjusted analysis was conducted for H1a because the primary ER effect was already null and pre-session ER was descriptively similar across conditions.

Pre-Post Changes in Episodic Richness and Self-Event Connections

For episodic richness, the phase × condition interaction was not statistically significant, $F(2, 46) = 0.89$, $p = .418$, indicating that pre-post change did not differ across the friend, stranger, and non-shared baseline conditions. H2a was therefore not supported. Planned contrasts likewise showed no significant difference in change between the friend and baseline conditions, $B = -0.42$, $SE = 1.90$, $t(46) = -0.22$, 95% CI [−4.24, 3.41], Holm-adjusted $p = .827$ (H2b), or between the stranger and baseline conditions, $B = 1.96$, $SE = 1.90$, $t(46) = 1.03$, 95% CI [−1.87, 5.78], Holm-adjusted $p = .616$ (H2c).

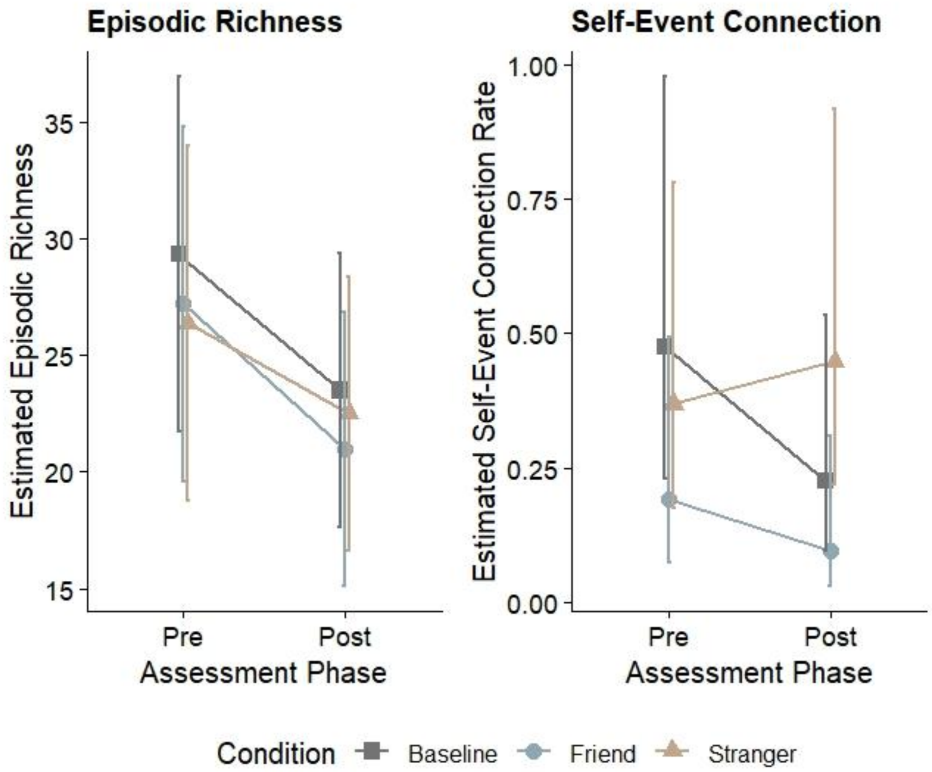
There was, however, a significant overall effect of phase, $F(1, 23) = 6.52$, $p = .018$, reflecting a general decrease in episodic richness from pre- to post-session. Because the phase × condition interaction was nonsignificant, the magnitude of this decrease did not differ significantly across conditions.

For SEC, the phase × condition interaction was also not statistically significant, $\chi^2(2) = 3.59$, $p = .166$, indicating that pre-post change did not differ across conditions. H2d was therefore not supported. Planned contrasts showed no significant difference in change between the friend and baseline conditions, $B = 0.05$, $SE = 0.73$, $z = 0.07$, Holm-adjusted $p = .941$, ratio of rate ratios = 1.06, 95% CI [0.25, 4.45] (H2e), or between the stranger and baseline conditions, $B = 0.94$, $SE = 0.54$, $z = 1.74$, Holm-adjusted $p = .165$, ratio of rate ratios = 2.56, 95% CI [0.89, 7.42] (H2f). Descriptively, SEC decreased in the friend and baseline conditions and increased slightly in the stranger condition, although these differences in change were not statistically significant.

The model-estimated pre-post patterns are shown in Figure 1, illustrating broadly similar declines in episodic richness across conditions and greater descriptive divergence in SEC.

Figure 1

Model-Estimated Pre-Post Changes in Episodic Richness and Self-Event Connections Rate by Condition



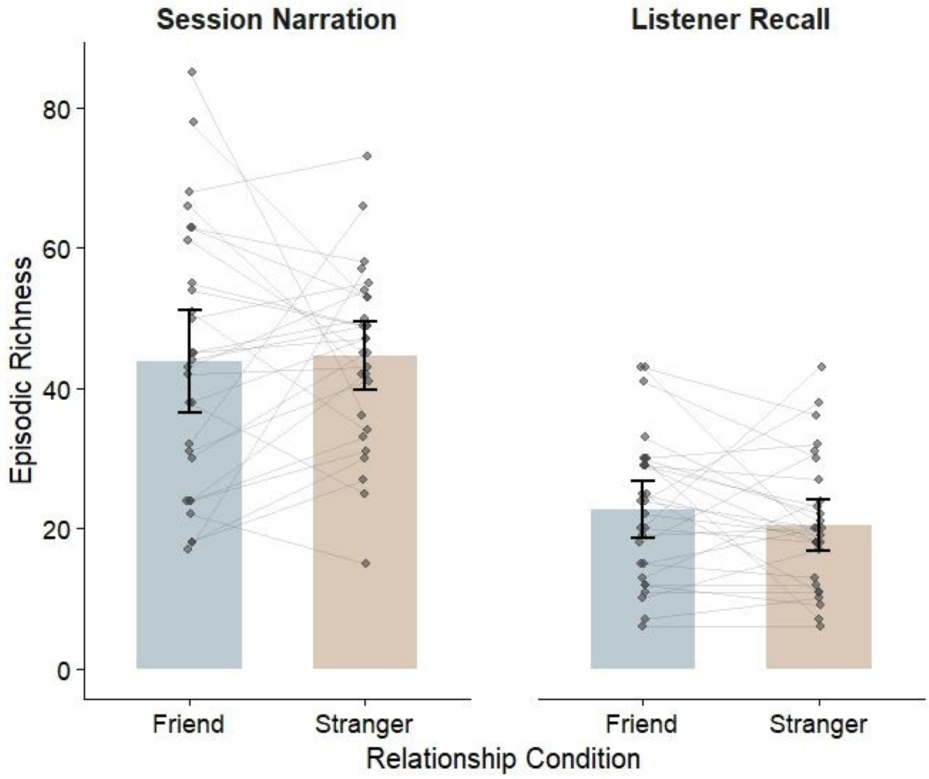
Note. Points represent model-estimated values, lines connect pre- and post-session estimates within each condition, and error bars represent 95% confidence intervals. Episodic richness estimates were derived from the linear mixed-effects phase × condition model. Self-event connection rates were derived from the corresponding Poisson generalized linear mixed-effects model and are presented on the response scale.

Listener Recall

H3 predicted greater listener-recall episodic richness for memories narrated by friends than by strangers. Listener recall was descriptively higher in the friend condition (Table 4), but the relationship effect was not statistically significant, $F(1, 27) = 1.22$, $p = .280$, $B = 2.18$, $SE = 1.98$, 95% CI [−1.87, 6.23]. H3 was therefore not supported. Figure 2 presents condition means and paired observations for session narration and subsequent listener recall.

Figure 2

Episodic Richness During Session Narration and Subsequent Listener Recall by Relationship



Note. Bars represent condition means with 95% confidence intervals. Points represent individual observations, and lines connect Friend and Stranger observations from the same participant.

In a secondary analysis adjusting for the episodic richness of the narration heard, narration richness showed a positive but nonsignificant association with listener recall, B = 0.14, SE = 0.07, $F(1, 41.79) = 3.52$, $p = .068$. The adjusted friend-stranger difference also remained nonsignificant, B = 2.54, SE = 1.91, 95% CI [−1.39, 6.47], $F(1, 26.55) = 1.77$, p = .195, leaving the H3 conclusion unchanged.

Exploratory analyses found no significant association between within-listener vividness and subsequent recall, B = −0.79, SE = 1.80, 95% CI [−4.31, 2.73], Holm-adjusted p = .662, or between emotional intensity and recall, B = 1.93, SE = 1.75, 95% CI [−1.50, 5.37], Holm-adjusted p = .560. Thus, listeners' subjective vividness and emotional intensity while receiving the memories were not clearly associated with subsequent episodic richness.

Discussion

The present study examined whether the relationship between a narrator and listener influences how autobiographical memories are communicated during sharing, whether any relationship-dependent effects persist into subsequent recall, and whether relationship context influences listeners' later reconstruction of the narrated memory. A secondary methodological aim was to evaluate the reliability and human agreement of GPT-based narrative coding. Overall, the findings provided limited evidence that relational closeness influenced the episodic content of autobiographical remembering. Across the hypotheses concerning episodic richness, relationship condition did not significantly influence the amount of episodic detail communicated during narration, the degree of pre-post change in narrators' memories, or the amount of episodic information subsequently recalled by listeners. In contrast, a relationship effect emerged for self-event connections during narration: narrators produced more explicit connections between the remembered event and the self when speaking to strangers than when speaking to friends. This effect remained after accounting for narrative length but was no longer statistically significant after adjusting for pre-session self-event connections. Furthermore, there was no evidence that this relationship difference resulted in a corresponding relationship-specific change in later recall. Taken together, the findings suggest that, within the relationship contexts and task used in the present study, relational closeness had little measurable influence on the amount of episodic information communicated or subsequently reconstructed but may have influenced how autobiographical experiences were framed in relation to the self during social communication.

Reliability and Human-GPT Agreement

The secondary methodological aim of evaluating GPT-5.5 Thinking for autobiographical narrative coding is considered first, as the substantive analyses relied on GPT-generated scores. Accordingly, the discussion begins by considering the reliability of the model's coding and its agreement with human coding before turning to the substantive findings.

Large language models (LLMs) may increase the efficiency of narrative coding, which is otherwise highly time-consuming. In the present study, GPT-5.5 Thinking made it possible to include extensive narrative analyses within the available timeframe. However, this approach also required evaluation of whether the model produced consistent scores across repeated runs and whether its coding corresponded with human coding.

Run-to-run reliability was strongest for memory-specificity classification and total episodic richness but varied across the episodic-richness subcomponents and was moderate for self-event connections (SECs). Entities present (EP) and sensory descriptions (SD) showed the lowest reliability. One possible explanation is that these categories placed greater demands on the model. EP was assessed across the full narrative while repeated references had to be excluded, which may have increased inconsistency across runs. SD may also have been difficult to distinguish from emotional expressions, despite explicit instructions separating the two categories. In addition, SD was relatively infrequent and was particularly zero-heavy in listeners' post-session recollections, which may have made reliability estimates more sensitive to small differences in coding across runs. The low run-to-run agreement might therefore also reflect the data itself, and its frequent zero scores.

SEC involved similar interpretive demands because the model had to identify explicit links between an event and the narrator's self. The large number of zero scores may partly reflect the conservative sentence-level coding framework, which prevented the model from inferring connections across the broader narrative. Although this approach may have missed some connections, it was intended to limit overinterpretation and support more consistent coding. Overall, the findings suggest that LLM reliability may vary according to the complexity and interpretive demands of the coding task.

Human-GPT agreement was generally weaker and less precise than GPT's internal run-to-run reliability. Agreement was limited for memory-specificity classification and moderate for SEC. Total episodic richness showed good agreement, although agreement varied considerably across its subcomponents.

The uneven distribution of memory-specificity categories may partly have affected the kappa estimate, as most memories in the validation subset were classified as specific and none were classified as categoric. However, disagreement was also concentrated among extended memories, which GPT frequently classified as specific, suggesting that the model and human coder may have applied the distinction between specific and extended memories differently.

The human-validation analyses were based on relatively small subsets of the available narratives, which was reflected in several wide confidence intervals. The findings therefore provide an initial indication of human-GPT correspondence but do not allow precise estimates of agreement across the full dataset. Furthermore, human-GPT agreement was assessed against coding conducted by a single researcher, whereas narrative coding typically involves at least two independent coders. Accordingly, both the methodological findings and the substantive analyses based on GPT-coded scores should be interpreted with caution.

The present findings provide partial support for previous research on automated analysis of autobiographical narratives. Consistent with Oded et al. (2026) and van Genugten and Schacter (2024), GPT showed that language models can identify psychologically meaningful narrative features, with good human-GPT correspondence for total episodic richness. However, agreement varied across subcomponents, suggesting that performance depends on both the structure of the narratives and the demands of the coding framework.

Human-GPT agreement for memory specificity was less favorable than the high classification performance reported by Mistica et al. (2024), although the findings are not directly comparable because different models, datasets, and performance measures were used. In the present study, disagreement was concentrated among extended memories, which GPT frequently classified as specific. Interestingly, Mistica et al. (2024) also identified ambiguous duration and event frequency as sources of classification error. Overall, the findings suggest that general-purpose LLMs may be useful for narrative coding, but their performance should be evaluated separately for each framework rather than assumed from strong performance on other narrative tasks.

These measurement properties informed how the GPT-coded variables were used and interpreted in the substantive analyses. Total episodic richness, which showed good human-GPT agreement and good run-to-run reliability, was retained as the primary measure of episodic narrative content when comparing experimental conditions, rather than relying on its individual subcomponents as primary outcomes. In contrast, findings involving SEC warrant greater caution given the more moderate reliability and less precise human-GPT agreement observed for this measure. Memory specificity was used primarily to characterize the distribution of memory types in the dataset rather than as a central outcome, particularly given the limited human-GPT agreement for this classification. Consistent with this approach, the individual subcomponents of episodic richness were not interpreted as separate substantive outcomes. Taken together, the reliability and validation analyses provide a basis for interpreting the GPT-coded data while indicating that confidence should vary across coding frameworks according to their respective measurement properties.

Relationship Context and Episodic Reconstruction

Across the different phases of memory recall examined in the present study, there was little evidence that relational closeness systematically influenced episodic richness. Memories narrated to strangers were descriptively richer than those narrated to friends, as predicted by H1a, but this difference was small and not statistically significant and remained nonsignificant after adjustment for narrative length. Similarly, sharing memories with friends or strangers did not produce differential pre-post changes in episodic richness relative to the non-shared baseline condition. Finally, listeners did not subsequently recall significantly more episodic information from memories narrated by friends than from memories narrated by strangers. Considered together, these findings suggest that, within the present paradigm, relational closeness was not sufficient to produce detectable differences in the amount of episodic information communicated, subsequently reconstructed by narrators, or retained by listeners.

This broader pattern contrasts with previous evidence that communicative context can shape autobiographical recall. Eckardt et al. (2024), for example, found differences in narrative content and linguistic characteristics when participants recalled an experience as though sharing it with a friend rather than under neutral recall instructions. An important distinction is that, for the immediate narration comparison in the present study, both relationship conditions involved genuine interpersonal communication. The contrast was therefore not between socially oriented and neutral retrieval, but between two social retrieval contexts differing in relational closeness. One possibility is that adopting a communicative orientation toward recall has a stronger influence on episodic elaboration than changing the familiarity of the audience once recall is already socially situated. This interpretation remains plausible given the methodological differences between the studies and it may help explain why the clear difference in interpersonal closeness between friends and strangers was not followed by a detectable difference in total episodic richness.

The findings also refine expectations based on common ground and shared reality: friends may require less contextual explanation because of their shared interpersonal history, whereas strangers may require more information to establish mutual understanding. Greater episodic elaboration was consequently expected when narrating to strangers. However, the present findings suggest that these differences in communicative demands were not sufficient to alter the overall amount of episodic information expressed. The exploratory subjective-rating analyses were consistent with this pattern: within-narrator variation in vividness and emotional intensity was not significantly associated with session episodic richness after accounting for pre-session richness and relationship condition. Accordingly, neither relational closeness nor the subjective experience of the memories clearly accounted for variation in overall episodic elaboration during narration. Selwood et al. (2020) similarly found that relationship type did not consistently produce broad differences in the qualities of autobiographical recall. Instead, collaboration itself had more consistent effects on memory qualities, whereas relationship type was reflected more strongly in the communication processes through which collaborative remembering unfolded. This distinction is relevant to the present study because Selwood et al.'s participants could actively contribute to and co-construct the remembered event, whereas listeners in the present narrator-listener paradigm had a deliberately restricted role. The absence of a friend-stranger difference in episodic richness may therefore partly reflect the limited opportunity for collaborative processes to shape recall. More broadly, Selwood et al.'s findings suggest that active collaboration may exert a stronger influence on broad qualities of autobiographical remembering than relationship type alone.

The pre-post findings extend this pattern by indicating that relationship context was also not associated with detectable differences in the subsequent episodic reconstruction of the memories. Unlike the immediate narration comparison, these analyses included a non-shared baseline memory, allowing changes following social narration to be distinguished from changes associated with repeated retrieval more generally. Neither memories shared with friends nor those shared with strangers changed differently in episodic richness from the non-shared baseline memory. This finding is broadly consistent with Grysman et al. (2023), who found that changes in narrative recall did not produce a corresponding persistent effect on later recall. At the same time, sociocultural accounts of autobiographical memory emphasize that the organization and elaboration of autobiographical narratives can be shaped through social interaction and repeated reminiscing (Fivush et al., 2011). The present findings therefore do not suggest that social interaction is irrelevant to autobiographical memory, but rather that the relationship context of a single sharing episode may not be sufficient to produce detectable lasting differences in episodic richness. Together, these findings highlight an important distinction between contextual influences on memory expression at a particular retrieval occasion and durable changes in how a memory is subsequently reconstructed. Although autobiographical retrieval is reconstructive and sensitive to current goals and context, contextual influences present during one retrieval episode may not necessarily become incorporated into later representations of the event.

Despite the absence of relationship-specific change, episodic richness decreased significantly from pre- to post-session across conditions. Importantly, this decline was also present for the non-shared baseline memory. The pattern therefore points to a process common to repeated retrieval rather than an effect produced specifically by social narration or relationship context. Several mechanisms could plausibly contribute to this decrease. Having recently described the same event may have reduced the extent to which participants elaborated their memories during subsequent recall, while reduced task engagement or fatigue across repeated assessments could also have contributed. Alternatively, the decrease may reflect changes in the accessibility or representation of episodic information itself. The present design cannot distinguish between these explanations, and the decline should therefore not be attributed to any single mechanism.

The timing of the retrieval phases is important for interpreting both the overall decline in episodic richness and the absence of relationship-specific effects. In this study, participants had already written about each autobiographical memory the day before the experimental session, meaning that all memories had been recently activated and structured prior to the social manipulation. Because of this prior rehearsal, the subsequent narration and recall tasks took place in a context where the memories were already accessible. This prior exposure may have reduced the extent to which interacting with either a friend or a stranger could further shape later recall.

Accordingly, the findings do not rule out the possibility that relational context may influence autobiographical memory over time. Rather, within the present design, which involved recently rehearsed memories and relatively short intervals between retrievals, there was no clear evidence of relationship-based differences in episodic richness.

Self-Event Connections and the Social Framing of Memory

The clearest relationship-dependent finding concerned SEC. Narrators produced significantly more SECs when speaking to strangers than to friends, and this difference remained after accounting for narrative length. This suggests that the effect was not simply attributable to narrators speaking for longer in the stranger condition. In contrast to episodic richness, which reflects the amount of event-specific information communicated, SECs capture how explicitly an autobiographical experience is linked to broader aspects of the self. Relationship context may therefore have influenced how memories were framed or interpreted in relation to the self, rather than the overall amount of episodic detail retrieved.

One possible explanation concerns the different levels of shared history between friends and strangers. Friends already possess considerable knowledge about one another's personalities, experiences, and identities, meaning that narrators may not need to explicitly explain why a memory is personally meaningful. When speaking to a stranger, however, this shared background is absent. Narrators may therefore be more likely to explain what an experience meant to them, how it affected them, or what it reveals about who they are. In this way, SECs may serve a communicative function by helping to establish personal context and shared understanding with an unfamiliar listener.

This interpretation should nevertheless be treated cautiously. Although the friend-stranger difference remained after adjustment for narrative length, it was attenuated and no longer statistically significant after accounting for the number of SECs already present in the corresponding pre-session memories. Pre-session SEC scores were also positively associated with session SEC scores. This suggests that the observed condition difference may partly reflect pre-existing differences in the self-related content of the memories rather than a difference that emerged specifically during narration. Consistent with this, the pre-post analyses showed no significant relationship-specific change in SECs. SECs were also relatively infrequent across conditions, which may partly reflect the use of strong positive memories rather than memories selected specifically for their relevance to identity or self-understanding. Finally, the comparatively lower human-GPT agreement for SEC coding introduces additional uncertainty. Taken together, the primary finding is consistent with the possibility that relationship context influenced how self-related meaning was communicated, but the baseline-adjusted analysis limits the extent to which this difference can be attributed specifically to the social context of narration.

Listener Reconstruction of Shared Memories

The third research question examined whether listeners reconstructed narrated memories differently depending on whether the narrator was a friend or stranger. Although listener recall was descriptively higher in the friend condition, the difference was not statistically significant, and the confidence interval was relatively wide. The findings therefore provide no clear evidence that relational closeness increased the amount of episodic information subsequently recalled.

This conclusion was unchanged after adjusting for the episodic richness of the narration heard. Narration-level richness was positively, but not significantly, associated with listener recall, suggesting that more detailed narrations may provide more information for later reconstruction, although the present data were too imprecise to establish this association. Likewise, within-listener variation in vividness and emotional intensity was not significantly related to subsequent episodic richness.

Several aspects of the design may have reduced the opportunity for a relationship effect to emerge. Listeners recalled both narratives shortly after hearing them, when both were likely to remain highly accessible. The explicit recall task may also have encouraged strategic attention to both stories, making the procedure more similar to a memory test than to everyday conversational listening.

Despite the nonsignificant condition effect, the listener analysis addresses a gap in previous narrator-listener research. Eckardt et al. (2024) examined how narrated memories influenced listeners' subsequent decision-making but did not directly assess what listeners retained of the narrated event. By measuring listeners' later reconstruction of the narratives, the present study extends this work and suggests that, under conditions of relatively immediate recall and constrained interaction, relational closeness alone was not associated with detectable differences in the episodic richness of listeners' recall.

Strengths and Limitations

Several features strengthen the study. Using actual friends and previously unfamiliar partners provided a direct manipulation of relationship context, which was confirmed by the closeness manipulation check. Including a non-shared baseline memory helped distinguish effects of social narration from changes due to repeated retrieval. Assessing both narrators and listeners captured both sides of the memory-sharing interaction, while repeated GPT coding with human validation allowed the reliability of automated coding to be evaluated directly. Together, these features enabled the study to examine interpersonal aspects of autobiographical memory sharing within a relatively controlled design.

A central limitation concerns the short intervals between retrieval phases. Narrators wrote each memory the day before the experimental session, meaning that all memories had been recently retrieved and rehearsed before narration and post-session recall. Any observed changes therefore occurred from a recently rehearsed rather than unrehearsed baseline. Although Eckardt et al. (2024) used a similarly short interval, their delay followed the encoding of a new event, whereas the present study involved older autobiographical memories that had recently been retrieved and written about. Therefore, despite reducing immediate carryover, the one-day interval may still have influenced subsequent recall.

The short interval may also have limited the emergence of relationship-specific differences in later recall. Listeners recalled both memories shortly after hearing them, when they were likely still highly accessible. A longer retention interval, such as the four-week delay used by Grysman et al. (2023), may have provided a stronger test of longer-term reconstruction, but would have increased attrition risk and would have required a longer data-collection period than was feasible within the scope of the project. The present design therefore favored feasibility and experimental control over sensitivity to longer-term memory change.

A further limitation concerns the restricted conversational reciprocity of the sharing sessions. Although listeners could respond naturally, they could not interrupt, ask questions, or elaborate, making the task closer to socially situated narration than unrestricted conversational remembering. Greater reciprocity may have increased ecological validity and co-construction, but at the cost of experimental control and comparability across conditions. The design therefore reflects a trade-off between interpersonal realism and comparability.

An additional limitation concerns the sample size, particularly for the narrator pre-post analyses, which were based on 24 participants with complete data. Several estimates consequently had relatively wide confidence intervals, limiting the precision with which the effects could be estimated. This uncertainty should be considered when interpreting the nonsignificant findings but does not imply that the effects would necessarily have reached statistical significance in a larger sample.

Finally, due to the sampling method, the researcher was familiar with a substantial proportion of participants. Due to this familiarity, although off camera and muted during narration sessions, the presence of a familiar researcher may have influenced the perceived interaction context, particularly in the stranger condition, and potentially reduced the distinction between conditions. As this was not directly assessed, any effect on narration remains uncertain.

Future Directions

Future research should adopt more longitudinal and ecologically valid approaches to conversational remembering. A single sharing episode may be insufficient to produce detectable lasting changes in autobiographical memory, whereas repeated retellings across multiple social interactions may gradually shape how events are reconstructed over time. Extending the present paradigm could therefore involve repeated in-person memory-sharing sessions over longer intervals and allow more reciprocal interaction between narrator and listener, similar to collaborative remembering paradigms such as Selwood et al. (2020). Larger samples would also improve statistical precision, particularly given the wide confidence intervals observed in the present study.

Finally, future work should refine and standardize LLM-based narrative coding, including prompting, model reporting, coding procedures, and validation. Although LLMs are developing faster than methodological standards can easily keep pace with, greater standardization would improve reproducibility, comparison across studies, and conclusions about their reliable use in psychological research.

Conclusion

In conclusion, the present study raises the possibility that social context may influence autobiographical remembering in more selective ways than expected while a single episode of memory sharing may not be enough to produce persistent changes in later recall. More lasting reconstruction may instead develop through repeated acts of remembering and sharing over time. The study also supports the potential of LLMs for coding narrative data, while showing that their performance differs across coding frameworks. LLM-based coding may therefore be a useful methodological tool, but its reliability and agreement with human coding should be evaluated for each framework before strong conclusions are drawn from its output.

Taken together, these findings contribute to understanding both the social conditions under which autobiographical memories are reconstructed and the methodological tools available for studying those processes.

Declaration of Generative AI Usage

During the preparation of this thesis, the author used ChatGPT (GPT-5.5) for language and grammatical support, the organization of information, and assistance with developing and troubleshooting R code. All analytical decisions were made by the author, and all AI-generated output was critically reviewed before use. The author assumes full responsibility for the content and conclusions of the thesis.

References

Alea, N., & Bluck, S. (2003). Why are you telling me that? A conceptual model of the social function of autobiographical memory. *Memory*, *11*(2), 165–178. https://doi.org/10.1080/741938207

Alea, N., & Bluck, S. (2007). I'll keep you in mind: The intimacy function of autobiographical memory. *Applied Cognitive Psychology*, *21*(8), 1091–1111. https://doi.org/10.1002/acp.1316

Beike, D. R., Brandon, N. R., & Cole, H. E. (2016). Is sharing specific autobiographical memories a distinct form of self-disclosure? *Journal of Experimental Psychology: General, 145*(4), 434–450. https://doi.org/10.1037/xge0000143

Bosch, S. E., Jehee, J. F. M., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *The Journal of Neuroscience, 34*(22), 7493–7500. https://doi.org/10.1523/JNEUROSCI.0805-14.2014

Boytos, A. S., & Costabile, K. A. (2023). Shared reality, memory goal satisfaction, and psychological well-being during conversational remembering. *Memory, 31*(5), 689–704. https://doi.org/10.1080/09658211.2023.2188643

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107*(2), 261–288. https://doi.org/10.1037/0033-295X.107.2.261

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition, 22*(5), 491–529. https://doi.org/10.1521/soco.22.5.491.50768

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). Shared reality: Experiencing commonality with others' inner states about the world. *Perspectives on Psychological Science, 4*(5), 496–521. https://doi.org/10.1111/j.1745-6924.2009.01161.x

Eckardt, D., Helion, C., Schmidt, H., Chen, J., & Murty, V. P. (2024). Storytelling changes the content and perceived value of event memories. *Cognition, 251*, Article 105884. https://doi.org/10.1016/j.cognition.2024.105884

Fivush, R., Habermas, T., Waters, T. E. A., & Zaman, W. (2011). The making of autobiographical memory: Intersections of culture, narratives, and identity. *International Journal of Psychology, 46*(5), 321–345. https://doi.org/10.1080/00207594.2011.596541

Ford, J. H., Rubin, D. C., & Giovanello, K. S. (2014). Effects of task instruction on autobiographical memory specificity in young and older adults. *Memory, 22*(6), 722–736. https://doi.org/10.1080/09658211.2013.820325

Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the closeness of relationships: A comprehensive evaluation of the "Inclusion of the Other in the Self" scale. *PLOS ONE, 10*(6), Article e0129478. https://doi.org/10.1371/journal.pone.0129478

Grysman, A., Baime, M., & Cantor, E. (2023). Listeners' effects on autobiographical memory for recent events. *Memory, 31*(10), 1425–1436. https://doi.org/10.1080/09658211.2023.2270778

Harris, C. B., Rasmussen, A. S., & Berntsen, D. (2014). The functions of autobiographical memory: An integrative approach. *Memory*, *22*(5), 559–581. https://doi.org/10.1080/09658211.2013.806555

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences of the United States of America, 104*(5), 1726–1731. https://doi.org/10.1073/pnas.0610561104

Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications, 6*, Article 7462. https://doi.org/10.1038/ncomms8462

Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory: Context-dependent updating. *Learning & Memory*, *15*(8), 574–579. https://doi.org/10.1101/lm.1022308

Jiang, T., Chen, Z., & Sedikides, C. (2020). Self-concept clarity lays the foundation for self-continuity: The restorative function of autobiographical memory. *Journal of Personality and Social Psychology, 119*(4), 945–959. https://doi.org/10.1037/pspp0000259

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Leng, J., Ma, B., & Guo, Q. (2025). Brief mindfulness practice influences the self-reference effect by caring for others. *Scientific Reports*, *15*(1), Article 43468. https://doi.org/10.1038/s41598-025-31049-3

Lepsius, M. (2026). The brain edits the past: Neural systems of memory reconstruction and self-stability. *New Ideas in Psychology, 80*, Article 101221. https://doi.org/10.1016/j.newideapsych.2025.101221

Liu, W., Deng, L., & Yang, H. (2025). Embracing strengths and avoiding weaknesses: A meta-analysis of the mnemic neglect effect. *Memory, 33*(2), 278–294. https://doi.org/10.1080/09658211.2024.2433680

Marlatte, H., Beaton, D., Adler-Luzon, S., Abo-Ahmad, L., & Gilboa, A. (2022). Scene construction and spatial processing in post-traumatic stress disorder. *Frontiers in*

Behavioral Neuroscience, 16, Article 888358. https://doi.org/10.3389/fnbeh.2022.888358

McLean, K. C., & Fournier, M. A. (2008). The content and processes of autobiographical reasoning in narrative identity. *Journal of Research in Personality, 42*(3), 527–545. https://doi.org/10.1016/j.jrp.2007.08.003

Mistica, M., Haylock, P., Michalewicz, A., Raad, S., Fitzgerald, E., & Hitchcock, C. (2024). A natural language model to automate scoring of autobiographical memories. *Behavior Research Methods, 56*(7), 6707–6720. https://doi.org/10.3758/s13428-024-02385-5

Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*(6797), 722–726. https://doi.org/10.1038/35021052

Oded, N., Rubin, M., Genzer, S., & Perry, A. (2026). Reading between the lines: LLMs match or exceed human empathic accuracy using text alone. *Computers in Human Behavior Reports, 22*, Article 100994. https://doi.org/10.1016/j.chbr.2026.100994

Posit team. (2026). *RStudio: Integrated development environment for R*. Posit Software, PBC

R Core Team. (2026). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://doi.org/10.32614/R.manuals

Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *eLife, 5*, Article e18260. https://doi.org/10.7554/eLife.18260

Rossignac-Milon, M., Bolger, N., Zee, K. S., Boothby, E. J., & Higgins, E. T. (2021). Merged minds: Generalized shared reality in dyadic relationships. *Journal of Personality and Social Psychology, 120*(4), 882–911. https://doi.org/10.1037/pspi0000266

Rossignac-Milon, M., & Higgins, E. T. (2018). Epistemic companions: Shared reality development in close relationships. *Current Opinion in Psychology, 23*, 66–71. https://doi.org/10.1016/j.copsyc.2018.01.001

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one's faults. *European Review of Social Psychology, 27*(1), 1–62. https://doi.org/10.1080/10463283.2016.1183913

Selwood, A., Harris, C. B., Barnier, A. J., & Sutton, J. (2020). Effects of collaboration on the qualities of autobiographical recall in strangers, friends, and siblings: Both remembering partner and communication processes matter. *Memory*, *28*(3), 399–416. https://doi.org/10.1080/09658211.2020.1727521

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243. https://doi.org/10.1162/jocn.1992.4.3.232

St. Jacques, P. L., & Schacter, D. L. (2013). Modifying memory: Selectively enhancing and updating personal memories for a museum tour by reactivating them. *Psychological Science, 24*(4), 537–543. https://doi.org/10.1177/0956797612457377

Sunet. (n.d.). *Teknisk dokumentation*. Sunet Wiki. Retrieved September 11, 2026, from https://wiki.sunet.se/spaces/scribe/pages/287801803/Teknisk+dokumentation

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

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114 van Genugten, R. D. I., & Schacter, D. L. (2024). Automated scoring of the Autobiographical Interview with natural language processing. *Behavior Research Methods, 56*(3), 2243–2259. https://doi.org/10.3758/s13428-023-02145-x

Wagner, U., Higgins, E. T., Axmacher, N., & Echterhoff, G. (2024). Biased memory retrieval in the service of shared reality with an audience: The role of cognitive accessibility. *Journal of Experimental Psychology: General, 153*(6), 1605–1627. https://doi.org/10.1037/xge0001580

Wank, A. A., Mehl, M. R., Andrews-Hanna, J. R., Polsinelli, A. J., Moseley, S., Glisky, E. L., & Grilli, M. D. (2020). Eavesdropping on autobiographical memory: A naturalistic observation study of older adults' memory sharing in daily conversations. *Frontiers in Human Neuroscience, 14*, Article 238. https://doi.org/10.3389/fnhum.2020.00238

Wei, Y., Zhang, L., Lian, T., Chen, Y., Liao, S., Guo, Q., & Hu, P. (2024). Happy storytelling promotes emotional contagion and interpersonal closeness. *Current Psychology*, *43*(4), 2993–3001. https://doi.org/10.1007/s12144-023-04544-x

Wilson, A., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory*, *11*(2), 137–149. https://doi.org/10.1080/741938210

Yang, C. M., Duncan, K. D., & Barense, M. D. (2026). Memory strength at reactivation, not memory age, governs prediction error-driven updating of naturalistic event memory. *Philosophical Transactions of the Royal Society B: Biological Sciences, 381*(1954), Article 20250244. https://doi.org/10.1098/rstb.2025.0244

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. *Cerebral Cortex*, *27*(10), 4988–5000. https://doi.org/10.1093/cercor/bhx202

Appendix A

Memory Experiences Questionnaire (MEQ)

Narrator MEQ:

Subscale: Vividness

Instructions: You have just described a personal memory. Please indicate how vivid the memory was for you while recalling and narrating it. Use the scale below to indicate your agreement with each statement about the clarity and vividness of your memory.

1. My memory for this event is clear.
2. My memory for this event is very vivid.
3. My memory for this event is very detailed.
4. My memory for this event is dim. (R)
5. My memory for this event is very vague. (R)
6. My memory for this event is sketchy. (R)

Subscale: Emotional Intensity

Instructions: You have just described a personal memory. Please indicate how emotionally intense the memory felt for you while recalling and narrating it. Use the scale below to indicate your agreement with each statement about the emotional intensity of your memory.

1. As I am remembering the experience now, my feelings are very intense.
2. My emotions are very intense concerning this event.
3. The memory of this event evokes powerful emotions.
4. I do not remember having particularly strong emotions at the time of this event. (R)
5. I do not have strong emotions about this memory. (R)
6. This memory does not evoke strong emotions in me. (R)

Memory Experiences Questionnaire (MEQ)

Listener MEQ:

Subscale: Vividness

Instructions: You have just listened to a narrator describe a personal memory. Please indicate how you experienced the vividness of the memory as it was told. Use the scale below to indicate your agreement with each statement.

1. The memory as it was told to me was clear.
2. The memory as it was told to me was very vivid.
3. The memory as it was told to me was very detailed.
4. The memory as it was told to me was dim. (R)
5. The memory as it was told to me was very vague. (R)
6. The memory as it was told to me was sketchy. (R)

Subscale: Emotional Intensity

Instructions: You have just listened to a narrator describe a personal memory. Please indicate how you experienced the emotional intensity of the memory as it was told. Use the scale below to indicate your agreement with each statement.

1. As I listened to the memory, the emotions conveyed felt very intense.
2. The memory as it was told to me evoked very intense emotions in me.
3. The narrated memory evoked powerful emotions.
4. The memory as it was told to me did not evoke particularly strong emotions in me. (R)
5. I did not feel strong emotions while listening to this memory. (R)
6. The narrated memory did not evoke strong emotions in me. (R)

Appendix B

B1. Prompts for Specificity Classification Coding

You are a strict annotation system for psychological narrative coding.
You are coding autobiographical memories.
Your task is to classify each memory into exactly 1 category of the following 3 categories.
Category definitions:
1. SPECIFIC
- A single event.
- Occurred at a particular time and place.
- Duration less than 24 hours.
- A memory can be SPECIFIC even if an exact date is not stated, provided it clearly refers to a unique occurrence.
2. EXTENDED
- A single continuous event.
- Duration greater than 24 hours.
3. CATEGORIC/GENERAL
- Repeated events, habits, routines, or summaries of multiple similar experiences.

Decision Rules:
If the memory mainly describes one particular occurrence, classify as SPECIFIC or EXTENDED.
Background information should not determine classification.
If repeated events are the primary focus, classify as CATEGORIC/GENERAL.
If a memory contains both specific and general information, classify according to the main event being described.
Choose exactly one category.
Do not invent or add information that is not explicitly present in the memory.
Base the classification only on the text provided.

Output Requirements:
Each memory is categorized into one category. The category is referenced to the following way: 1 = specific. 2 = Extended. 3 = Categoric/general.

B2. Prompts for Episodic Richness Coding

You are a strict annotation system for psychological narrative coding.
Your task is to identify and count occurrences of five categories in autobiographical narratives at the sentence level.

Definitions of the 5 categories:

1. SPA: Spatial References
   References to locations, places, physical settings, spatial positions or spatial relationships.
   Examples:
   "at school" - location
   "next to you" - spatial relationship
   "under the table" - spatial positions
   "Exam room" - place
2. EP: Entities Present
   References to distinct people, objects and animals.
   Examples:
   "My sister" - person
   "flowers" - object
   "A dog" - animal
3. SD: Sensory Descriptions
   References to sensory experiences, environmental conditions, or physically observable descriptive properties of entities
   Examples:
   "It was cold" - sensory experience
   "It felt soft" - sensory experience
   "The room was warm" - environmental conditions
   "The building was tall" - descriptive properties of entities
4. TE: Thoughts and emotions
   References to thoughts, feelings, intentions or subjective evaluations made by the narrator or other people
   Examples:
   "I felt nervous" - feelings
   "I thought it was sweet" - thoughts
   "I was going to help" - intentions
5. A: Actions
   References to observable actions of the narrator or other people.
   Examples:
   "We talked"
   "He ran"

Unit of Analysis:

Process one sentence at a time.
For SPA, SD, TE, and A, always evaluate each sentence using the target sentence only.
Context sentences are not used for classification. Do not evaluate full narratives globally.

For EP only, use the full narrative to determine whether a person, object, or animal has already been counted earlier. EP counts should represent newly introduced unique entities, not repeated mentions.
A sentence can contain multiple categories and multiple counts within the same category.

Counting Rules:
For SPA, SD, TE, and A, count each distinct explicit instance within a sentence separately. Multiple locations, sensory descriptions, thoughts/emotions, intentions, evaluations, or actions within the same sentence each receive their own count.

Special EP counting rule:
Unlike SPA, SD, TE, and A, which are counted sentence by sentence, EP is counted globally across the entire memory.
Maintain a global entity register for EP across the full narrative. Count each unique concrete person, object, or animal only once in the full memory, at the first explicit mention. Later references to the same entity, including repeated names, noun phrases, pronouns, possessives, or role descriptions, do not receive additional EP counts.
For example, if one sentence says "my boyfriend and I" and a later sentence again says "my boyfriend and I," these refer to the same two people and should receive a total EP count of 2 across the full memory, not 4.
For sentence-level output, EP should reflect only the number of new unique entities first introduced in that sentence. Repeated entities already counted earlier in the memory receive EP = 0 in later sentences.
This global tracking rule applies only to EP. All other categories must still be evaluated using only the target sentence and counted by explicit instances within that sentence.

Decision Making Rules:
Locations and places count only as SPA and not as EP. (e.g. "exam room" = SPA, not EP).
For EP, count only concrete people, concrete objects, and animals. Do not count abstract concepts, events, activities, institutions used as places, time periods, emotions, achievements, qualities, or general outcomes as EP. Repeated references to the same person/object/animal within one sentence, including pronouns and possessives, count as one EP.
Emotions are only TE and not SD (e.g. "I felt joyous" = TE, not SD).

Internal Rules:
Do not infer meaning beyond the explicit text.
Only count explicit clearly stated instances. Do not infer implied categories.

Two-Step Process:
Step 1: Detection
Determine the presence of the 5 categories in the sentence.
Step 2: Confirmation
Confirm that the detected cases are present, by referring to the categories' definitions, the counting rules and the decision-making rules.

Output:
Output exactly in the following format (no additional text):
The totals include each count for each category for the whole memory.
TOTALS:
SPA: X, EP: X, SD: X, TE: X, A: X

B3. Prompts for Self-event connections

You are a strict annotation system for psychological narrative coding.
Your task is to identify and count self-event connections (SECs) in autobiographical text at the sentence level.

Definition of SECs:

A self-event connection is an explicit statement that connects the described event within the memory to an aspect of the narrator's self. Every self-event connection must belong to one of the following four categories:

1. Dispositions: traits, stable characteristics and personality
   Examples: "This experience reminded me that I am shy"
2. Values: moral reasoning, right and wrong, ethics and principles
   Examples: "I believe it is right to…", "I think lying is wrong"
3. Outlook: worldview and attitudes about the world and people
   Examples: "I think the future is bright", "This experience changed how I view refugees"
4. Personal growth: identity development. Change and maturation
   Examples: "over time I changed my mind" "I learned what matters to me" "Now I feel more confident".

A detected self-event connection must clearly fit at least one of the four categories above. Use the categories only to determine whether a self-event connection is present; do not output the category.
A self-event connection will almost always contain a first-person reference (I, me, my, myself, we, us, our). However, the presence of a first-person reference alone is not sufficient.
Only code self-event connections when the narrator is explicitly referenced in the sentence. Statements that do not explicitly refer to the narrator should be coded as NO.
Do not count statements about another person's dispositions, values, outlook, or personal growth unless the narrator is explicitly included (e.g., we, us, our).

Unit of Analysis:

Process one sentence at the time.
Always evaluate each sentence using the target sentence only. Context sentences are not used for classification.
Do not evaluate full narratives globally.

Two-Step Process:

Step 1: Detection
Determine whether the sentence explicitly links the remembered event to the narrator.
Step 2: Confirmation
Confirm that the link belongs to one of the four SEC categories, by referring back to the definitions and the rules provided. If it does not, mark NO.

Internal rules:

Do not infer meaning beyond the explicit text.
A sentence can contain a maximum of one SEC.

If the connection between the event and the narrator must be inferred rather than explicitly stated, mark NO.
Not SECs:
Do not count self-event connections that are about others than the narrator
Do not count emotions (e.g. “I felt happy”) as SECs

Output:
SEC_Count: <number>