COGNITIVE SCIENCE

How Much of a Day Do We Remember? Sensory Input, Attention, and the Development of Memory

A focused evidence review and participant-clustered reanalysis of mental replay in everyday life

Revanth Matha · Matha Labs | Preprint awaiting review 1.0 | 2026-11-20

Abstract

A person can remember where a conversation happened, recognize the other speaker, and retain an idea from it while losing the exact words. Which of these outcomes should count as remembering the conversation? We examine how researchers measure sensory input, selective attention, developmental learning, object recognition and episodic recall. We then reanalyze a public wearable-camera study in which 40 young adults recorded everyday events and later mentally replayed them. Applying the original duration exclusion left 347 events. Median event duration was 13.13 minutes and median replay was 31.73 seconds (participant-clustered 95% interval, 27.16-38.04 seconds). The median within-event ratio of elapsed to replay time was 24.25. These observations quantify the tempo of retrieval under a cue. They cannot establish a daily fraction of sensory input stored or a biological storage rate in gigabytes. We identify measurable alternatives, including the fraction of prespecified details recovered after a delay and the ability to recognize a familiar object while distinguishing two encounters with it. We conclude with experimental predictions for a real-time weight-memory system and explain which comparisons to human development can be tested.

Keywords: autobiographical memory, attention, cued recall, interoception, infant development, temporal integration, secondary data analysis

Contributions

1. A comparison of the units and denominators used to measure sensation, attention, recognition and recall.

2. An independently reproducible descriptive analysis of 347 everyday events with participant-clustered uncertainty and two sensitivity checks.

3. An account of how early sensorimotor responses, prenatal exposure, social feedback and cultural learning constrain analogies to initialized model weights.

4. Testable predictions for preserving object identity, distinct episodes and response speed under ongoing memory consolidation.

1. What would count as remembering a day?

Suppose you spend lunch with someone who is wearing a blue jacket. Tomorrow you may identify their face in a crowd and explain an idea they shared. You might misremember the jacket or be unable to reproduce the order of the sentences. Years later, you might still use the idea without recalling that lunch. Counting minutes of retained video would assign one unit to four different outcomes. A defensible question names the content to be recovered, the delay, the cues, and a rule for deciding whether a response is correct.

The same principle applies before recall. A pressure change impinges on the ear, a particular voice dominates attention, a listener understands a promise, and tomorrow that promise changes a decision. Acoustic power, neural activity, correct words and a changed decision each require a different instrument. A decrease in information reported between two tasks does not reveal how many physical inputs were discarded. A signal can influence action without a person later describing it.

This paper asks three questions. What do experiments establish about selection and later memory? Which early abilities give researchers a justified starting point for a developmental analogy? What numerical account of everyday recall can a researcher reproduce? We use the last question to test a specific interval-based measure and then set out an experiment in which the idea of a daily fraction would have an identifiable denominator.

2. Evidence and measurement methods

We conducted a focused narrative review on 20 November 2026 using web searches of PubMed, PubMed Central, journal sites and author repositories. Search phrases included "partial report visual presentations," "attended speaker multi-talker cortical response," "heartbeat predicts visual detection," "newborn differential rooting," "preterm suck swallow respiration," "development babbling contingent responses," and "wearable camera real-life event duration mental replay." We followed references to original reports and included a study when its stimulus, task, outcome and species could be identified. We selected contrasting tasks to show how measurement choices change the answer. This approach neither estimates study prevalence nor pools effect sizes across laboratories.

We classified results by their unit of observation. A sensory encoding experiment measures an input or a response to it. An attention experiment changes a goal or cue while comparing responses. A recognition experiment asks whether an item was previously present. A recall experiment asks a person to produce material under specified cues. The strength of a claim depends on its denominator. In a two-choice recognition test, the number tested and the 50% chance level are known. In an ordinary afternoon there is no natural count of all potentially reportable details.

Our quantitative component is a secondary analysis of Colson and D'Argembeau's published numeric workbook and accompanying analysis script. We verified the workbook's SHA-256 digest before parsing it and used three columns: participant identifier, experienced event duration in minutes and mental replay duration in seconds. Workbook identifiers serve only to group observations. All new figures and reported results are aggregate. We did not access camera photographs, participant names or recordings. [1][2]

Table 1. A quantity needs a measurement rule

Question	Observable	What determines its denominator?
How much was captured?	File size or digital sample count	A declared instrument, recording interval and encoding
What was attended?	Task-specific detection, selection or accuracy	Presented alternatives, cues and scoring rule
What was recognized?	Correct choices, hits and false alarms	Target and foil sets, response format, chance level
What was recalled?	Correct prespecified details and intrusions	A ground-truth event sample, delay, and cue
How long did recall take?	Seconds spent replaying a selected event	Number of eligible timed replays

Two studies can both use the word memory while estimating different quantities. Failure to report a detail is compatible with several mechanisms, including a missing cue.

3. From a physical signal to a selected event

Light intensity, sound pressure, skin deformation and chemical concentration can vary continuously in the environment. The human observer moves, blinks, sniffs and samples only portions of that environment. Receptor responses, retinal spike patterns and a microphone's digital stream are already different representations. The interior of the body provides another source of changing signals. A list of external senses alone would omit information about breathing, cardiac activity, posture and bodily need. These inputs have different units and often share information about the same event.

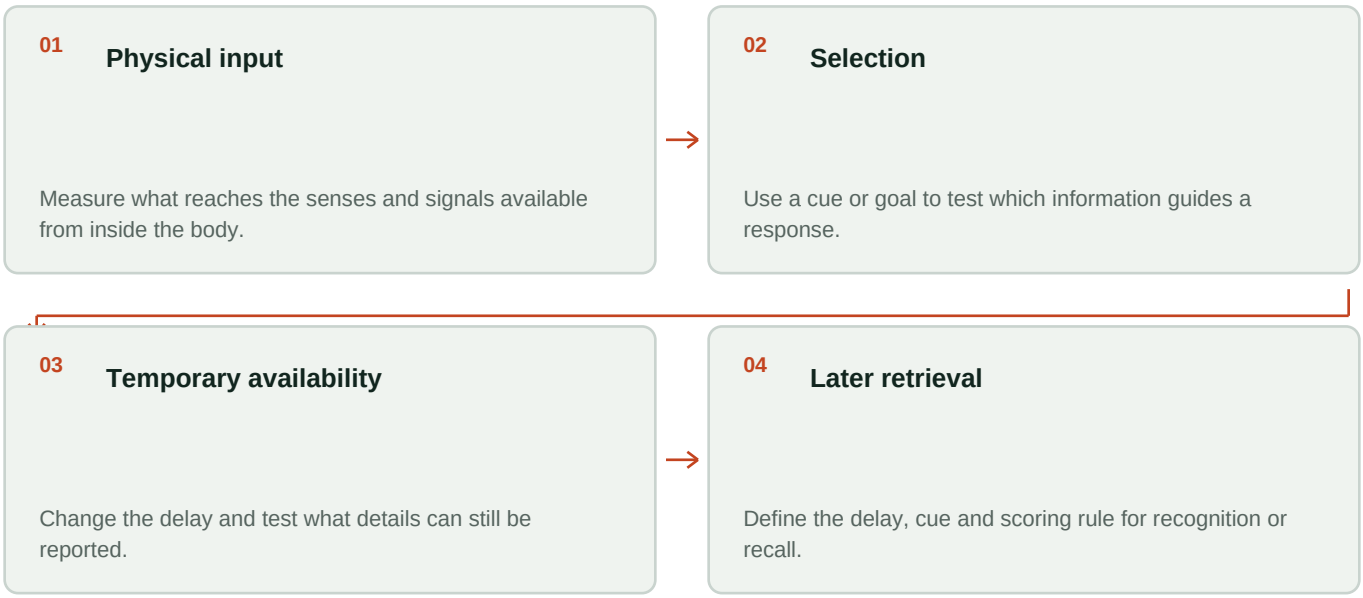
Unit conversion has a precise failure point. Koch and colleagues recorded seven types of guinea-pig retinal ganglion cells responding to moving images. Under an independence assumption, they estimated roughly 875,000 bits per second for an entire guinea-pig optic nerve. The estimate refers to distinguishable responses under their stimulus distribution. Its human counterpart was an extrapolation. Even a direct human optic-nerve estimate would still concern visual transmission. It supplies neither an auditory or bodily input rate nor a later recall count. [3]

Attention can alter how a fixed input is represented. Mesgarani and Chang presented a mixture of two voices to three people with clinically implanted electrodes. When the listener's task changed, responses in non-primary auditory cortex carried features of the attended speaker. The microphone would record the same mixture

across attention instructions. Park and colleagues found that the neural response accompanying a heartbeat before a faint visual grating appeared predicted its detection. This association provides an example of internal state interacting with external perception. It does not identify all pathways through which that interaction occurred. [4][5]

Motivation can act before the event is encoded. Adcock and colleagues used a cue that promised high or low monetary reward for remembering the upcoming scene. High-value scenes were more often recognized about 24 hours later. A laboratory incentive is one measurable instance of goal-dependent selection. Long-term concerns such as a relationship or a career involve changing goals over many encounters, so an incentive effect in one task cannot assign them a separate biological processing clock. [6]

Measurements from stimulus to later recall



The boxes identify measurements. Selection and temporary availability can overlap, and later retrieval depends on cues as well as earlier experience. [3][4][8]

4. A buffer has a time scale and a test

Sperling showed arrays of symbols briefly and asked people to report the whole array or a row designated by a tone. In a whole-report task after a 50-millisecond exposure, participants reported about 4.5 items on average. An immediate row cue revealed more available information than participants could write down in a whole report. Within roughly one second the extra advantage had largely disappeared. This is evidence of rapidly decaying visual availability. Its unit is recoverable symbols under a cue. Its time course depends on the display, task and reporting procedure. [7]

Working memory asks a different question. Luck and Vogel compared a sample display with a later test display and asked whether a feature changed. Participants performed close to perfectly on one to three simple items,

and performance declined at larger set sizes. Brady and colleagues, in contrast, presented 2,500 object photographs over several hours. Participants later chose an old object against a foil. Accuracy was approximately 92% against an unrelated object and 88% and 87% against another exemplar or another state, respectively. A small active comparison set can therefore coexist with a large collection of discriminable past items. The tested target-foil pairs supply the denominator. The experiment does not show reconstruction of 2,500 pixel arrays. [8][9]

Neither a screen full of symbols nor a stream of pictures can stand in for every sense. Cornell Kärnekull and colleagues tested 83 young participants on familiar and unfamiliar odors and faces immediately and after intervals extending to 64 days. Recognition changed with delay and familiarity for both stimulus classes. To measure daily multisensory retention, one would need defined odor, tactile and auditory probes as well as visual probes. A method that records photographs alone has no ground truth for the flavor of lunch or its emotional meaning. [10]

5. Which capabilities come first?

An infant turning toward a touched cheek and an infant successfully finishing a feed demonstrate different abilities. In a small study, Rochat and Hespos compared contact from an experimenter with contact from the infant's own hand. The five newborns rooted more to external stimulation. The eleven four-week-olds showed the opposite pattern. Here rooting meant a visible orienting response of the head and mouth. Amaizu and colleagues measured bottle-feeding performance in 16 infants born at 26-29 gestational weeks. They recorded sucking, swallowing and breathing as oral feeding progressed. Their component measures developed at different rates. A touch-triggered orientation, rhythmic sucking and coordinated transfer of milk are separate outcomes. The preterm cohort places a limit on generalization to healthy term newborns. [11][12]

Genetic instructions and early learning are plausible at different levels of explanation. A survey of gene expression in 1,340 samples from 57 human brains found patterns that varied with region and developmental time. That survey did not test rooting or recall. In a separate experiment, Partanen and colleagues exposed fetuses to variants of a repeated speech sound starting in late pregnancy. After birth, the previously exposed group showed a stronger neural response to a pitch change. A behavior or response present at birth can therefore reflect a mixture of inherited developmental organization, bodily maturation, and experience before observation began. Calling the organization a base memory is a proposed engineering vocabulary. Neither study shows that DNA stores a recollectable episode of feeding. [13][14]

Infants acquire capabilities along overlapping trajectories. Goldstein and colleagues studied 30 eight-month-olds. Caregiver responses timed to a child's vocalizations elicited more mature features of babbling than yoked responses delivered with an unrelated timing pattern. In an independent experiment, 33 six-to-nine-month-olds looked toward familiar items in pictures after a parent named them. Receptive word knowledge can appear while vocal production is still developing. These experiments give separate measures for producing a sound, responding to feedback and recognizing what a word refers to. [15][16]

Body and environment change the training distribution. In Adolph and colleagues' free-play samples, 12-19-month-old walkers averaged 2,368 steps and 17 falls per observed hour. Most observations came from a laboratory playroom, and sessions lasted 15-60 minutes. The measurements expose substantial practice as

infants learn balance and locomotion. Later, reading adds a culturally taught mapping. Dehaene and colleagues compared 63 adults with three literacy histories, including 22 who learned to read in adulthood. Brain responses to printed and spoken forms differed with literacy. Because education and life history were not randomized, group differences cannot isolate every cause. Together these cases favor a developmental model with changing sensors, actions, feedback and instruction. They leave open how to distinguish stored knowledge from a newly available motor skill. [17][18]

6. Recognition time, integration time and planning time

Recognizing an animal quickly does not answer the question of whether it is the same individual seen yesterday. Thorpe and colleagues showed previously unseen photos for 20 milliseconds to 15 adults who had to decide whether each contained an animal. Electrical responses to target and non-target trials diverged around 150 milliseconds. That is a latency for a task-related neural signal. It differs from a completed button press, a stable object identity decision, or retrieval of where a specific animal was encountered. An AI system claiming near-real-time recognition should specify the event that starts its clock and the completed output that stops it. [19]

Hasson and colleagues scrambled silent films at scales of approximately 4, 12 and 36 seconds. In different parts of the cortex, fMRI response reliability depended on preserving different durations of temporal structure. Ho and colleagues found that participants in planning tasks selectively represented obstacles that mattered for reaching a goal. These studies motivate multiple ways to integrate context and direct computation. They supply no evidence for a universal brain update every half, one or two seconds, or independent networks dedicated to business and relationships. A cortical integration window, a planning horizon and a software worker's scheduling period are three different parameters. [20][21]

Consider a street crossing on the way to meet a friend. An approaching bicycle demands a rapid response. The customary route, the remembered meeting place and the longer social purpose may all influence the next action. A hypothesis worth testing is whether a responsive route can continue producing correct actions while a slower route consolidates past encounters. The biological observations identify plausible constraints and competing mechanisms. They supply no engineering latency target by themselves.

7. How we analyzed everyday replay

Colson and D'Argembeau asked 40 young adults to record three everyday events per day for three consecutive days with wearable cameras. Their protocol included a later cued memory test, described by the team as taking place two to four days later. Participants mentally replayed selected experiences. The workbook contains 358 event rows. We used the authors' script rule retaining replay durations strictly below 187 seconds. The resulting sample contains 347 events across all 40 people. The missing two rows relative to nine planned events per person have no records in the workbook. We make no assumption about what occurred in those slots. The sample describes events deliberately recorded during a study. [1][2][22]

For each included event we divided elapsed event time, converted to seconds, by the reported replay time. We separately calculated the median event length, median replay time and median of the event-specific ratios.

Dividing the first two sample medians gives a different quantity because the events in their middle ranks may differ. To display the distribution, we grouped event duration into four ranges with boundaries at 5, 10 and 20 minutes. The bins are descriptive choices made for this reanalysis. They serve no role in the authors' model.

We estimated pointwise 95% percentile intervals with 5,000 bootstrap replicates. Each replicate draws 40 participant identifiers with replacement and includes all eligible events from each drawn person. This preserves correlations among one person's events. It retains the original estimator's event weighting. Our sensitivity checks reinstate the 11 excluded events and weight people equally by taking the median of each person's median replay. Fixed random seed 20261120 and the pinned workbook make the results reproducible. The original authors fit nonlinear mixed models. We offer a new descriptive analysis of the same observations. Selection of people, choice of events, possible cueing effects and unrecorded intervals remain sources of uncertainty beyond the bootstrap.

```
R_i = (60 x T_i) / Q_i
```

T_i is experienced event duration in minutes, Q_i is timed mental replay in seconds, and R_i is an event-specific elapsed-to-replay time ratio. Its unit cancels. It contains no measure of representational fidelity or storage size.

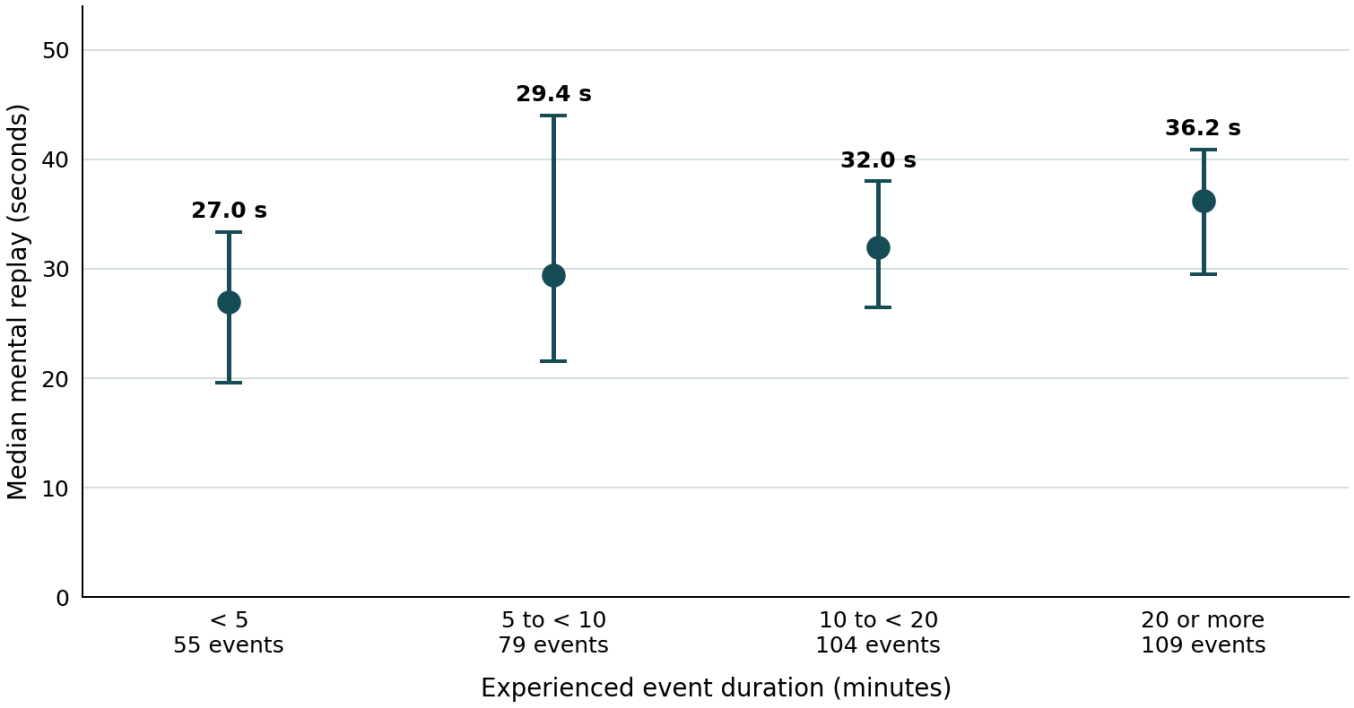
8. Results from 347 cued events

Median event duration was 13.13 minutes (95% interval 11.57-14.99) and median replay duration was 31.73 seconds (27.16-38.04). The median paired duration ratio was 24.25 (19.77-29.94). A simple interpretation is that the middle observed ratio corresponded to about 24 seconds of elapsed experience per second a participant spent mentally replaying that event. That arithmetic says nothing about which parts of an episode the participant could recognize later.

Median replay times in the four duration bands were 27.0, 29.4, 32.0 and 36.2 seconds. The widest interval was in the five-to-under-ten-minute band. The four intervals describe uncertainty for each median separately. The points alone cannot establish an individual-level growth curve or a discontinuity at a particular duration. Three included events lasted at least an hour, so extrapolating to a workday would give the longest events disproportionate influence.

The median including all 358 rows was 32.71 seconds. The median of the 40 participant-specific medians after exclusions was 32.54 seconds. These estimates are close in this dataset. Both still depend on which episodes participants recorded and which cue was available. Neither sensitivity check recovers unobserved events, measures completeness of the remembered scene or calibrates the meaning of a second of subjective replay.

Figure 1. Mental replay duration for shorter and longer sampled events



Dots show event-level medians. Vertical bars are pointwise 95% percentile intervals from 5,000 whole-participant bootstrap samples. Each range is lower-inclusive and upper-exclusive. Source observations and study protocol are those of Colson and D'Argembeau. [1][2]

Table 2. Duration bands and clustered uncertainty

Event length (min)	Events	People represented	Median replay, s (95% interval)
Under 5	55	26	27.0 (19.6-33.4)
5 to under 10	79	36	29.4 (21.6-44.0)
10 to under 20	104	38	32.0 (26.5-38.0)
20 or more	109	35	36.2 (29.5-40.9)

People can occur in several bands. Counts of events sum to 347. There is no imputed value for a day or a participant who contributed no event to a particular band.

9. The fraction we could estimate, and the one we cannot

A wearable-camera study by Finley and Brewer illustrates an alternative denominator. Nine undergraduates wore cameras on two days. At one week and one month they recalled information about selected intervals, then examined photographs and added or corrected information. Participants self-scored their responses. The authors reported 89% mean accuracy and treated their 79% mean completeness as an upper bound under this procedure. The latter depends on the visible, eligible material initially recalled or later identified through picture review. A camera misses taste, much of what was heard, goals, and thoughts. A person may also never describe a detail that both the eyes and the camera encountered. [23]

Misra and colleagues used a more demanding test of episode-specific detail. Participants walked a specified route for about an hour. Later, one-second video clips from the person's own experience had to be distinguished from similar clips captured by another walker. Performance was only modestly above chance. This result sits naturally beside impressive recognition of deliberately studied object photographs. Familiarity with an object, recognition of its previous state, and identifying an exact occasion from a similar route test different information. [24][9]

A researcher who knows a camera's fixed bitrate can answer a different daily-quantity question exactly. If a recording averaged 5 megabits per second for 16 hours, its payload would be 36 decimal gigabytes, excluding container overhead. At 20 megabits per second it would be 144 gigabytes. Changing a codec setting changes those numbers while the person's experience stays fixed. Uncompressed frames likewise require an analyst to select a resolution, color depth and sample rate. Neither bandwidth nor a count of available synapses tells us how many particular experiences this person could later retrieve. There is no basis in these studies for multiplying the replay ratio by a camera rate to estimate biological storage.

A feasible future estimate begins with a declared waking observation window and random time sampling. Before delayed testing, blind coders would annotate a fixed set of verifiable details per sampled interval. For identical target sets one could compare unprompted recall, category cues and recognition with matched foils at a day and a week. Record both correct details and confidently reported errors. Estimate day-level retention of those sampled details with inverse sampling weights and person-level uncertainty. Separate permission would be needed to collect recordings from new participants and any third parties. The result would answer a specified retention question and expose areas the recording device could never score.

10. What the biology motivates for programmable memory

A model should be tested on both identity and episode. Present the same mug at different times and locations, then a near-identical mug with a different marking. Count duplicate identities for the same mug, false merges across mugs, and accuracy when asked where each encounter occurred. A useful consolidated representation might recognize one object's repeated appearances while still supporting two different answers about its history. An aggressive merge might improve identification speed and erase precisely the temporal distinctions the episode test requires.

To test parallel processing, interleave a rapid recognition task with a delayed planning task and run a background worker that consolidates accumulated observations. Compare 0.5-, 1- and 2-second intervals with

work triggered by a measured event, under equal inputs and work budgets. Measure tail response latency, missed deadlines, unresolved backlog, false merges, and retrieval of older episodes. Score goals explicitly on the planning task. Separate success at one-step recognition from success over a long-horizon decision. The clock periods are experimenter choices. The measured outcomes are computer performance.

Development adds a second test. Initialize some sensor-to-action responses and give the model a stream with recurrent objects, corrective feedback and gradually expanded action possibilities. Contrast the same eventual action set made available all at once. Compare generalization to new objects, retention of old associations and calibration after changes to the sensor or body. A separate condition should offer instruction in a culturally learned task only after new symbols appear. If the system performs equally well under shuffled exposure and state-free writes, a developmental or memory-dependent explanation loses support.

11. Conclusions and review questions

For a specified, cued sample of young adults' everyday events, mental replay took about half a minute at the median. For brief symbol displays one can measure available items. For object photographs one can measure forced-choice accuracy. For daily-life intervals one can score a prespecified set of details. Putting these findings side by side identifies the missing ingredients in a claim about an entire day: a common sampling frame, a recoverability criterion, a delay and a way to validate each sensory modality. The current evidence supports useful measurements without an estimate in gigabytes.

For methodological review, the most consequential choices are the event-weighted estimand, the original replay cutoff, and the decision to resample whole people. A reviewer can reproduce the chart and evaluate those choices from the published workbook. One could additionally fit the original authors' nonlinear model or assess alternative sampling schemes. A separate physiological experiment would be required to connect a neural transmission rate to a particular later memory outcome.

The quantitative work uses aggregate outputs of a secondary analysis of the authors' released numeric data. Individual recordings and the source workbook are not distributed by this project. The review and implementation are prepared for independent review. AI tools assisted with literature discovery, manuscript drafting and analysis code. Independent scrutiny of the original sources, code and interpretation is invited.

References

[1] Colson C, D'Argembeau A (2025). Memory in motion: how real-life event features influence the tempo of episodic recall. Memory 33, 1180-1195. Original wearable-camera study. Accessed 2026-11-20.

[2] Colson C, D'Argembeau A. Data, analysis script and materials for Memory in motion. Open Science Framework. Numeric workbook and authors' R analysis. Reuse here is limited to new summaries and figures. Accessed 2026-11-20.

[3] Koch K et al. (2006). How much the eye tells the brain. Current Biology. Species and model assumptions matter for information-rate comparisons. Accessed 2026-11-20.

[4] Mesgarani N, Chang EF (2012). Selective cortical representation of attended speaker in multi-talker speech perception. Nature. Human intracranial recordings during selective listening. Accessed 2026-11-20.

[5] Park H-D et al. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nature Neuroscience. Human magnetoencephalography and faint visual targets. Accessed 2026-11-20.

[6] Adcock RA et al. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron 50, 507-517. Incentivized scene memory. Accessed 2026-11-20.

[7] Sperling G (1960). The information available in brief visual presentations. Psychological Monographs 74(11), 1-29. Partial report with cue delays. Accessed 2026-11-20.

[8] Luck SJ, Vogel EK (1997). The capacity of visual working memory for features and conjunctions. Nature 390, 279-281. Visual change detection. Accessed 2026-11-20.

[9] Brady TF et al. (2008). Visual long-term memory has a massive storage capacity for object details. PNAS 105, 14325-14329. Forced-choice object recognition. Accessed 2026-11-20.

[10] Cornell Kärnekull S et al. (2015). Long-term memory for odors: influences of familiarity and identification across 64 days. Chemical Senses. Olfactory and facial recognition with varied retention intervals. Accessed 2026-11-20.

[11] Rochat P, Hespos SJ (1997). Differential rooting response by neonates: evidence for an early sense of self. Early Development and Parenting 6, 105-112. Self-touch and external touch. Accessed 2026-11-20.

[12] Amaizu N et al. (2008). Maturation of oral feeding skills in preterm infants. Acta Paediatrica. Longitudinal assessment of components of feeding. Accessed 2026-11-20.

[13] Kang HJ et al. (2011). Spatio-temporal transcriptome of the human brain. Nature. Gene expression across brain regions and developmental stages. Accessed 2026-11-20.

[14] Partanen E et al. (2013). Learning-induced neural plasticity of speech processing before birth. PNAS. Prenatal auditory exposure and neonatal mismatch responses. Accessed 2026-11-20.

[15] Goldstein MH, King AP, West MJ (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. PNAS. Contingent social responses and eight-month-olds' vocalizations. Accessed 2026-11-20.

[16] Bergelson E, Swingley D (2012). At 6-9 months, human infants know the meanings of many common nouns. PNAS. Eye movements in response to familiar words. Accessed 2026-11-20.

[17] Adolph KE et al. (2012). How do you learn to walk? Thousands of steps and dozens of falls per day. Psychological Science. Spontaneous locomotion measured during free play. Accessed 2026-11-20.

[18] Dehaene S et al. (2010). How learning to read changes the cortical networks for vision and language. Science. Adult literacy histories and cortical responses. Accessed 2026-11-20.

[19] Thorpe S, Fize D, Marlot C (1996). Speed of processing in the human visual system. Nature 381, 520-522. Timing of visual categorization. Accessed 2026-11-20.

[20] Hasson U et al. (2008). A hierarchy of temporal receptive windows in human cortex. Journal of Neuroscience. Temporally scrambled movies and fMRI. Accessed 2026-11-20.

[21] Ho MK et al. (2022). People construct simplified mental representations to plan. Nature. Goal-dependent choice of task representations in mazes. Accessed 2026-11-20.

[22] Colson C, D'Argembeau A (2025). Temporal compression in episodic memory for real-life events: the role of event features. University of Liège. Authors' conference protocol description for the same study. Accessed 2026-11-20.

[23] Finley JR, Brewer WF (2024). Accuracy and completeness of autobiographical memory: evidence from a wearable camera study. Memory. Timeslice recall followed by photo-cued assessment and comparison. Accessed 2026-11-20.

[24] Misra P et al. (2018). Minimal memory for details in real life events. Scientific Reports. Source-specific recognition of brief clips from a walked route. Accessed 2026-11-20.

Reproducibility record

Private review materials: https://github.com/RevanthM/human-sensory-memory | Research snapshot: 16a024eeb85c7a8452b54609d6575d6f922a41eb