Sleep Differentially and Profoundly Impairs Recall Memory in a Patient with Fornix Damage

Nelly Matorina1, Julie Tseng2, Natalia Ladyka-Wojcik1, Rosanna Olsen3, Donald J. Mabbott1,2, and Morgan D. Barense1,3

Abstract

In March 2020, C.T., a kind, bright, and friendly young woman underwent surgery for a midline tumor involving her septum pellucidum and extending down into her fornices bilaterally. Following tumor diagnosis and surgery, C.T. experienced significant memory deficits: C.T.’s family reported that she could remember things throughout the day, but when she woke up in the morning or following a nap, she would expect to be in the hospital, forgetting all the information that she had learned before sleep. The current study aimed to empirically validate C.T.’s pattern of memory loss and explore its neurological underpinnings. On two successive days, C.T. and age-matched controls watched an episode of a TV show and took a nap or stayed awake before completing a memory test. Although C.T. performed numerically worse than controls in both conditions, sleep profoundly exacerbated her memory impairment, such that she could not recall any details following a nap. This effect was replicated in a second testing session. High-resolution MRI scans showed evidence of the transcallosal surgical approach’s impact on the mid-anterior corpus callosum, indicated that C.T. had perturbed white matter particularly in the right fornix column, and demonstrated that C.T.’s hippocampal volumes did not differ from controls. These findings suggest that the fornix is important for processing episodic memories during sleep. As a key output pathway of the hippocampus, the fornix may ensure that specific memories are replayed during sleep, maintain the balance of sleep stages, or allow for the retrieval of memories following sleep.

INTRODUCTION

In March 2021, we first became aware of case C.T., a kind, bright, and friendly young woman who had an unusual form of amnesia that seemed to be related to sleep. In March 2020, at the age of 16 years, she underwent surgery to remove a brain tumor that was directly above both fornices and appeared to invade the right fornix. C.T. emerged from the surgery with severe anterograde amnesia, with her last stable episodic memory reported as the moment that the anesthesia mask was fit to her face. Remarkably, C.T.’s family described that she could remember things throughout the day, but when she woke up in the morning or following a nap, she would expect to be in the hospital, forgetting all the information that she had learned before sleep. To cope with the amnesia, C.T. wrote letters to herself explaining the memory loss and the decisions she had made on previous days. Some of these items were big life decisions, including education and relationships. This case was puzzling to us for two main reasons: First, to our knowledge, there have been no previously reported cases of dense hippocampal system amnesia [Smith et al., 2010], as well as some anecdotal accounts [e.g., Bett, 2022]; these cases have no known neurological origin. Second, sleep has been shown to be beneficial to memory consolidation.

Sleep plays an important role in episodic memory consolidation. For example, a meta-analysis showed that in young adults, slow-wave sleep (SWS) and non-rapid eye movement sleep significantly predicted episodic memory performance (Hokett, Arunmozhi, Campbell, Verhaeghen, & Duarte, 2021), suggesting that these specific sleep stages promote episodic memory consolidation. In animals, waking experiences are replayed during SWS in the hippocampus (Ji & Wilson, 2007). A similar process is thought to take place during non-rapid eye movement sleep in humans (Saletin & Walker, 2012), whereby recently encoded hippocampal memories are reactivated and consolidated. In terms of memory for everyday life events and naturalistic stimuli, participants had better memory for stories and personal events after a period of sleep compared with wake (Aly & Moscovitch, 2010) and improved free recall memory for naturalistic videos after a period of sleep (Coutanche, Koch, & Paulus, 2020). Thus, based on the previous literature, we would expect sleep to be beneficial to episodic memories, including memories of everyday life.

What happens during sleep when there is damage to the hippocampal system? People with hippocampal damage
have reduced SWS and slow-wave activity compared with healthy controls (Spanò et al., 2020), suggesting that the hippocampus influences in sleep physiology. Sleep disturbances are prevalent in both Alzheimer disease and mild cognitive impairment (MCI), both of which consistently result in hippocampal atrophy (Tapiola et al., 2008; Devanand et al., 2007; McCurry & Ancoli-Israel, 2005; Mega et al., 2002; McCurry et al., 1999). These sleep disturbances included sleep fragmentation, rapid eye movement (REM) sleep decrease, disruption of sleep/wake rhythms, and lower sleep efficiency (D’Rozario et al., 2020; Burke, Maramaldi, Cadet, & Kukull, 2016; Peter-Derex, Yammine, Bastuji, & Croisile, 2015; McCurry et al., 1999). Higher levels of SWS were associated with relatively better memory recollection in patients with Alzheimer disease (Rauchs et al., 2013). Relatedly, increased slow-wave activity in patients with amnestic MCI predicted an improvement in word recall the next morning. Together these data show that SWS and slow-wave activity play an integral role in the preservation of our memories. Importantly, CA1 and CA3 hippocampal atrophy was found to correlate with poorer sleep-dependent memory consolidation in those with MCI (Lam et al., 2021), showing that damage to the hippocampal system can alter the relationship between sleep and memory processing.

Sleep is typically associated with a memory benefit; however, sleep has been found to be detrimental to memory in one patient population—those with Down syndrome. Down syndrome is associated with a reduction in myelination in the hippocampal formation (Abraham et al., 2012), a reduction in hippocampal volumes (Koenig et al., 2021; Aylward et al., 1999), and impairments in types of memory dependent on the hippocampus (Lavenex et al., 2015; Visu-Petra, Benga, Tincas, & Miclea, 2007; Nadel, 2003). Surprisingly, whereas typically developing children show a memory benefit from sleep, children with Down syndrome show either a memory impairment following sleep (Spanò et al., 2018) or no benefit (Ashworth, Hill, Karmiloff-Smith, & Dimitriou, 2017). One explanation for a sleep impairment could be that in children with Down syndrome, the hippocampal system replays incorrect or nonspecific activity patterns during sleep, interfering with memory consolidation (Spanò et al., 2018). In addition, children with Down syndrome retained more memory at longer delays if they remained awake for several hours after encoding, suggesting that naps prevented effective memory consolidation (Spanò et al., 2018).

The fornix, which was implicated in C.T. S. original surgical report, is a key efferent pathway for the hippocampus (Benear, Ngo, & Olson, 2020), linking the hippocampus with the septal nuclei, the basal forebrain, the mammillary bodies of the diencephalon, and several anterior hypothalamic areas, including indirect connections to the anterior thalamic nuclei (Senova, Fomenko, Gondard, & Lozano, 2020; Aggleton et al., 2010). The fornix may also connect the hippocampus to the medial prefrontal cortex (Aggleton, Wright, Rosene, & Saunders, 2015); however, there is some dispute over whether these fibers constitute the fornix proper or the hippocampal-to-prefrontal cortex pathway (Godsil, Kiss, Spedding, & Jay, 2015). Damage to the fornix typically results in anterograde amnesia without retrograde amnesia. This pattern has been observed in cases of damage to the anterior columns of the fornix (Rizek, Pasternak, Leung, & Jenkins, 2013; Murr, Thaisetthawatkul, Helvey, & Fayad, 2012; Park, Hahn, Kim, Na, & Huh, 2000; Calabrese, Markowitz, Harders, Scholz, & Gehlen, 1995; Hodges & Carpenter, 1991), the body of the fornix (Chen, Zayas, & Gold, 2008), and the posterior portion of the fornix (D’Esposito, Verfaellie, Alexander, & Katz, 1995). However, some cases of damage to the anterior columns of the fornix have also resulted in both anterograde and retrograde amnesia (Baweja, Mensinkai, Reddy, & Sahlas, 2015; Adamovich, Guaitberto, Roberts, Haut, & Gutmann, 2009). These cases indicate that the fornix is necessary for the consolidation of new memories and sometimes also the retrieval of older memories. There are also a handful of studies that suggest an association between fornix damage and sleep disturbances, yet to our knowledge, these sleep disturbances have not been linked to memory impairment. Disrupted sleep patterns, including excessive daytime sleepiness and REM sleep behavior disorder, are associated with lower connectivity in the fornix (Dolatshahi et al., 2021; Ghazi Sherbaf et al., 2018; Matsui et al., 2006), demonstrating that alterations to the fornix may cause changes in sleep quality. Fornix damage was also associated with abnormal sleep behavior in patients with amyotrophic lateral sclerosis (Gabery et al., 2021).

Diffusion tensor imaging metrics, such as fractional anisotropy (FA), can be derived from diffusion-weighted imaging to characterize white matter in the fornix. FA describes the degree to which a principal diffusion direction exists within a voxel, and an existing body of work has related decreased FA in the fornix to cognitive deficits in aging (Chen, Strauss, Hayes, Davis, & Hodaie, 2015), Alzheimer disease (Mielke et al., 2012), and Parkinson disease (Matsui et al., 2006). Although FA is a reasonable metric when a voxel contains a singular white matter bundle, voxels containing crossing fibers would appear to have low FA. As a result, higher-order diffusion models that allow for more specificity in characterizing white matter changes with the traditional diffusion tensor imaging model are growing in popularity. Fixel-based analysis (FBA) is one such approach that uses constrained spherical deconvolution to model multiple fiber bundle populations (i.e., fixels) within a voxel (Dhollander, 2020). Although FBA is a relatively new approach, recent work has found an association between decreased fornix fiber density (FD) and age-related decline in memory performance (Radhakrishnan, Stark, & Stark, 2020), decreased fornix FD and Alzheimer disease (Mito et al., 2018), and increased fornix fiber cross-section (FC) in Parkinson disease (Rau et al., 2019).

In this study, we wished to empirically validate C.T.’s pattern of memory loss, as well as characterize structural
changes to her fornix and surrounding structures. We were most interested in understanding the role of sleep on C.T.’s memories of everyday life. For this reason, we tested C.T. and age-matched controls with a naturalistic stimulus (a TV episode) that approximated rich, temporally extended, real-life memories. To characterizestructural changes in the brain, we collected T1- and diffusion-weighted MRI scans from C.T. and a second age-matched control group. In addition to comparing hippocampal volumes between C.T. and the control group, we combined tractography with FBA metrics to localize fiber bundles differences along the fornix as well as corpus callosum, as well as to explore the fornix’s potential role in connecting the hippocampus to other brain areas.

METHODS

Case History

C.T.’s history is significant for surgical resection and diagnosis of an intraventricular pilocytic astrocytoma. She presented in March 2020 at the age of 16 years with a 3-year history of persistent headaches. Over time, the headaches escalated to daily frequency that were associated with dizziness and light-headedness—typically starting a few hours after waking in the morning. In March 2020, she started to experience nausea and episodes of vomiting and was admitted to the hospital. Upon hospital admission, MRI scanning revealed a midline lesion in the septum pellucidum, which involved the fornices. This lesion was located above the anterior commissure and was causing some compression of the foramen of Monro bilaterally and mild obstructive hydrocephalus. C.T. was taken for surgery where a trans-callosal subtotal excision was carried out using neuro-navigation. Specifically, surgery involved bifrontal exposure and right frontal craniotomy for an interhemispheric approach to the intraventricular tumor. During surgery, it became clear that the tumor was intimately associated with the right fornix. The surgical team reported that they achieved 70% debulking of the tumor and elected to not go any further because of concerns regarding impact on memory. Postoperative MRI showed residual tumor involving the inferior and posterior aspect of the septum pellucidum on the right side. Immediately postoperatively, C.T. displayed significant anterograde memory deficits.

Following surgical admission in the hospital, C.T. was then admitted as an inpatient to a brain injury rehabilitation program at a rehabilitation center for 6 weeks in April–May 2020. During this admission, she received occupational therapy, physiotherapy, and neuropsychological consultation services. An overnight sleep study from 2020 indicated that C.T. had normal sleep latency (20 min), normal sleep maintenance efficiency (90.5%), normal REM sleep latency (122 min), normal apnea hypopnea index (0.3/hr), and normal periodic limb movements. However, she had decreased amount of SWS (15%), increased stage N1 (7.6%), and increased arousal index because of spontaneous arousals (21.3 arousals/hr). This information was shared with us by C.T.’s family from a report by YouthDale Sleep Center. Anecdotally, C.T.’s family reported that she had insomnia before the surgery but not following the surgery.

C.T.’s family first contacted the present research team in March 2021 when she was 17 years old, about 1 year after her surgery. C.T. was finishing her final year of high school with accommodations. Her family indicated that she could successfully complete assignments if she worked on them all in one session with no periods of sleep in between. However, if she took a nap or slept at night, she would forget the assignment and any progress she had made up until that point. C.T. would wake up in the morning or following a nap expecting to be in the hospital directly after surgery. However, she was able to remember things throughout the day if she did not sleep (e.g., in the afternoon she could discuss what she had eaten for breakfast that morning). Following a suggestion from one of her doctors, her family made several attempts to reduce the length of her naps on the chance this might help her retain information. Unfortunately, however, if C.T. did not nap, she experienced severe headaches. Moreover, reducing the length of her naps did not appear to improve her memory. Over time, around June 2022, we learned that C.T.’s expectation to be in the hospital upon waking up gradually went away, although she did not appear to have any new long-term episodic memories.

C.T. was between 18 and 19 years old at the time of behavioral tests and MRI scan reported in this article. On our demographics questionnaire, C.T. indicated that her first language is English and that she speaks both English and French. She is right-handed, wears glasses, and is not color-blind. She has a family history of Alzheimer disease or other related dementias (onset at age 85 years). C.T. scored 3 on the Pittsburgh Sleep Quality Index (PSQI; scores below 5 are considered to be “good” sleepers; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). She scored 57 on the Morningness–Eveningness Questionnaire (MEQ; Horne & Östberg, 1976, intermediate morningness–eveningness type). Her family assisted with filling out the PSQI and MEQ questionnaires because the questions asked about frequency of events in the last month. C.T. gave written informed consent, which was approved by the University of Toronto Ethics Board.

Overview of Current Case Report

We conducted a number of different assessments with C.T. to characterize her case. These include: (1) a standardized neuropsychological test battery, chosen to capture attention, memory, verbal learning, and executive functioning; (2) an autobiographical interview (Levine, Svoboda, Hay, Winocur, & Moscovich, 2002) to assess memory for events that occurred before and after surgery; (3) a formal characterization of her sleep-related memory...
deficit, testing memory recall and comprehension for a TV episode across nap and wake delays in C.T. and a group of age-matched controls; and (4) T1- and diffusion-weighted MRI scans to assess hippocampal size, conduct fornix tractography, and FBA to localize fiber bundle differences along the fornix in C.T. relative to a second group of controls. The memory for a TV episode was conducted in 2021. The MRI scans were completed a year later in 2022. The neuropsychological tests and the autobiographical interview were conducted a few months after the MRI scans on two adjacent days. As of February 2023, C.T.’s family confirmed that her pattern of memory loss had not changed.

Standard Neuropsychological Tests

We tested C.T. in June 2022 on a battery of tests that included: Rey Auditory Verbal Learning Test (Schmidt, 1996); Trail Making Test Trail A and Trail B (Reitan & Wolfson, 1985); Digit Symbol Substitution (Wechsler, 2012); Brief Visuospatial Memory Test–Revised (BVMT-R; Benedict, 1997); 1-min phonemic verbal fluency (F-A-S; Spreen & Benton, 1977); and 1-min categorical verbal fluency (Animal Naming; Spreen & Benton, 1977).

Autobiographical Interview

Instructions and Free Recall

In the autobiographical interview (Figure 1), C.T. was asked to recall five events from the following age ranges: early childhood to age 6 years; age 7–10 years; age 11–13 years; age 13–16 years; and age 17–19 years (age range following tumor diagnosis and surgery). C.T. was 19 years old at time of testing. The interview was audiotaped to allow for transcription and scoring.

C.T. was given a list of typical life events and asked to either select one event from the list for each time period or to choose a different event. She was asked to select events that were specific to a particular time and place and to provide as much detail as possible. She was instructed to select events that she was personally involved in, had a recollection of being personally involved in, and that she was comfortable discussing in detail. C.T. was informed that it was not the events she chose that were of interest, but how she described them. After the instructions, C.T. was given an opportunity to ask any questions. We then asked C.T. to recall memories beginning with the earliest age category. C.T. recalled the entire event without any interruptions before moving on to the next event.

Scoring and Results

Internal details (those directly related to the specific event) and external details (not related to the specific event) were summed. Two independent raters (including the first author) scored event recall for internal and external details. Intraclass correlation coefficients (ICCs), a measure of reliability, were calculated only for the four life events that C.T. could recall (not including the memory from the same day). ICCs for a fixed number of raters were calculated using the psych package in R Studio Version 4.0.3 (R CoreTeam, 2020; RStudio Team, 2020). The ICCs were as follows: .89 for internal details and .87 for external details, indicating good reliability (Koo & Li, 2016). Following independent ratings, raters discussed any differences in their ratings to determine final detail scores on which both agreed.

Memory Test for TV Episode

Overview

We assessed C.T.’s memory for rich, temporally extended events across nap and wake delays relative to neurologically intact controls. Participants were tested on both their recall memory for a TV episode in response to a series of prompts, as well as a multiple-choice comprehension test.

Control Participants

We recruited 10 healthy age-matched controls (\(M_{\text{age}} = 18.20\) years, \(SD_{\text{age}} = .40\) years) from local high schools, the University of Toronto, and through a flyer advertising the study. The sample size for control participants was chosen to be similar to previous single-case studies (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004; Schacter, Curran, Galluccio, Milberg, & Bates, 1996).
Participants were included if they met the following criteria to ensure regular sleep patterns: no history of sleep disorders, sleeping a minimum of 7 hr per night, usually going to sleep no later than 2 a.m. and waking up no later than 10 a.m., an MEQ score between 31 and 69 (to exclude extreme chronotypes), a PSQI score ≤ 5 (to exclude poor sleep quality), and drinking three or fewer servings of caffeine per day. Participants were also only included if they were comfortable napping and had access to a computer with a camera and a microphone for Zoom meetings. PSQI and MEQ score cutoffs were taken from a previous sleep study (Schapiro, McDevitt, Rogers, Mednick, & Norman, 2018). Participants were also asked if they were familiar with the TV show Poirot and, if so, if they had heard of the specific episodes used in the experiment. Those familiar with the episodes were excluded from the study.

Participants indicated their gender as female (n = 8) or male (n = 2). Controls had a mean of 14.2 years of education (SDeducation = 1.62). Two participants indicated that they thought they had greater memory problems than their peers. Participants scored an average of 52.5 on the MEQ (SD = 3.95). They scored an average of 4.2 on the PSQI (SD = 0.63, range = 3–5), indicating that our control participants had good sleep quality (Buysse et al., 1989). All participants gave written informed consent, which was approved by the University of Toronto ethics board.

**Stimuli**

We tested participants on two episodes of the television show Poirot, a detective show that ran from 1989 to 2013. Episodes were selected from previous work in our laboratory because of their memorability and lack of familiarity to most participants. Episode A “The Adventure of Clapham Cook” (1989) was 50 min long, and Episode B “The Theft of the Royal Ruby” (1991) was 48 min long. Both episodes were produced by London Weekend Television. Each episode stands on its own and their order is interchangeable in terms of the overall storyline.

**Design**

Our dependent measures were number of details recalled in response to prompts from the episodes and scores on a multiple-choice comprehension test. A prompted recall measure was selected to scaffold memory recall as much as possible for C.T. We employed a within-subject design, in which all participants completed both nap and wake conditions (Figure 2A). Participants were randomly assigned to either complete the nap condition on Day 1 and wake condition on Day 2, or the reverse arrangement. C.T. underwent the same procedure, with the addition that she completed both conditions again after a 10-day delay with condition assignment flipped. Specifically, on Day 1, she watched Episode A in the wake condition and, on Day 2, she watched Episode B in the nap condition. Sixteen days later, on Day 17, C.T. watched Episode A in the nap condition and, on Day 18, she watched Episode B in the wake condition.

**Procedure**

Participants met with the experimenter over Zoom either at 10:20 a.m. (nap condition) or 12:45 p.m. (wake condition). Timings were dictated by C.T.’s nap schedule, which was consistently at 1:45 p.m. to 3:15 p.m. every day. Participants watched the TV episode on their computer while sharing their screen with the experimenter and leaving their camera on to ensure that they were paying attention. If a participant appeared to be distracted, the experimenter would politely ask them to pay attention to the episode. Participants were asked to report any technical difficulties they encountered.

Following the episode, participants took a 100-min break. Participants in the wake condition were asked to stay awake and refrain from napping during this interval. They were also asked not to rewatch or research anything related to the episode they just had seen. Participants were not monitored during this interval and were not asked to complete any specific tasks. We chose this design to match C.T.’s normal schedule and to make the results generalizable to her everyday life and memory loss. Participants in the nap condition were asked to nap during the first 90 min of the break and wake up 10 min before the memory test. The memory test comprised two components, a prompted recall section and a comprehension test. Following the comprehension test, control participants also rated the difficulty of the comprehension test questions. Lastly, participants also completed a demographics questionnaire either after the first test session or at the beginning of the second test session. Participants were asked to refrain from consuming caffeine 1 hr before the experiment began and to refrain from consuming alcohol within 24 hr of the experiment. Participants were compensated between $10 and $15 per hour depending on where they were recruited.

**Qualitative interview (C.T.).** Before beginning each memory test, the first author asked C.T. if she had heard of Poirot before, and assessed whether C.T. recognized the first author from the previous testing session.

**Prompted recall.** At test, participants were given four prompts corresponding to different sections of the episode (e.g., “On Christmas, Poirot finds something in the pudding and makes a plan to catch the thief.”) We chose to use a prompted recall design rather than free recall to provide additional support for C.T.’s memory retrieval. The experimenter read the prompts out loud and then instructed participants to tell them anything they could remember about that part of the episode.
Comprehension test. Participants were next given a comprehension test composed of 25 multiple-choice questions. These multiple-choice questions were created by the first author while consulting a script of the episode for accuracy. Twenty questions involved the content of the episode directly, and five questions asked about the personality of characters that were unique to that episode. Each question was accompanied by a corresponding screenshot taken from the episode and labelled with the names of the characters to assist C.T. with memory retrieval as much as possible.

Difficulty ratings (controls). Finally, at the end of Day 1, control participants were asked to rate the difficulty of the set of Day 1 comprehension questions on a scale of 1–7 (1 = very difficult, 7 = very easy). On Day 2, participants were asked to rate the difficulty of the set of Day 2 comprehension questions and then to rate whether the comprehension test from Day 2 was easier, more difficult, or about the same level of difficulty compared with the test from Day 1.

Scoring

Prompted recall responses were transcribed and scored for number of correct unique details. Details were operationalized as unique pieces of information that were accurate and did not include repetitions from the prompt or any preceding prompts. In cases where two details were closely related, they were counted as two unique details if they could reasonably have been remembered independently from one another.

Statistical Analysis

We analyzed prompted recall and comprehension test results for C.T. and age-matched controls using multilevel models (MLMs). Bayesian MLMs offer several advantages for single case designs, including that Bayesian analysis is not based on asymptotic theory, so are better suited to making inferences with small sample sizes (Rindskopf, 2014). For this reason, we ran Bayesian MLMs in the probabilistic programming language Stan (Carpenter et al., 2017) using the brms package (Bürkner, 2017) in RStudio Version 4.0.3 (RStudio Team, 2020). Posterior distributions were obtained using Markov Chain Monte Carlo sampling with chains of 10,000 iterations each, 2000 of which were the warm-up phase. To assess convergence, we checked that the Rhat values were less than 1.1.

We ran models with a random slope for condition (nap vs. wake) and a random intercept for each participant. We investigated each dependent variable as a function of condition (nap vs. wake), group (patient vs. control),
episode (A vs. B), as well as an interaction between condition and group. We estimated fixed effects for condition, group, episode, and the interaction between condition and group, random slopes for condition, and a random intercept for each participant. We used the following formula:

\[
\text{Memory (prompted recall or comprehension test) } \sim \text{condition (nap vs wake) } \times \text{group (patients vs control)} + \text{condition (nap vs wake) } + \text{group (patient vs control)} + \text{episode (A vs B)} + (1 + \text{condition}|\text{group})
\]

(1)

Condition (nap vs. wake), group (patient vs. control), and episode (A vs. B) were all contrast coded (nap = 1, wake = −1; patient = 1, control = −1; episode a = 1, episode b = −1). Intercept and \(\beta\) were given vague (weakly informative) Gaussian priors centered on 0 with a standard deviation of 10. The residual standard deviation was given a Half-Cauchy prior, which restricts the range of positive values. We calculated the Savage-Dickey Bayes Factor, which is the 95% chance that the parameter is in this interval; and the posterior distribution of an effect against a standard deviation of 10. We use the following formula:

\[
\text{Savage-Dickey Bayes Factor} = \frac{\int_{0}^{\infty} L(\theta) \, d\theta}{\int_{-\infty}^{\infty} L(\theta) \, d\theta}
\]

where \(L(\theta)\) is the likelihood of the posterior distribution. We report beta estimate; 95% credible interval, and the posterior distribution of an effect against a standard deviation of 10. We use the following formula:

\[
\text{Posterior Distribution} = \frac{\int_{0}^{\infty} L(\theta) \, d\theta}{\int_{-\infty}^{\infty} L(\theta) \, d\theta}
\]

We acquired the following MRI sequences: a. A sagittal T1-weighted 3-D magnetization prepared rapid gradient echo sequence (1-mm iso, TR/TE = 2300/2.98 msec, flip angle = 9°, 192 slices, 1-mm slice thickness, matrix = 256 × 256); b. Three sets of axial diffusion-weighted single shot spin-echo sequences with EPI readout (2-mm iso, TR/TE = 3900/73 msec, flip angle = 90°, 70 slices, 2-mm slice thickness, matrix = 122 × 122, P ➔ A phase encoding direction) corresponding to three b-value and gradient direction pairs: b = 1000 sec/mm² at 30 directions, b = 1600 sec/mm² at 40 directions, and b = 2600 sec/mm² at 60 directions, with interleaved b = 0 images. Two additional b = 0 images were collected before and after the main sequence with reverse phase encoding direction (P ➔ A).

MRI Scan

Control Participants

We recruited 11 healthy age-matched controls (M\(_{\text{age}}\) = 19.36 years, SD\(_{\text{age}}\) = .92 years) from the University of Toronto (via department listservs and research communities) and the broader Toronto community (via social media). As described above for controls in the memory test, the sample size was chosen to be similar to previous single-case studies (King et al., 2004; Schacter et al., 1996). Participants were included if they met the following criteria: between 18 and 21 years of age, not claustrophobic, able to keep still for 45 min to an hour, and no contraindications for MRI scanning. Participants indicated their gender as woman (n = 8), man (n = 2), and nonbinary person (n = 1), and their sex as female (n = 9) and male (n = 2). Participants indicated that they were right-handed (n = 10) or left-handed (n = 1).

All participants gave written informed consent to participate in the study, which was approved by the University of Toronto ethics board. Participants met the experimenter outside of the MRI facility and were screened a second time for MRI safety. They watched a nature documentary during the scan and were not given any specific instructions. Scanning took approximately 45–60 min. Participants also completed a short demographics questionnaire.

Clinical MRI Acquisition

We obtained C.T.’s presurgical clinical MRI scan from March 2020 for visual comparison to the postsurgical MRI scans acquired in April 2022. Presurgical images were collected on a Philips Achieva 3 T scanner with an 8-channel head coil and included a sagittal 3-D T1-weighted turbo field echo sequence (0.5 × 0.53 × 0.53 mm³ reconstructed voxel resolution, repetition time/echo time [TR/TE] = 4.97/2.3 msec, flip angle = 8°, 1-mm slice thickness with 0.5-mm spacing between slices, matrix = 220 × 220 reconstructed to 432 × 432).

Research MRI Acquisition

MRI data were acquired on a 3 T Siemens Magnetom Prisma scanner (Siemens Medical Solutions) at the Toronto Neuroimaging Facility using a 32-channel head coil. Foam padding was used to minimize head motion. We acquired the following MRI sequences:

a. A sagittal T1-weighted 3-D magnetization prepared rapid gradient echo sequence (1-mm iso, TR/TE = 2300/2.98 msec, flip angle = 9°, 192 slices, 1-mm slice thickness, matrix = 256 × 256); b. Three sets of axial diffusion-weighted single shot spin-echo sequences with EPI readout (2-mm iso, TR/TE = 3900/73 msec, flip angle = 90°, 70 slices, 2-mm slice thickness, matrix = 122 × 122, P ➔ A phase encoding direction) corresponding to three b-value and gradient direction pairs: b = 1000 sec/mm² at 30 directions, b = 1600 sec/mm² at 40 directions, and b = 2600 sec/mm² at 60 directions, with interleaved b = 0 images. Two additional b = 0 images were collected before and after the main sequence with reverse phase encoding direction (P ➔ A).

Structural MRI Processing

Structural T1-weighted images from the research MRI scan were processed using FastSurfer (v1.1.0, Henschel, Kügler, & Reuter, 2022; Henschel et al., 2020), which is an implementation of FreeSurfer’s recon-all structural processing pipeline that employs deep learning to reduce runtime of volumetric segmentation and cortical surface reconstruction. The pipeline produced a segmentation of subcortical regions with the aseg atlas (Fischl et al., 2002), from which bilateral hippocampal segmentations were obtained for volume comparison between C.T. and the control group and inclusion masks for fornix tractography. A measure of total brain volume (TBV) was also automatically calculated as the sum of aseg structure.
volumes including the cerebellum but excluding ventricles, cerebrospinal fluid, and dura (FreeSurfer output name BrainSegVolNotVent). The asseg atlas segmentation for each participant was manually inspected by J. T. for any issues and to ensure good delineation of the hippocampus. To account for differences in head size, we divided hippocampal volume by TBV and multiplied by 100 to express hippocampal volume as a percentage of TBV. We used single-case study $t$ tests (Crawford & Howell, 1998) to compare volumetric measures. In addition to the asseg atlas, participants’ structural scans also underwent surface-based registration to the Human Connectome Project’s Multimodal Parcellation (1.0; Glasser et al., 2016), which groups individual ROIs into lobes (e.g., lateral temporal).

**Diffusion MRI Processing**

**Preprocessing.** Diffusion-weighted images were preprocessed using a Pythonic implementation of the DESIGNER pipeline (Ades-Aron et al., 2018). Pipeline software dependencies include MRtrix3 (v3.0.3, Tournier et al., 2019), FMRIB Software Library (v6.0.5.1, Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012), and Python (v3.7.13). Briefly, preprocessing steps included PCA-based denoising (Cordero-Grande, Christiaens, Hutter, Price, & Hajnal, 2019; Veraart, Fieremans, & Novikov, 2016; Veraart, Novikov, et al., 2016), Rician bias correction (Veraart, Van Hecke, & Sijbers, 2011), Gibbs unringing (Kellner, Dhital, Kiselev, & Reisert, 2016), and EPI distortion, motion, and field inhomogeneity corrections (Smith et al., 2004).

Using MRtrix3, preprocessed diffusion-weighted images were upsampled to 1.25-mm isotropic resolution using a cubic b-spline interpolation. Upsampled data were used to estimate tissue response functions with the dbollander algorithm (Dhollander, Mito, Raffelt, & Connelly, 2019; Dhollander, Raffelt, & Connelly, 2016). Then, control-group-averaged tissue response functions were used to generate fiber orientation distribution (FOD) images for each individual with a multishell multituessue constrained spherical deconvolution approach (Jeurissen, Tournier, Dhollander, Connelly, & Sijbers, 2014; Tournier, Calamante, & Connelly, 2004). FOD images were further preprocessed with joint bias field correction and intensity normalization (Dhollander, Tabbara, et al., 2021; Raffelt, Dhollander, et al., 2017).

**Fixel-based group analysis.** The following steps were carried out using MRtrix3 (v3.0.3) to derive fixel metrics (for review and details, see Dhollander, Clemente, et al. [2021] and Raffelt, Tournier, et al. [2017]). Briefly, following individual FOD image generation, a population template was generated from the control group’s FOD images using an iterative registration and averaging approach (Raffelt et al., 2011). Then, apparent FD was estimated from the group template, identifying the number and orientation of fiber bundle elements (i.e., fixels) in each voxel. Each individual’s FOD image was nonlinearly registered to the group template, thereby achieving spatial correspondence between participants. This model yields three metrics to describe white matter: FD (a microstructural measure of white matter approximating intra-axonal volume), FC (a macrostructural measure of white matter sensitive to atrophy), and FD and cross-section (FDC: a measure combining both FD and FDC into an overall characterization of a white matter bundle). Using the whole-brain tractogram described in the following section, a fixel–fixel connectivity matrix was generated and used to smooth FD, FC, and FDC values. Finally, we subtracted the control group mean smoothed FDC from C.T.’s smoothed FDC and divided by the control group mean: This expression of FDC difference as a percentage relative to the control group mean was used to characterize fiber bundle differences in the corpus callosum, as well as projected onto the control-group-derived representative fornix tract.

**Tractography.** Whole-brain probabilistic tractography was carried out for C.T. and the group (using the population FOD template) with suggested parameters: iFOD2 algorithm (Tournier, Calamante, & Connelly, 2010), 20 million streamlines selected, angle 22.5°, min/max length = 10/250 mm, FOD amplitude power cutoff = 0.06. The resulting tractograms was further filtered by the spherical-deconvolution informed filtering of tractograms algorithm (Smith, Tournier, Calamante, & Connelly, 2013), which uses the FOD images to reduce reconstruction biases and decreases the number of streamlines to a requested total of 2 million streamlines.

The left and right fornices were traced out from the whole-brain tractograms with the following procedure: First, a fornix mask was obtained by warping FMRIB Software Library’s built-in fornix template (Brown et al., 2017) to C.T. and the template space, then dilating it to ensure coverage of bilateral fornices. Streamlines that met the following criteria were retained: (1) wholly contained within the fornix mask, (2) endpoints in either the left or right hippocampus, and (3) passes through a 5-mm sphere manually placed on the body of the fornix.

Tractograms describing hippocampal connectivity to different bilateral cortices were obtained with the following procedure: First, the Human Connectome Project’s Multimodal Parcellation 1.0 atlas parcellation was transformed to the group FOD template space using FreeSurfer and MRtrix3. Then, tract files were obtained by filtering the whole-brain tractograms for streamlines connecting the hippocampus and each ROI. These individual hippocampus-ROI files were combined into one tractogram describing connectivity from the hippocampus to the cortex of interest by grouping regions by their cortex identifier (as identified in the atlas documentation).
Statistical comparison. We employed a tractometry approach (Chandio et al., 2020): The control-group derived left and right fornices were resampled such that all streamlines within each fornix were defined by 20 coordinates. For each participant, FDC was sampled and averaged across streamlines, thereby allowing for statistical comparison of FDC differences at points along the fornix. Error bars representing 95% bootstrapped confidence intervals were generated for C.T.’s average FDC values (i.e., across streamlines), as well as for the control group mean FDC values (i.e., across participants).

Single-case study t tests (Crawford & Howell, 1998) were used to test which points along the fornix differed significantly between C.T. and the control group. Results were Bonferroni-corrected by dividing alpha (0.05) by the total number of coordinates (40 comparison points across the left and right fornix) for a significance threshold of 0.00125.

RESULTS

Behavioral Results

Standard Neuropsychological Tests

C.T.’s results on standardized neuropsychological tests are given in Table 1. In summary, she performed well (at 50th percentile or above) on tests of visual attention and task switching (the Trail Making Test), on an assessment of visual scanning, perceptual speed, and motor memory (main component of Digit Symbol Substitution), and on an assessment of recent memory (free recall section of Digit Symbol Substitution). However, she performed poorly (closer to or less than 1st percentile) on a verbal learning and memory task (the Rey Auditory Verbal Learning Test [RAVLT]), including both repetition (i.e., across streamlines), as well as for the control group mean FDC values (i.e., across participants).

Single-case study t tests (Crawford & Howell, 1998) were used to test which points along the fornix differed significantly between C.T. and the control group. Results were Bonferroni-corrected by dividing alpha (0.05) by the total number of coordinates (40 comparison points across the left and right fornix) for a significance threshold of 0.00125.

RESULTS

Behavioral Results

Standard Neuropsychological Tests

C.T.’s results on standardized neuropsychological tests are given in Table 1. In summary, she performed well (at 50th percentile or above) on tests of visual attention and task switching (the Trail Making Test), on an assessment of visual scanning, perceptual speed, and motor memory (main component of Digit Symbol Substitution), and on an assessment of recent memory (free recall section of Digit Symbol Substitution). However, she performed poorly (closer to or less than 1st percentile) on a verbal learning and memory task (the Rey Auditory Verbal Learning Test [RAVLT]), including both initial learning and a 25-min delayed recall. She also performed poorly on delayed recall and percent retained after a 16-min delay in a visuospatial memory task (BVMT-R). These results show that C.T. exhibits a memory impairment for information after short delays that do not include sleep.

C.T. also completed a spatial ability questionnaire—the Navigation Strategies Questionnaire (NSQ; Brunec et al., 2019). The NSQ is a 14-item self-reported measure that assesses whether participants are “scene-based” navigators (for an overall negative score) or “map-based” navigators (for an overall positive score). C.T. scored 2 on the NSQ, indicating that she is more of a “map-based” navigator, suggesting that she relies on an allocentric view of the environment in navigational contexts. Indeed, map-based navigators may rely more on an allocentric view of the environment and have greater flexibility in navigation than scene-based navigators (Marchette, Bakker, & Shelton, 2011). Anecdotally, her family reported that C.T.’s spatial navigation has perhaps improved but has certainly not worsened since her surgery.

Autobiographical Interview

Results for the autobiographical interview are given in Figure 1. C.T.’s results indicate that she has anterograde amnesia with no retrograde amnesia. For the final category (ages 17–19 years, which followed tumor diagnosis and surgery), C.T. could not recall any events that did not take place on the day of testing. When prompted to generate a memory from this recent time category, she thought for some time and then asked if she could recall an event from earlier in the day that had not been separated by sleep (approximately 6–8 hr before testing). This was a vivid memory with 15 internal details and three external details. Notably, this vivid memory stands in sharp contrast to C.T.’s inability to recall a memory even from the day prior. We also tested C.T. on her autobiographical memory using the HippoCamera smartphone app (Martin et al., 2022), in which automated text messages prompted her to record and replay memory cues of daily events over the course of several weeks. We tested her memory for 10 of these events, and she could not provide any episodic details.

Memory Test for TV Episode

Qualitative interview (C.T.). During both memory tests that followed a nap, C.T. neither recognized the experimenter nor remembered ever hearing of Poirot. In contrast, she remembered both the experimenter and Poirot during the two memory tests that following 1 hr 40 min of being awake. These observations are consistent with the other testing sessions reported in this article in which the first author tested C.T. after a day of being awake, immediately after sleep, or after many weeks had passed. Whenever the delay included an interval of sleep, C.T. never appeared to recognize the first author. However, when tested twice in the same day with no intervening sleep, C.T. had no noticeable memory deficit in general conversation. She would recognize the first author and refer to specific details in previous interactions (e.g., she would recall having coffee together in the morning 12 hr earlier).

Prompted recall. C.T. was not able to recall any details in both testing sessions after a nap, but was able to recall 7 and 11 details after an equivalent period of wake. Control participants recalled, on average, 43.6 details in the nap condition (SD = 29.32) and 45.8 details in the wake condition (SD = 23.84). We modelled the number of correct unique details recalled with a Poisson distribution because the number of details was measured in counts. Model results for prompted recall are given in Table 2. All Rhat values were 1, suggesting posterior distribution convergence. We found evidence of an interaction between Condition and Group (b = −114.97, 95% CrI [−299.59, −7.88], BF < 0.01). Follow-up analyses to investigate this interaction revealed that C.T. recalled fewer details than...
controls in the nap condition, \( b = -102.19 \), 95% Crl \([-267.46, -3.65]\), but not the wake condition, \( b = -1.53 \), 95% Crl \([-3.18, .07]\). Estimated marginal means for controls were 40.73 details in the nap condition, 95% Crl \([22.65, 62.89]\), and 35.44 details in the wake condition, 95% Crl \([17.03, 58.61]\). Estimated marginal means for C.T. were 0.00 details in the nap condition, 95% Crl \([.00, .00]\), and 8.73 details for the wake condition, 95% Crl \([.45, 31.56]\). Condition, Group, and Episode did not predict number of details recalled. Although the main effect of Group was not significant, C.T. performed numerically worse than all controls, suggesting a memory impairment in all conditions. Thus, in summary, we found evidence that sleep has a disproportionate impact on C.T.’s memory.

### Table 2. Mean, Standard Error, 95% Credible Interval, and Rhat Statistic for Each Parameter of a Model Predicting Prompted Recall

<table>
<thead>
<tr>
<th>Parameter</th>
<th>M</th>
<th>SE</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Rhat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.80</td>
<td>.25</td>
<td>3.31</td>
<td>4.30</td>
<td>1.00</td>
</tr>
<tr>
<td>Condition (nap vs. wake)</td>
<td>-.14</td>
<td>.20</td>
<td>-.55</td>
<td>.27</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (patient vs. controls)</td>
<td>-1.52</td>
<td>.81</td>
<td>-3.12</td>
<td>.15</td>
<td>1.00</td>
</tr>
<tr>
<td>Episode (A vs. B)</td>
<td>-.22</td>
<td>.18</td>
<td>-.54</td>
<td>.16</td>
<td>1.00</td>
</tr>
<tr>
<td>Condition × Group</td>
<td>-114.97</td>
<td>80.04</td>
<td>-299.59</td>
<td>-7.88</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Values in **bold** indicate 95% credible intervals that fall outside of 0.
imporation, but her memory is not fully intact over periods of wake.

Comprehension test. Scores on the comprehension test are given as proportion of correct responses. C.T. scored 0.48 and 0.28 on the comprehension test in the nap condition and 0.60 and 0.44 in the wake condition. C.T. performed numerically better in the wake condition in both testing sessions. Control participants scored, on average, 0.92 in the nap condition (SD = 0.07) and 0.92 in the wake condition (SD = 0.10). We investigated responses on each trial of the multiple-choice comprehension test with a binomial distribution. Model results for the comprehension test are given in Table 3. All Rhat values were 1, suggesting posterior distribution convergence. There was no evidence that Condition, Group, or Episode played a role in responses on comprehension test memory. It is important to note that chance for this test was at 25% and the differences between C.T.’s performance across conditions were small. For example, the difference between her comprehension test score for Episode B in the nap condition (0.48) and Episode B in the wake condition (0.44) was only one question. C.T. performed numerically worse than all controls on the comprehension test.

Difficulty ratings (controls). On average, control participants rated the difficulty of the comprehension test for Episode A as 5.3 on a 1–7 point scale (slightly easy, SD = 1.16). They rated the difficulty of the comprehension test for Episode B as 6 on a 1–7 point scale (fairly easy, SD = 0.94). A paired t test between difficulty ratings for Episode A and Episode B showed no significant differences in difficulty between episodes, t(9) = −2.09, p = .07. When comparing difficulty between Episode A questions and Episode B questions directly, participants indicated an average of 4.7 on a 7-point scale (Episode A as slightly easier, SD = 1.13). A t test comparing relative difficulty ratings against 4 (neither more difficult nor easier) was not significant, t(9) = 1.91, p = .09. Overall, these ratings suggest that the comprehension tests for Episode A and Episode B were well-matched and not significantly different from each other in terms of difficulty.

Table 3. Mean, Standard Error, 95% Credible Interval, and Rhat Statistic for Each Parameter of a Model Predicting Comprehension Test Performance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>M</th>
<th>SE</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Rhat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.27</td>
<td>.58</td>
<td>2.27</td>
<td>4.56</td>
<td>1.00</td>
</tr>
<tr>
<td>Condition (nap vs. wake)</td>
<td>−.18</td>
<td>.54</td>
<td>−1.28</td>
<td>.89</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (patient vs. controls)</td>
<td>−2.86</td>
<td>1.44</td>
<td>−5.96</td>
<td>.06</td>
<td>1.00</td>
</tr>
<tr>
<td>Episode (A vs. B)</td>
<td>−.60</td>
<td>.31</td>
<td>−1.21</td>
<td>.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Condition × Group</td>
<td>−.46</td>
<td>1.23</td>
<td>−3.23</td>
<td>2.04</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Values in bold indicate 95% credible intervals that fall outside of 0.

MRI Results

Impact to Fornix and Corpus Callosum, But No Change to Hippocampus

Initial visual inspection of the sagittal slices from T1-weighted structural scans showed differences in the fornices in C.T. pre- and postsurgery, as well as in comparison to a control participant (Figure 3). Evidence of the transcallosal approach is appreciated in the comparison of the corpus callosum from pre- to postsurgical scans, as well as observations of reduced FDC in the anterior corpus callosum (Figure 3B) from FBA.

C.T.’s left hippocampal volume was 4309 mm³ (0.37% of TBV), and here right hippocampal volume was 4267 mm³ (0.36% of TBV). The control group’s average left hippocampal volume was 4094 mm³ (0.36% of TBV, SD = 305.7 mm³), and their average right hippocampal volume was 4254 mm³ (0.37% of TBV, SD = 332.1 mm³). Volumetric analysis with hippocampus volume as a percentage of TBV showed no significant differences between C.T. and the control participants in the left or right hippocampus (left: t(10) = 0.422, p = .341; right: t(10) = −0.224, p = .414), nor did C.T. differ in TBV from the control group, t(10) = 0.422, p = .341.

Characterizing Fornix White Matter Using Tractography and FBA

Probabilistic tractography showed that C.T. had fewer streamline counts in the right versus left fornix (246 vs. 644 streamlines; Figure 4A and B), as well as in C.T. relative to control participants (control group right vs. left: 966 vs. 1087 streamlines).

Initial inspection of FBAs showed that the greatest percentage decrease in FDC in C.T. relative to the control group was in the right anterior fornix column, but those differences extended through the body and crus (Figure 4C). Consistent with this observation, sampling at 20 points along the right fornix revealed that FDC was significantly decreased compared with controls in 11 out of 20 of the coordinates (−11.62 ≤ ts ≤ −4.17, Bonferroni-corrected 7.89e−6 ≤ ps < .039; Figure 4D).
Figure 3. Fornix and corpus callosum pre-/post-surgery in C.T. compared with control participant(s). (A) Red line on axial image identifies the midline sagittal slices. White arrows indicate the corpus callosum and fornix on a control participant scan. (B) Percentage decrease in FDC in C.T. relative to control group overlaid on the control-group-derived white matter template. Crosshair indicates slice positions on (cropped) sagittal, coronal, and axial planes.

Figure 4. Differences in fiber bundle density and cross-section (FDC) in fornix. (A) Tractography of fornix in C.T. overlaid on sagittal T1-weighted image, (B) as well as 3-D volume renderings showing fewer streamlines in right versus left fornix. (C) Percentage decrease of FDC in C.T. relative to control group projected onto group-derived template fornix. (D) Fiber bundle differences localized to right fornix from anterior column through body to crus. *Indicates $p < .00125$. 
In contrast, only 3 out of 20 points in the column and body of the left fornix had significantly decreased FDC compared with the control group ($-6.60 \leq t_s \leq -4.25$, Bonferroni-corrected $0.001 \leq p_s < 0.034$; Figure 4D).

As FDC combines both FD and FC, examining the constituent metrics revealed that almost all the coordinates with decreased FDC also had decreased FD (left fornix: two out of three coordinates, right fornix: 8 out of 11 coordinates). In contrast, there were no significant FC differences in the left fornix, and only two coordinates with increased FC in the right fornix. The near complete overlap between decreased FD and FDC suggest that overall differences in the right fornix are driven by microstructural changes (i.e., lower intra-axonal volume).

To explore the potential impact of fornix perturbation on hippocampal connectivity with distributed brain regions, FDC was also projected onto control-group-derived filtered tractograms representing structural connectivity between the hippocampus to bilateral regions grouped by cortex. As seen in Figure 5, C.T.’s fornix damage may disrupt connections between the hippocampus and lateral temporal, inferior frontal, anterior cingulate, and medial prefrontal regions.

**DISCUSSION**

Here, we report case C.T., a young woman experiencing anterograde amnesia following impact to the mid-anterior corpus callosum and the right fornix, beginning from the column but extending through the body and crus. Her hippocampus was intact. Because the fornix acts as a key output pathway for the hippocampus, this lesion may be associated with widespread disrupted connectivity of the hippocampus to other brain regions, including the lateral temporal, inferior frontal, anterior cingulate, and medial prefrontal lobes. In two separate testing sessions, we found that sleep differentially and profoundly impaired her memory for information acquired before sleep. Following a nap, she was unable to recall any details from a TV episode watched immediately before sleep. In contrast, C.T. could recall several details from the episode if her memory was tested following an equivalent time period of wakefulness, and her recall in this wake condition did not differ statistically from controls. Moreover, after a nap, C.T. neither recognized the experimenter nor remembered ever hearing of the TV show. In contrast, she showed no noticeable memory impairment in general conversation when she stayed awake for the same amount
of time. The autobiographical interview revealed anterograde amnesia with no retrograde amnesia, and she was unable to generate a memory from the time period following her surgery, unless that event had occurred on the day of testing and was not separated by sleep. These findings reveal a robust memory impairment in C.T. following sleep. However, we wish to highlight that although she could remember some details in the wake condition, her memory performance was nonetheless numerically worse than all controls. In addition, C.T. performed poorly on short waking delays for auditory verbal learning and visuospatial memory. Thus, her memory also appears to be impacted during periods of wake, albeit to a much lesser extent relative to sleep. We conclude that the fornix plays a critical role in processing episodic memories during sleep, and in C.T., sleep impaired memory more than the mere passage of time.

As far as we know, there has only been one other case of sleep-related amnesia described in the previous literature. Patient F.L. is a woman who developed functional amnesia with no obvious neural correlate following a car accident (Smith et al., 2010). During the day, F.L.’s memory was normal. However, following a night of sleep, she reported that she did not remember the previous day. On formal tests, F.L. performed moderately well on material that she had learned that day and showed no memory for material she had learned the previous day, but also performed moderately well on material that had been covertly presented from previous days (Smith et al., 2010). Although there are several commonalities across F.L. and C.T., including an absence of retrograde amnesia, losing memories after a night of sleep, and waking up with the sense that the current day is the day after the accident or surgery, there are also several important differences. Critically, there was no visible neural damage in F.L. and naps did not impair her memory. F.L.’s condition improved when she began to interrupt her sleep at 4-hr intervals, which has allowed her to retain memories over multiple days. C.T.’s memory did not improve from reducing the length of her naps and her profile of memory impairment has not changed in the more than 3 years since her surgery. Therefore, although some aspects of their memory presentation overlap, we consider these cases to be too distinct to have a common cause.

It is important to consider whether rehearsal may explain C.T.’s relative sparing of memory over periods of wake. Although we certainly accept that C.T. might have been rehearsing some of the material, we think that this is unlikely to explain the entirety of her memory profile. For example, there were several instances in which C.T. provided details that she did not know would be part of a memory test or were so incidental it is unlikely she would have deemed them sufficiently important to rehearse. For example, in one case, the first author spoke to C.T. in the morning and, that evening, C.T. referred to drinking coffee together that morning. In another case, she described a memory from several hours ago to the first author that evening at 9 p.m. She was not informed that she was going to be asked about her autobiographical memories that day, so it is unlikely that she was rehearsing that information. For these reasons, we believe that rehearsal cannot be the primary explanation for C.T.’s superior performance in the wake condition. Moreover, the extent of her memory loss following sleep was so profound that it is hard to imagine how rehearsal could have protected against this memory loss. For example, she lost not only memory for the name of the experimenter, but also memory for every episodic aspect of that morning’s testing session.

It is also possible that the differences we see in C.T.’s memory are exacerbated by the greater contextual difference between sleeping and waking up compared with staying awake. The hippocampal system is important for maintaining memories across contexts, such as across tasks (e.g., Scoville & Milner, 1957) and after a delay (e.g., Baddeley & Wilson, 2002), as well as reinstating context to retrieve a memory (e.g., Tanaka et al., 2014). However, there were not necessarily more contextual changes in the nap condition compared with the wake condition. C.T. was almost certainly exposed to more changes in context during the wake condition, as she completed various tasks in a variety of environments, compared with only going to her room and taking a nap. Importantly, C.T.’s encoding and testing for both the wake and nap conditions took place in the same spatial context. It is possible that C.T. mitigated the changes in context in the wake condition by rehearsing the episode for the duration of the waking interval. However, as discussed above, based on other test sessions with C.T., we do not believe that her memory retention over long waking intervals can be solely attributed to rehearsal.

It is currently unclear what specifically about C.T.’s sleep may be impairing her memory. A limitation of the current study is that we were not able to collect sleep physiology data in C.T. during her naps and therefore cannot directly speak to whether C.T.’s pattern of memory impairment is related to differences in specific sleep stages or other sleep disruptions. However, in a previous sleep study, C.T. showed reduced SWS, in line with patients with hippocampal damage (Spanò et al., 2020). It is possible that damage to the broader hippocampal system reduces SWS, which may influence memory consolidation. A recent meta-analysis found that SWS positively predicts memory in young adults (Hokett et al., 2021), suggesting that a reduction in SWS may reduce memory consolidation. One proposed mechanism for how SWS consolidates memories is the reactivation of recently encoded hippocampal representations (Paller, Mayes, Antony, & Norman, 2020; Lewis, Knoblich, & Poe, 2018; Oudiette, Antony, Creery, & Paller, 2013), which is thought to both strengthen the hippocampal-dependent aspects of the memory representation and project reactivated memory information to neocortical and striatal networks (Inostroza & Born, 2013). Specifically, hippocampal representations are thought to be reactivated during the transition.
between cortical down-states marked by slow-oscillations (~0.75 Hz) and cortical up-states (Maingret, Girardeau, Todorova, Goutierie, & Zugaro, 2016). At the slow-oscillation down-state, thalamic cortical spindles (9–15 Hz) trigger sharp-wave ripple activity (100–250 Hz) in both the hippocampus and the cortex, promoting hippocampal–cortical communication (Maingret et al., 2016; Navarrete, Valderrama, & Lewis, 2020). Although these processes may be impacted in C.T., it is unlikely that a reduction in SWS could explain the scale of C.T.’s memory impairment following periods of sleep. Furthermore, selective SWS deprivation does not result in changes to memory consolidation in healthy participants (Genzel, Dresler, Wehrle, Grözinger, & Steiger, 2009).

One speculative explanation for our results is that when C.T. is awake, she can encode and partially retain some memories in her intact hippocampus. However, during sleep, the coordination between the hippocampus and the cortex may be disturbed by damage to her fornix, disrupting memory replay and making the memories no longer retrievable. Incorrect or nonspecific memory replay in the hippocampal circuit during sleep could explain why children with Down syndrome show a memory impairment following sleep (Spanò et al., 2018) or no benefit (Ashworth et al., 2017) compared with typically developing children who show a memory benefit from sleep (Spanò et al., 2018). C.T.’s hippocampal system may also be replaying incorrect or nonspecific activity. Importantly, memory replay has also been shown to take place in animals during wake (Kudrimoti, Barnes, & McNaughton, 1999), which may play a role in both memory consolidation and memory retrieval (Findlay, Tononi, & Cirelli, 2020; Carr, Jadhav, & Frank, 2011). Disrupted memory replay during wake could explain why C.T.’s memory also seems to be affected after waking intervals.

The impact to C.T.’s brain is localized to the fornix and the middle of the corpus callosum. Aspects of her memory presentation, anterograde amnesia without retrograde amnesia, are in line with previous cases of fornix damage (Rizek et al., 2013; Murr et al., 2012; Park et al., 2000; Cabalrese et al., 1995; Hodges & Carpenter, 1991). However, damage to the corpus callosum is also associated with impairment to recent memory. In particular, reductions in white matter integrity in subregions of the corpus callosum have been associated with poorer verbal and visuospatial working memory in healthy individuals (Treble et al., 2013) and in patients with multiple sclerosis (Bodini et al., 2013; Hasan, Gupta, Santos, Wolinsky, & Narayana, 2005). Consistent with these findings, some epilepsy patients who undergo anterior corpus callosotomies have moderate memory deficits, particularly in verbal and visuospatial memory (Phelps, Hirst, & Gazzaniga, 1991; Spencer, 1988; Zaidel & Sperry, 1974). However, the moderate level of impairment and inconsistency across patients leads us to believe that the corpus callosum is not likely to be the main explanatory factor in C.T.’s anterograde amnesia (Wong et al., 2006; Mamelak, Barbaro, Walker, & Laxer, 1993). Furthermore, existing case reports describing amnesia alongside corpus callosum damage are often associated with damage to the fornix or surrounding structures (Zhang, Zhang, Jiang, Lv, & Dong, 2022; Ren et al., 2018; Saito, Matsumura, & Shimizu, 2006; Park et al., 2000; Clark & Geffen, 1989). Lastly, many brain surgeries are accessed through the corpus callosum without resulting memory loss.

In addition, the role of the corpus callosum in sleep and sleep-dependent memory consolidation is not well researched. Patients who underwent callosotomy had significantly reduced probability of the propagation of slow-waves across hemispheres, suggesting the corpus callosum is necessary for cross-hemispheric slow-wave propagation (Avvenuti et al., 2020). Greater axial diffusivity in the anterior corpus callosum was associated with higher spindle power (Piantoni et al., 2013), indicating a link between the structure of the anterior corpus callosum and sleep physiology. However, it is unknown whether the structure of the corpus callosum and its link to sleep physiology is related to memory consolidation.

Another patient group that has been thought to have an altered relationship between sleep and memory are patients with transient epileptic amnesia (TEA). Some patients with TEA experience accelerated long-term forgetting, a memory impairment in which learning and memory appear normal at initial test but are followed by rapid forgetting. Although an impairment in memory consolidation during sleep was proposed as a possible explanation for accelerated long-term forgetting, several studies found that sleep equally benefited memory retention in both controls and patients with TEA (Atherton et al., 2016; Atherton, Nobre, Zeman, & Butler, 2014) and that patients with TEA displayed a memory impairment following 3–8 hr of wake (Hoefeziers, Dewar, Della Sala, Butler, & Zeman, 2015). Furthermore, patients with TEA performed worse than controls only in the 12-hr wake delay condition and not the sleep-delay condition (Atherton et al., 2014). However, more SWS predicted better memory consolidation in healthy controls, whereas more SWS predicted a smaller memory benefit in patients with TEA (Atherton et al., 2016), suggesting that memory consolidation during SWS in patients with TEA may be distinct from healthy controls.

Notably, in the current study, we did not see a difference in recall memory between sleep and wake conditions for control participants. Although this could suggest that sleep does not play a role in the consolidation of naturalistic stimuli in healthy controls, we believe that our design does not allow us to answer this question. Our prompted recall measure was selected to scaffold memory recall as much as possible for C.T. and controls were likely performing at the top of their memory capacity. If sleep does play a role in the consolidation of naturalistic stimuli, it would be best captured by either a free recall measure (as in Coutanche et al., 2020) or the inclusion of more videos to mitigate ceiling effects in controls. Future work
could investigate memory for more video clips across longer sleep and wake delays.

Future research could also investigate the role of sleep in other memory disorders with hippocampal system damage to determine whether there is additional evidence of sleep impairing memory. As discussed above, Spanò and colleagues (2018) suggested that in children with Down syndrome, the hippocampal circuit may replay incorrect or nonspecific activity during sleep that could interfere with memory consolidation. It would be interesting to investigate activity in the hippocampal circuit during sleep in patients with an altered sleep and memory relationship. In addition, it would be interesting to investigate the role of the fornix in memory consolidation during sleep, such as with MRI studies looking at whether fornix microstructure predicts differences in memory consolidation across sleep and wake delays. Furthermore, testing C.T.’s memory using implicit methods, such as eye-tracking, could help to determine whether C.T. can retain memories after sleep that she cannot consciously express (Ryan & Shen, 2020). Testing C.T.’s procedural or priming memory after sleep and wake delays could also help clarify whether implicit and motor memories are also subject to memory loss during sleep.

In conclusion, we present C.T., a young woman experiencing anterograde amnesia following impact to the midsection of the corpus callosum and the right fornix. Although C.T. also experienced difficulties in learning and memory during wake, sleep differentially and profoundly impaired her memory recall. This work provides evidence that the fornix plays a critical role in processing memories during sleep. We speculate that the fornix may ensure that specific memories are replayed during sleep, maintain the balance of sleep stages, or allow for the retrieval of memories following sleep.

Acknowledgments
We would like to thank C.T. and her family for their collaboration. We would also like to thank Miranda Chang for assistance with data collection and scoring the autobiographical interview, Aaliyah Mulla for data collection and scoring, and Dr. Bryan Hong for his assistance with statistics.

Reprint requests should be sent to Nelly Matorina, Department of Psychology, University of Toronto, or via e-mail:nelly.matorina@mail.utoronto.ca.

Data Availability Statement
Behavioral data are available on the Open Science Framework – https://osf.io/cnf69/?view_only=4e4dd5310ff4ed7b2239742d1610e9f.

Author Contributions
Nelly Matorina: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Writing—Original draft; Writing—Review & editing. Julie Tseng: Data curation; Formal analysis; Methodology; Software; Visualization; Writing—Original draft; Writing—Review & editing. Natalia Ladyka-Wojcik: Methodology; Resources. Rosanna Olsen: Investigation; Methodology; Supervision; Validation; Writing—Review & editing. Donald J. Mabbott: Methodology; Project administration; Resources; Supervision; Writing—Review & editing. Morgan D. Barense: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Validation; Writing—Review & editing.

Funding Information

Diversity in Citation Practices
Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the Journal of Cognitive Neuroscience (JoCN) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, 2021), the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, JoCN, 34:1, pp. 1–3). Consequently, JoCN encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: M/M = .433; W/M = .268; M/W = .155; W/W = .144.

REFERENCES
Adamovich, B. L., Gualberto, G., Roberts, T., Haut, M. W., & Gutmann, L. (2009). Teaching NeuroImages: Amnesia due to
for the first time since the day is slowly getting her memory back.


Matorina et al. 1651


