ORIGINAL ARTICLE

Sleep deprivation impairs binding of information with its context

Courtney A. Kurinec[1,2,✉], Paul Whitney[1,2,*], John M. Hinson[1,2,✉], Devon A. Hansen[2,3] and Hans P. A. Van Dongen[2,3,✉]

[1]Department of Psychology, Washington State University, Pullman, WA, [2]Sleep and Performance Research Center, Washington State University, Spokane, WA and [3]Elson S. Floyd College of Medicine, Washington State University, Spokane, WA

*Corresponding author. Paul Whitney, Department of Psychology, Washington State University, P.O. Box 644820, Pullman, WA 99164-4820. Email: pwhitney@wsu.edu.

Abstract

Binding information to its context in long-term memory is critical for many tasks, including memory tasks and decision making. Failure to associate information to its context could be an important aspect of sleep deprivation effects on cognition, but little is known about binding problems from being sleep-deprived at the time of encoding. We studied how sleep deprivation affects binding using a well-established paradigm testing the ability to remember auditorily presented words (items) and their speakers (source context). In a laboratory study, 68 healthy young adults were randomly assigned to total sleep deprivation or a well-rested control condition. Participants completed an affective item and source memory task twice: once after 7-hour awake during baseline and again 24 hours later, after nearly 31 hours awake in the total sleep deprivation condition or 7 hours awake in the control condition. Participants listened to negative, positive, and neutral words presented by a male or female speaker and were immediately tested for recognition of the words and their respective speakers. Recognition of items declined during sleep deprivation, but even when items were recognized accurately, recognition of their associated sources also declined. Negative items were less bound with their sources than positive or neutral items, but sleep deprivation did not significantly affect this pattern. Our findings indicate that learning while sleep-deprived disrupts the binding of information to its context independent of item valence. Such binding failures may contribute to sleep deprivation effects on tasks requiring the ability to bind new information together in memory.

Statement of Significance

Making associations between information and its context in long-term memory is important for many common tasks, including learning stimulus-response associations from feedback or remembering who committed an action. While it is generally accepted that learning during sleep deprivation affects retention, item memory deficits are dissociable from the process of binding items to the context in which they were produced. Our results show that learning during sleep deprivation not only reduces recognition of items, but also interferes with the binding of correctly recognized items to their source contexts. Source memory errors produced by sleep deprivation's effect on binding could critically impact real-world situations such as eyewitness accounts and the influence of misinformation on people's decision making.

Introduction

Being able to remember the context something was presented in—the who, when, and where—is key to the ability to form functional and detailed memories of events. Having functional memory for specific items and their associated contexts is essential for navigating daily life. Context memory failures are often frustrating but harmless, such as being unable to recall where you left your house keys, but in certain situations, inadequate context memory can have serious consequences. For example, if an eyewitness fails to encode the color and model of a car and associate it with the crime they witnessed, it could impede a criminal investigation. Similarly, if an individual correctly encodes an erroneous claim made by a speaker but associates it with the wrong person, they may attribute the information to an actual expert rather than a debunked source, resulting in the perpetuation of misleading information.

An important factor influencing memory processes is sleep [1, 2]. Studies that have examined the role of sleep in the retention of memory for contexts have generally found that sleeping, compared to foregoing sleep, results in better contextual memory performance [3–6]. As with many investigations of memory and sleep, these studies have focused primarily on how sleep (or the lack thereof) affects memory consolidation processes. That is, these studies examined how well information that was learned when rested had been stabilized and retained in long-term memory after a period of sleep or wake. However, consolidation is not the only memory process that is potentially vulnerable to sleep loss. Insufficient sleep has also been found to impair the ability to *encode* or convert information into a format that can be stored in the first place [7–9]. Conversely, limited research investigating sleep loss and retrieval, or the process of using internal or external cues to access stored memories, suggests that this process may be generally preserved during total sleep deprivation (TSD) [10, 11].

Contextual memory is particularly important to investigate under conditions of TSD at the time of encoding because remembering the context in which an item was presented is inherently dependent on attending and processing that information while it was present [12]. If the item and its context are not encoded in the first place, there will be nothing to later consolidate or retrieve from memory. Compared to research on consolidation, however, only a few studies have investigated how TSD at the time of encoding impacts memory for context [13–15]. Use of idiosyncratic paradigms and the introduction of possible confounds, such as sleep (and consolidation) after learning and before testing, preclude any definitive conclusions about the effect of learning while sleep-deprived on contextual memory. Given that people are often not fully rested when they need to commit an item's context to memory and that the consequences of poor memory for context can be substantial, more conclusive research is needed to establish how learning during TSD affects contextual memory.

Investigating contextual memory also provides an opportunity to assess the effect of TSD on the ability to *bind* the encoded information in long-term memory. Binding, or the linking of independent neuronal representations to form associations, is an essential part of cognition at multiple levels, from perception to memory to decision making [16]. Several lines of research suggest that sleep-deprived individuals show deficits on a variety of cognitive tasks that require binding in long-term memory,

including decision-making tasks where learning from feedback is necessary for advantageous performance [17–21]. For example, decision making on the Iowa Gambling Task (IGT), a task that requires associating outcomes of choices with the options that produced the outcomes, is impaired by TSD [18]. Notably, memory for context has been found to account for sizeable variance in performance on the IGT [22] because people have to bind the outcomes of their choices to the source (i.e. the deck chosen on each trial) [23]. Thus, if the binding of items and contexts is impaired by TSD, failures of binding may explain some of the effects of TSD on cognitive tasks that rely on memory for associations between choices and outcomes [20].

Contextual memory is often assessed using what is known as a source memory paradigm. In a typical source memory paradigm, a series of items are presented in context, and participants are later tested for recognition of both *item memory*, or memory for the specific stimuli, and *source memory*, or memory for the context the stimuli were originally presented in. For example, in a common implementation of a source memory task, to-be-remembered words are spoken by either a female or male speaker. Testing recognition memory for the words assesses item memory, and, for words that are correctly recognized, the ability to identify the voice of the speaker assesses whether items were bound to their sources. Although correct memory for a source requires that the item is remembered, neuroimaging, clinical, and behavioral studies have found item and source memory to be dissociable processes [24–32].

There is suggestive evidence from neuroscience that source memory in particular, and long-term memory binding in general, may be vulnerable to TSD. The hippocampus, which is critical for binding items and contexts [33, 34] and is often involved in their retrieval [35, 36], has been found to show decreased activation during TSD or disrupted sleep, leading to impairments in encoding and retention of new episodic memories [4, 8, 37, 38]. Additionally, TSD at the time of encoding can result in less stable and poorer quality memory representations [9]. These findings point to a vital role of sleep in preparing the hippocampus for learning and retaining items and their contexts. Unlike memory for items, which can sometimes be supported by mediotemporal lobe (MTL) structures other than the hippocampus [25, 30, 39–41] that may be spared during sleep loss [42], source memory depends on binding processes supported by the hippocampus [43]. As a result, disruptions to hippocampal functioning caused by TSD may be particularly detrimental to source memory.

The prefrontal cortex (PFC) has also been implicated in the binding of items and contexts, in both neuroimaging and patient studies [25, 34], such that patients with frontal lobe lesions who have similar item memory performance as control participants still show deficits in source memory compared to controls [31]. Like the hippocampus, the PFC is believed to be vulnerable to TSD deficits [44], with lower activation during sleep deprivation and potentially worse behavioral performance on tasks that recruit the PFC [45, 46]. Interactions between the frontal and hippocampal regions contribute to forming and retrieving episodic memories [47], and source memory has been found to depend more heavily than item memory on the functional connections between the hippocampus and PFC [27]. This suggests that source memory, and other cognitive tasks that depend on the binding of novel associations [34, 48], may be particularly susceptible to disruptions of the hippocampus and PFC by TSD.

In the present study, we used a well-established item and source memory paradigm [24, 49–54] to test whether binding of sources to items that are correctly recognized is affected by TSD. We also manipulated the affective nature of items in long-term memory because the affective nature of the stimuli may influence how well sources are bound to items. For example, sleep-deprived persons may show selective enhancement of item memory for negative items [15, 55] and subsequently their sources, so that later recognition could be biased toward the context of negative events. Such potential valence effects merit serious attention, as studies outside of the TSD literature have commonly found that the affective valence of an item influences item memory and may also influence how items are bound to neutrally valenced sources in memory [50, 56–68]. Affective valence, particularly negative valence, has consistently been shown to promote memory for items as compared to neutral valence [56–60], and research suggests this pattern persists in sleep-deprived individuals [15, 55, 69–71]. However, how an item's valence affects its binding to (neutral) sources is not clear. Previous research has found source memory to be enhanced [57, 61–64], as well as unaffected or even hampered [50, 65–68] by item valence, and work exploring memory for the source of valenced items encoded and bound under TSD has been likewise inconsistent [14, 15].

In the version of the item and source memory task we employed, participants were presented with positive, negative, or neutral words spoken in a male or female voice and were immediately tested on their memory for the words (i.e. items) and their respective speaker (i.e. sources). By immediately testing participants' memory after study, we isolated the effect of TSD on the initial association of sources with items from the long-term consolidation of memories shown to be strengthened by sleep, which generally takes place over the course of several hours [72]. We employed a number of manipulations to separate both encoding and binding processes from other potential confounds as well. To limit the influence of context reinstatement, which refers to enhanced memory for items when the conditions at study match the conditions at test [73–75], we presented the words in different modalities in the study phase (i.e. verbal) versus the test phase (i.e. written). Additionally, we tested participants' item and source memory with a recognition test, which requires individuals only to choose the correct response among options rather than needing to retrieve the correct response from memory, as they would with a recall test, thereby minimizing TSD effects on retrieval processes as a possible confound.

To separate binding from item encoding processes, we assessed source memory performance only for those studied items that were correctly recognized. Because the word and speaker were processed as one during study in this task, recognized items indicated that both the item and therefore voice that produced it were attended, but the source memory test uniquely assessed whether the word was bound to the voice that produced it. We also statistically controlled for the potential effect of vigilant attention deficits on both item and source memory performance to account for the possibility that items and/or sources may not be encoded due to attentional lapsing associated with TSD [76]. With this rigorous method, our objective was to provide clarity to the literature on the effects of learning during TSD on source memory—and begin to bridge the gap in the literature on the role of binding in TSD-related deficits on tasks that require making associations in long-term memory.

Methods

Participants

$N = 68$ healthy adults (35 female, 33 male) aged 22–40 years (mean ± SD: 26.8 ± 4.8 years) completed a 4-day in-laboratory study. They were randomly assigned to either a total sleep deprivation (TSD) condition ($n = 38$) or a well-rested control (WRC) condition ($n = 30$). Data from one additional participant assigned to the WRC condition were excluded from analyses because, during baseline, this person did not understand the item and source memory task or did not comply with task instructions (giving the same response to every test item).

Participants were screened to be physically and psychologically healthy by means of physical exam, history, blood chemistry, and questionnaires. They were free of drugs (except for oral contraceptives) and were not currently on medical treatment or pregnant. They reported good habitual sleep of between 6 and 10 hours per night, with regular bedtimes and habitually getting up between 06:00 and 09:00 hours. They did not have any sleep or circadian disorders, as assessed by baseline polysomnography and questionnaires. They had not traveled across time zones within 1 month and had not engaged in shift work within 3 months of entering the study. All participants had normal or corrected-to-normal vision and hearing and were native English speakers.

For a week before and throughout the duration of the study, participants abstained from caffeine, tobacco, alcohol, and drugs. Adherence was confirmed by breathalyzer and urine drug screen during a screening session and upon admission to the laboratory. Participants were asked to maintain their habitual sleep pattern and refrain from napping in the week leading up to the study. Pre-study sleep schedules were verified by means of wrist-worn actigraphy, sleep diary, and called-in sleep and wake times.

The study was approved by the Washington State University's Institutional Review Board. Participants gave written, informed consent to the study procedures.

Experimental design

Participants were inside the laboratory for four consecutive days (three nights). They entered the laboratory in the late afternoon of day 1 and had a baseline nighttime sleep opportunity. On day 2, they were informed of their condition assignment. Participants in the WRC condition were given another nighttime sleep opportunity, whereas those in the TSD condition were kept awake for 38 hours continuously. All participants had a nighttime (recovery) sleep opportunity on day 3 in order to recuperate before leaving the laboratory in the morning of day 4. All sleep opportunities were 10 hours long (time in bed: 22:00–08:00 hours), and sleep was recorded polysomnographically. Total sleep time during baseline sleep was more than 8 hours in both groups, with a small but significant difference between the groups ($t_{46.6} = 2.44$, $p = 0.018$) such that the TSD group had a little more baseline sleep ($M = 534.4$ minutes, $SD = 41.1$) than the WRC group ($M = 502.9$ minutes, $SD = 64.9$).

The affective item and source memory (AISM) was administered twice—once at 14:50 hours on day 2 (baseline), when all participants were rested (session 1), and once 24 hours later on day 3, when TSD participants had been awake for nearly 31 hours and WRC participants were again rested (session 2). Two equivalent versions of the 15-minute task were used, and the order in which participants completed the two versions of the task was randomized and counterbalanced across participants. The number of participants who completed the task in each order was approximately equivalent within conditions (order 1 vs. 2: *n* = 20 vs. *n* = 18 in the TSD condition; *n* = 14 vs. *n* = 16 in the WRC condition).

To evaluate participants' vigilant attention and self-reported affect over the course of the study, a 10-minute psychomotor vigilance test (PVT) [76] and the Positive and Negative Affect Schedule (PANAS) [77] were administered at 2–4 hours intervals during scheduled wakefulness. These tasks served to verify the effectiveness of the TSD manipulation and to ascertain that participants' subjective affect did not influence their processing of affectively valenced stimuli [78].

For the entirety of the laboratory study, participants remained inside the sleep research facility of the Sleep and Performance Research Center at Washington State University Health Sciences Spokane, with fixed ambient temperature (21 ± 1°C) and light levels (<100 lux during scheduled wakefulness, lights off during sleep). Up to four participants participated in the study at the same time, each having their own room for sleep and performance testing. Meals were provided at 4 hours intervals during scheduled wakefulness. Between performance testing and meals, participants were only allowed to engage in non-vigorous activities. They were not allowed to use any laptops, tablets or cell phones, watch live television, listen to live radio, or receive visitors. Participants' compliance with sleep and behavioral conditions was monitored continuously by trained research assistants.

AISM task

The AISM task consisted of a study phase and a test phase. During the study phase, participants were presented with 68 words, at a rate of one word per second with a 1-second inter-trial interval, spoken by either a male or female speaker using a neutral tone of voice. The words were taken from the Affective Norms for English Words (ANEW) list [79]. Sixty of the 68 words served as target words. These were equally divided into categories by affective valence (how negative or positive words are), with 20 negative, 20 neutral, and 20 positive words. The target words were also equally divided over two sources (speakers), with 10 spoken by a male speaker and 10 spoken by a female speaker in each affective valence category. All target words were presented twice during the study phase, in randomized order, with both instances from the same speaker. Of the remaining eight words, four were presented before and four were presented after the target words to minimize serial position effects on later memory performance (as items presented at the beginning and end of lists tend to be better recognized than those in the middle of the list due to greater processing time and continued presence in working memory, respectively [80, 81]).

Immediately following the study phase, participants were tested for recognition of the target words. The test phase consisted of the 60 target words with the addition of 60 lures. The lures were negative, neutral, and positive words (20 each) also taken from the ANEW list. Test words were presented visually, in random order, with a 1-second interval between test trials. On each trial, participants indicated whether they recognized a word as having been heard before ("old") or not ("new") using the left and right mouse buttons. If they indicated that the word had been heard before, they then used the left and right mouse buttons to indicate if the word had been spoken by a male or female voice. Afterward, they had 10 seconds to rate their confidence in their voice choice on a 4-point Likert-type scale (1 = "very sure" to 4 = "just guessing"). There were no time limits to participants' responses.

The AISM task was programmed using E-Prime (version 2.0) [82]. In each iteration of the task, participants first completed a short practice block of the study and test phases. The practice study block consisted of six words, and the practice test block consisted of six target words and six lures. Participants were informed before and after the practice phase that they would be tested on whether the word was heard during study and if so, whether the speaker was male or female.

We compared the words used in the two versions of the AISM task on their affective valence, arousal (how calming or exciting words are), and frequency of use in US English. For both versions, negative and positive words had higher arousal ratings than neutral ones (version 1: t_{78} = 6.93, $p < 0.001$, and t_{78} = 8.55, $p < 0.001$; version 2: t_{78} = 8.91, $p < 0.001$, and t_{78} = 6.41, $p < 0.001$), and the arousal ratings did not differ significantly between negative and positive words in either version (t_{78} = 1.39, $p = 0.168$ and t_{78} = 1.89, $p = 0.062$, respectively). All words were 3–11 characters in length (mean ± SD: 5.75 ± 1.56). Their frequency of use, checked using the SUBTLEXus database [83], ranged from 2.47 to 5.89 (mean ± SD: 4.20 ± 0.62) on the Zipf scale [84], which expresses word frequency on a logarithmic scale from 1 (0.01 per million words) to 7 (10 000 per million words). Affective valence, arousal, and frequency did not differ significantly between the target words and lures (t_{238} = 0.24, $p = 0.81$; t_{238} = –0.24, $p = 0.81$; t_{238} = 1.17, $p = 0.24$) or between the two versions of the task (t_{238} = –0.36, $p = 0.72$; t_{238} = –1.50, $p = 0.14$; t_{238} = –0.35, $p = 0.72$).

Psychomotor vigilance test

The PVT measures the ability to sustain vigilant attention [76]. During this 10-minute task, participants were required to respond as quickly as possible, by pressing a button, to a visual stimulus presented at random 2- to 10-second intervals. Performance was measured by the number of lapses of attention, defined as reaction times > 500 ms. The PVT is highly sensitive to vigilant attention deficits caused by TSD [85, 86].

Positive and negative affect schedule

The PANAS is a 20-item self-report measure designed to assess participants' affect for a given time frame [77], in this case at the present moment. Using a 5-point Likert-type scale (1 = "very slightly or not at all" to 5 = "extremely"), participants indicated the degree to which 10 positive emotions, such as "excited," and 10 negative items, such as "irritable," described their current affect. The PANAS produces separate scores for positive and negative affect, and shows good internal consistency and construct

validity [77, 87]. The PANAS has been used previously to capture changes in self-reported affect associated with TSD [88–90].

Statistical analyses

Item memory

For item memory, recognition results were expressed in terms of the discriminability index (d') [91]. This index assessed participants' ability to discriminate between the target and lure words. We calculated d' from the standardized difference between participants' proportion of hits and false alarms for each of the three valence categories.[1] Hits were classified as the proportion of old (previously studied) words correctly recognized as targets, and false alarms were classified as the proportion of new (lure) words incorrectly recognized as targets. We also computed the criterion index, c, to evaluate whether participants were biased toward responding that words were "old" or "new." This measure was calculated by multiplying –0.5 by the standardized sum of participants' proportion of hits and false alarms [92]. The d' and c values, as well as the proportions of hits and false alarms, were analyzed with a linear mixed-effects ANOVA model with fixed effects for condition (TSD or WRC), session (1 or 2), affective valence (negative, neutral, or positive), and their interactions, and a random intercept for participants.

Source memory

For source memory, participants were able to make eight possible responses: a binary source judgment (male or female voice) and a 4-point confidence judgment. This allowed us to calculate recognition accuracy from receiver operating characteristic (ROC) curves [93, 94], which plot the hit rate against the false alarm rate at each confidence rating [92] (using the female voice as the target). Specifically, the area under the ROC curve (*AUC*) provides a nonparametric measure of participants' discriminability, with *AUC* values closer to 1 representing better ability to discriminate between sources. *AUC* values were calculated using the ROC Toolbox for MATLAB [95], which fit each participant's source memory responses to an unequal-variance signal detection (UVSD) model [96]. One WRC participant's data failed to fit the model, leaving 67 participants for the source memory analyses. We also calculated participants' response criterion, c, using the same method as described for item memory. The *AUC* and c values, proportions of hits and false alarms, and confidence ratings for source recognition were analyzed using the same linear mixed-effects ANOVA model as for item memory.

PVT and PANAS

PVT lapses of attention and PANAS positive and negative affect data were analyzed with a linear mixed-effects ANOVA model with fixed effects for time point (13 levels), condition (TSD or WRC), and their interaction, and a random intercept for participants. Only time points shared between the two conditions were included, as WRC participants did not have equivalent time points to the nighttime test administrations during TSD.

Covariates

To control for any effects of attentional lapses or subjective affect on participants' item and source memory performance, the aforementioned linear mixed-effects ANOVAs for item d' and source *AUC* were repeated, with participants' PVT lapses, positive affect, and negative affect from the bout closest to the AISM task administration in each session (at 13:00 hours) included as time-varying covariates.

The fixed effects reported here used the Satterthwaite approximation for degrees of freedom, and pairwise comparisons were Bonferroni-adjusted when comparing within levels of valence or within levels of bout (reported p values are multiplied by the number of comparisons, i.e. 3 or 13). The order of the AISM task administrations was included as a covariate in all primary item and source memory analyses and not found to be statistically significant ($p \geq 0.147$) nor materially influencing the pattern of effects. Therefore, order was not included in the results presented below.

Results

There were no significant two- or three-way interactions between condition and valence for any of our item or source memory outcome measures. We conducted follow-up analyses split by session to investigate any potential condition by valence interactions. The results from these analyses did not provide any evidence that our TSD manipulation influenced how item valence affected memory for items or sources.[2] These interactions are therefore not discussed further.

Item memory

Figure 1 shows the results for item memory performance on the AISM task. Results from the mixed-effects ANOVAs are reported in Table 1.

The d' analyses revealed significant main effects of condition ($F_{1, 66} = 7.20$, $p = 0.009$), session ($F_{1, 330} = 53.04$, $p < 0.001$), and word valence ($F_{2, 330} = 21.50$, $p < 0.001$), as well as the important condition by session interaction ($F_{1, 330} = 84.00$, $p < 0.001$). Further examination of the condition by session interaction revealed that in session 1, when all participants were rested, TSD and WRC participants' performance did not differ significantly ($t_{88.5} = -0.92$, $p = 0.360$). However, TSD participants' performance deteriorated during session 2, when they were sleep-deprived. Their performance in session 2 was worse than that of their WRC counterparts ($t_{88.5} = 5.90$, $p < 0.001$) and worse than their own baseline performance in session 1 ($t_{330} = 12.38$, $p < 0.001$). Rested participants' performance did not differ significantly between sessions ($t_{330} = -1.26$, $p = 0.209$). The decline in TSD participants' d' scores in session 2 was driven by a decrease in hits ($t_{330} = 8.12$, $p < 0.001$) and an increase in false alarms

---

[1]   Before calculating d', we converted proportions of hits and false alarms of 0 and 1 to 0.5/m and 1–0.5/m, respectively, to avoid undefined values [92]; m represents the number of trials for each condition, i.e. 20.

[2]   The condition by valence interaction was not significant for any item memory measures in session 1 or session 2 ($p \geq 0.420$ and $p \geq 0.160$, respectively) or any source memory measures in session 1 or session 2 ($p \geq 0.115$ and $p \geq 0.136$, respectively) save for source hits in session 1 ($F_{2, 130} = 4.50$, $p = 0.013$). However, pairwise comparisons indicated the only significant difference in session 1, when all participants were rested, was that WRC participants had significantly more source hits for positive words ($p = 0.048$) and neutral words ($p = 0.008$) than negative words.
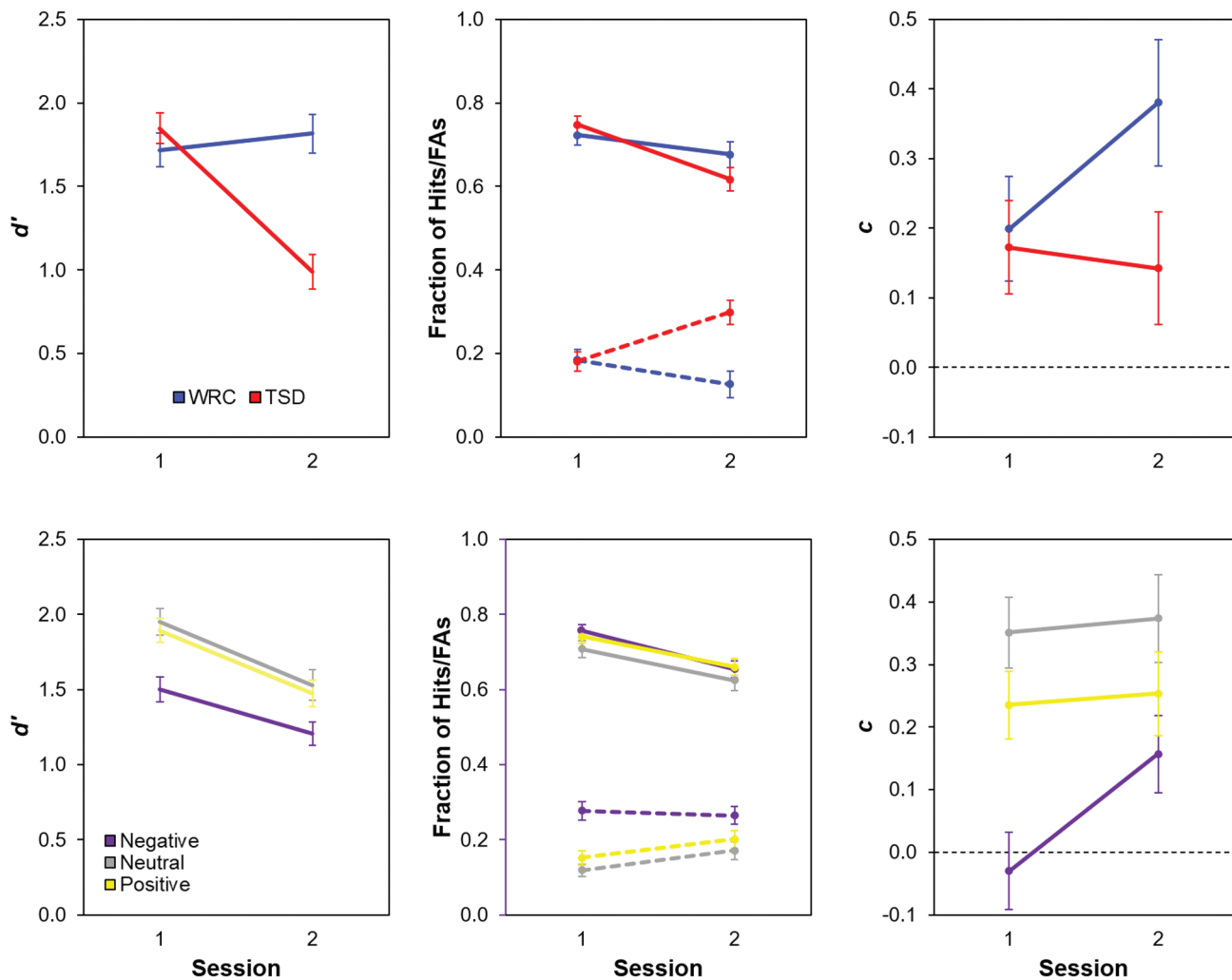
Figure 1. Effects of sleep deprivation on *item memory* in the AISM task. The top panels show means (± SE) for discriminability (*d'*), hits (solid lines) and false alarms (FAs, dashed lines) as a proportion of total trials, and criterion (*c*) values across sessions 1 (baseline) and 2 as a function of study condition (TSD or WRC), collapsed across affective valence. The bottom panels show means (± SE) for *d'*, hits and FAs, and *c* across sessions as a function of affective valence, collapsed across condition.

(t_{330} = −8.00, *p* < 0.001) compared to their rested performance. Although TSD participants showed an increase in false alarms during session 2, this increase did not reflect a significant shift in response criterion, *c* (t_{330} = 0.76, *p* = 0.449). WRC participants showed a decrease in false alarms during session 2 (t_{330} = 3.54, *p* = 0.001) and became more likely to respond that items were "new" (t_{330} = −4.08, *p* < 0.001). Taken together, these results show that TSD impaired participants' ability to discriminate between studied and non-studied words.

Participants overall had better *d'* scores for positive and neutral words compared to negative words (t_{330} = −5.19, *p* < 0.001, and t_{330} = −6.07, *p* < 0.001, respectively), while performance did not differ significantly between positive and neutral words (t_{330} = −0.87, *p* > 0.99). Although participants made more hits on negative and positive words than neutral words (t_{330} = 2.65, *p* = 0.026, and t_{330} = 2.31, *p* = 0.065), the poorer discriminability on negative words was driven by significantly greater false alarms to negative words than both positive words (t_{330} = 6.99, *p* < 0.001) and neutral words (t_{330} = 9.31, *p* < 0.001). Additionally, despite our intention for words spoken by the male and female speaker to

be recognized at similar rates, participants had better item recognition for negative words spoken by the female speaker.[3]

Response criterion was greater than zero in the direction to respond "new" for positive (t_{67} = 4.33, *p* < 0.001) and neutral words in session 1 (t_{67} = 6.27, *p* < 0.001) and for negative (t_{67} = 2.17, *p* = 0.033), positive (t_{67} = 3.56, *p* < 0.001), and neutral words in session 2 (t_{67} = 5.17, *p* < 0.001). Response criterion in session 1 was smaller for negative words than positive words (t_{330} = −5.15, *p* < 0.001) and neutral words (t_{330} = −7.40, *p* < 0.001), and response criterion for negative words increased from session 1 to session 2 (t_{330} = −3.63, *p* < 0.001). Valence effects did not differ significantly

---

[3] A mixed-effects ANOVA on the ability to correctly identify old items with fixed effects for condition, session, valence, speaker, and their interactions, and a random intercept for participants, revealed a significant effect of speaker ($F_{1, 726}$ = 8.28, *p* = 0.004) and a significant speaker by valence interaction ($F_{2, 726}$ = 4.66, *p* = 0.010), such that recognition was significantly greater for negative words spoken by the female rather than male speaker (t_{726} = 4.15, *p* < 0.001). Speaker did not significantly interact with condition ($F_{1, 726}$ = 3.02, *p* = 0.082) or session ($F_{1, 726}$ = 0.02, *p* = 0.878).

by condition, as seen in our lack of condition by valence and condition by valence by session interactions (Table 1). Thus, regardless of study condition, participants were worse at discriminating whether negative words had been presented during study.

Source memory

Figure 2 shows the results for source memory performance on the AISM task. Results from the mixed-effects ANOVAs are reported in Table 2. Note that to correctly recognize sources, participants must have first correctly recognized that an item was presented during the study phase of the task. Because participants did not have perfect item memory, only 69.0% of trials could be used meaningfully for analysis of source memory (subsets ranged from 68.2% to 69.9% by condition, from 64.3% to 73.7% by session, and from 66.5% to 70.6% by valence). Across those trials, the overall mean of participants' accuracy in identifying the source was 71.6% (subsets ranged from 68.4% to 75.5% by condition, from 70.0% to 73.0% by session, and from 68.9% to 73.5% by valence). Separately, the confidence ratings indicated that participants had higher confidence in correct source responses, showing that participants used the confidence ratings as intended.[4]

For participants' AUC, our nonparametric discriminability measure, the main effects of condition ($F_{1, 65} = 4.54$, $p = 0.037$) and session ($F_{1, 325} = 5.36$, $p = 0.021$) were significant, as were the interactions of condition by session ($F_{1, 325} = 16.98$, $p < 0.001$) and valence by session ($F_{2, 325} = 3.49$, $p = 0.032$). Exploring the condition by session interaction revealed that in session 1, the TSD and WRC groups did not differ significantly in their memory for sources ($t_{92.5} = 0.27$, $p = 0.786$). TSD participants' performance worsened during session 2, such that TSD participants' source performance in session 2 was significantly lower than that in the WRC condition ($t_{92.5} = 3.62$, $p = 0.001$) and their own baseline performance in session 1 ($t_{325} = 4.81$, $p < 0.001$). Much like the item discriminability results, WRC participants' discriminability for sources did not differ significantly between sessions ($t_{325} = -1.22$, $p = 0.225$).

To explain TSD participants' change in source memory AUC during session 2, we looked at participants' hits, false alarms, and confidence ratings. TSD participants made fewer source hits ($t_{325} = 3.46$, $p < 0.001$) and more source false alarms ($t_{325} = -3.80$, $p < 0.001$) in session 2 compared to their own baseline performance. In addition, TSD participants' confidence in their responses decreased from session 1 to session 2 ($t_{325} = -4.22$, $p < 0.001$). However, the two conditions did not differ significantly in the proportion of cases in which they endorsed the "Just Guessing" confidence level in session 2 ($M_{WRC} = 0.079$, $M_{TSD} = 0.096$; $\chi^2_1 = 2.44$, $p = 0.119$), suggesting that the TSD group's decrease in confidence was not simply due to an increase in item memory guesses. Furthermore, the increase in false alarms and decrease in confidence during TSD were not accompanied by a significant change in c for condition, session, or their interaction. Thus, sleep-deprived participants were not biased to respond that most

4   A mixed-effects ANOVA on participants' confidence ratings with fixed effects for condition, session, valence, source accuracy, and their interactions, and a random intercept for participants indicated that participants were more confident in correct than incorrect source responses ($F_{1, 5574.66} = 338.28$, $p < 0.001$). The condition by source accuracy interaction approached significance ($F_{1, 5514.03} = 3.36$, $p = 0.067$), but an examination of this interaction indicated that both groups were less confident in incorrect responses (both $p < 0.001$).

Table 1. ANOVA table of the mixed-effects model results for each measure of *item memory* performance

Measure	Effect	df	F	p
	Condition	1, 66	7.20	**0.009**
	Session	1, 330	53.04	**<0.001**
	Valence	2, 330	21.50	**<0.001**
d'	Condition × Session	1, 330	84.00	**<0.001**
	Condition × Valence	2, 330	0.73	0.481
	Valence × Session	2, 330	0.64	0.527
	Condition × Valence × Session	2, 330	1.09	0.337
	Condition	1, 66	1.58	0.213
	Session	1, 330	6.46	**0.011**
	Valence	2, 330	34.19	**<0.001**
c	Condition × Session	1, 330	12.60	**<0.001**
	Condition × Valence	2, 330	0.38	0.684
	Valence × Session	2, 330	3.50	**0.031**
	Condition × Valence × Session	2, 330	0.53	0.590
	Condition	1, 66	0.26	0.615
	Session	1, 330	53.62	**<0.001**
	Valence	2, 330	4.15	**0.017**
Hits	Condition × Session	1, 330	11.97	**<0.001**
	Condition × Valence	2, 330	0.50	0.607
	Valence × Session	2, 330	0.30	0.738
	Condition × Valence × Session	2, 330	1.51	0.222
	Condition	1, 66	5.51	**0.022**
	Session	1, 330	7.12	**0.008**
	Valence	2, 330	46.98	**<0.001**
FAs	Condition × Session	1, 330	63.43	**<0.001**
	Condition × Valence	2, 330	0.37	0.694
	Valence × Session	2, 330	3.67	**0.027**
	Condition × Valence × Session	2, 330	0.56	0.571

df, degrees of freedom (Satterthwaite approximation for degrees-of-freedom); d', discriminability index; c, criterion index; FAs, false alarms. The p values of significant effects ($p < 0.05$) are marked bold.

sources were male or female. Rather, sleep deprivation worsened TSD participants' source discriminability by decreasing hits and response confidence and increasing false alarms.

Separately, the valence by session interaction indicated that participants, regardless of condition, had better AUC values in session 1 for sources of positive and neutral items than for negative items ($t_{325} = 2.85$, $p = 0.014$, and $t_{325} = 3.01$, $p = 0.008$, respectively). Discriminability for sources of positive and neutral items did not significantly differ in session 1 ($t_{325} = -0.16$, $p > 0.99$). In session 2, AUC did not significantly differ by valence (all $p > 0.99$), and the lack of differences in session 2 was driven in part by greater source hits for negative than neutral words ($t_{325} = 2.80$, $p = 0.016$). Participants had more source false alarms in both sessions for negative words than for positive ($t_{325} = 2.73$, $p = 0.020$) and neutral words ($t_{325} = 4.41$, $p < 0.001$). They were also less confident in their source decisions for negative than positive words ($t_{325} = 2.93$, $p = 0.011$), although confidence for neutral words did not differ significantly from that for negative or positive words ($t_{325} = 2.08$, $p = 0.116$, and $t_{325} = -0.85$, $p > 0.99$, respectively). Response criterion only differed significantly from zero in session 2, when participants showed an increased tendency to assign negative words to the female speaker ($t_{66} = -3.05$, $p = 0.003$) and an increased tendency to assign neutral words to the male speaker ($t_{66} = 2.50$, $p = 0.015$). Thus, both TSD and WRC participants' poorer source discriminability on negative words in session 1 could only be attributed to their greater source false alarms and decreased confidence in source decisions for negative words.

Figure 2. Effects of sleep deprivation on *source memory* in the AISM task. The top panels show means (± SE) for nonparametric discriminability (*AUC*), hits (solid lines) and false alarms (FAs, dashed lines) as proportion of trials with correctly recognized items, criterion (*c*) values, and confidence ratings across sessions 1 (baseline) and 2 as a function of study condition (TSD or WRC), collapsed across affective valence. The bottom panels show means (±SE) for *AUC*, hits and FAs, *c*, and confidence across sessions as a function of affective valence, collapsed across condition. Note that confidence is plotted in reverse for ease of interpretation (lower values indicate greater confidence).

Vigilant attention and positive and negative affect

Figure 3 shows the performance of the TSD and WRC groups on the PVT and PANAS over the course of the experiment, and Table 3 shows the results from the mixed-effects ANOVAs.

For the PVT, there were significant main effects of bout ($F_{12, 792}$ = 20.98, p < 0.001) and condition ($F_{1, 66}$ = 30.06, p < 0.001) and a significant bout by condition interaction ($F_{12, 792}$ = 23.41, p < 0.001). As expected, TSD participants had significantly more lapses than WRC participants in each of the sleep-deprived bouts (all p ≤ 0.002), indicating that the TSD manipulation was successful. For positive affect on the PANAS, there was a significant main effect of bout ($F_{12, 792}$ = 41.86, p < 0.001) and a significant bout by condition interaction ($F_{12, 792}$ = 10.53, p < 0.001), such that TSD participants reported lower positive affect than WRC participants during the first bout in the morning following their night spent awake (p = 0.001 for the first bout, all other p ≥ 0.112). Finally, for negative affect, the main effect of bout ($F_{12, 792}$ = 4.68, p < 0.001) and the bout by condition interaction were significant ($F_{12, 792}$ = 1.91, p = 0.030), but TSD participants did not significantly differ from WRC participants during any of the sleep-deprived bouts (all p ≥ 0.485). Thus, sleep-deprived participants' change in affect during sleep deprivation was characterized mainly by a loss of positive affect rather than an increase in negative affect.

Covariate analyses

When covariates were added to the item *d*′ and source *AUC* ANOVAs, the number of lapses during the PVT bout closest to the AISM task was a significant covariate for item memory ($F_{1, 383.87}$ = 4.63, p = 0.032), but not for source memory ($F_{1, 374.89}$ = 1.10, p = 0.294). For session 2 in the TSD condition, the correlation with PVT lapses was significant for item *d*′ (r_{35} = −0.37, p = 0.026) but not for source *AUC* (r_{35} = −0.04, p = 0.837); these correlations were significantly different (z = −2.48, p = 0.013) [97]. Neither positive affect (all p ≥ 0.123) nor negative affect (all p ≥ 0.072) were significant covariates for item *d*′ or source *AUC*. Adding the covariates did not substantively change the reported results for either item or source memory.

Discussion

Our study was motivated by the possibility that sleep-deprived people may have poorer ability to bind items with their contexts. We examined the effects of TSD on contextual binding using a well-established source memory paradigm, the AISM. Our findings support three important conclusions about item and source memory under TSD: (1) encoding items and binding them in memory to their sources are both impaired by TSD; (2) even when items are recognized accurately, recognition of their association to sources is reduced by TSD, reflecting that item and source memory are related but separate processes; and (3) the affective valence of individual stimuli influences how items are bound with sources, but this process does not appear to be moderated by TSD.

Our work replicates previous findings that encoding new episodic memories for individual items is decreased during TSD

Table 2. ANOVA table of the mixed-effects model results for each measure of *source memory* performance

Measure	Effect	df	F	p
AUC	Condition	1, 65	4.54	**0.037**
	Session	1, 325	5.36	**0.021**
	Valence	2, 325	2.74	0.066
	Condition × Session	1, 325	16.98	**<0.001**
	Condition × Valence	2, 325	1.56	0.212
	Valence × Session	2, 325	3.49	**0.032**
	Condition × Valence × Session	2, 325	0.46	0.632
c	Condition	1, 65	0.16	0.691
	Session	1, 325	0.81	0.370
	Valence	2, 325	7.80	**<0.001**
	Condition × Session	1, 325	0.10	0.756
	Condition × Valence	2, 325	0.71	0.491
	Valence × Session	2, 325	3.42	**0.034**
	Condition × Valence × Session	2, 325	1.92	0.148
Hits	Condition	1, 65	3.54	0.064
	Session	1, 325	1.25	0.264
	Valence	2, 325	1.18	0.307
	Condition × Session	1, 325	12.33	**<0.001**
	Condition × Valence	2, 325	2.23	0.109
	Valence × Session	2, 325	4.86	**0.008**
	Condition × Valence × Session	2, 325	2.35	0.097
FAs	Condition	1, 65	5.03	**0.028**
	Session	1, 325	4.71	**0.031**
	Valence	2, 325	9.89	**<0.001**
	Condition × Session	1, 325	8.48	**0.004**
	Condition × Valence	2, 325	0.43	0.653
	Valence × Session	2, 325	1.18	0.310
	Condition × Valence × Session	2, 325	0.46	0.632
Confidence	Condition	1, 65	0.17	0.684
	Session	1, 325	2.34	0.127
	Valence	2, 325	4.54	**0.011**
	Condition × Session	1, 325	16.95	**<0.001**
	Condition × Valence	2, 325	0.53	0.589
	Valence × Session	2, 325	0.86	0.426
	Condition × Valence × Session	2, 325	1.06	0.347

df, degrees of freedom (Satterthwaite approximation for degrees-of-freedom); AUC, area under the curve (nonparametric discriminability index); c, criterion index; FAs, false alarms. The p values of significant effects (p < 0.05) are marked bold.

[7, 55, 98], but importantly shows that *even if item representations are intact in memory, memory for the item's specific source—that is, the association between the item and its context—is impaired by TSD*. This is in line with previous work indicating that brain regions implicated in both item and source memory, the hippocampus and PFC, are disrupted by TSD [17, 27, 38]. The findings provide strong support to a small but growing body of research that has found similar deficits under sleep loss on tasks that required participants to create associations between stimuli in long-term memory [13–15, 98]. Our finding of impaired item and source memory was due to a pattern of fewer hits and more false alarms for both items and sources compared to rested controls. The increased false alarms during sleep deprivation were not explained by shifts in response criterion for TSD participants, indicating that the effects of sleep deprivation on source memory could not be explained as an epiphenomenon of changes in response bias. Also, it is possible that guessing on the item memory test would have contributed substantially to

the degradation of source memory performance during TSD, but if that were the case then there should have been a concomitant increase in "Just Guessing" ratings for the confidence level of source memory responses, as participants without strong item memory would be highly unlikely to nonetheless possess strong memory for that item's source. However, ratings for the confidence level of source memory as "Just Guessing" in session 2 did not differ significantly between the two conditions, making a methodological confound from item memory guesses an unlikely explanation for the observed deficits in source memory beyond those seen in item memory.

Indeed, despite both item and source memory being degraded by TSD, encoding an item (or a source) is a distinct process from binding that item to its source. Research in populations with frontal or temporal lobe deficits and research using neuroimaging techniques have repeatedly found that item and source memory are dissociable processes that depend on different PFC and MTL structures [24–27, 30–32]. This dissociation between item and source memory is further evidenced in our study by the fact that PVT lapses of attention were a significant covariate for item memory but not for source memory, and the correlation with PVT lapses during TSD was significantly greater for item memory than for source memory. This dissociation suggests that problems with item and source memory during TSD are associated with deficits in distinct underlying processes.

Our findings also revealed that although item valence affected source memory, we did not observe any moderation of the valence effects by TSD. Whether sleep-deprived or rested, participants were less able to recognize the source of negative words than neutral words, particularly during the first administration of the AISM. Instead of emotional items showing enhanced memory and consequently bolstering memory for their sources [99], our results suggest that negative items are associated with both more hits and more false alarms, and consequently with more source false alarms and decreased confidence in source decisions.

These results are consistent with the idea that negative items result in greater attentional capture [58, 100]. We would expect increased attention toward negative items to cause people to not only have better memory for those items and their integrally processed sources, but also be more likely to conflate the increased saliency of those items with actual memory, resulting in increased false alarms for items and sources. Although we did not see an overall tendency to respond that negative items had been studied, also known as emotion-induced recognition bias [101–104], we did see a shift in response criterion for the source of negative items during the second administration of the AISM, such that negative words were more likely to be attributed to the female speaker. This tendency to associate negative words with the female speaker echoes the finding that negative words spoken by the female speaker were more frequently recognized correctly overall. As female voices are more perceptually salient than male voices [105], our findings suggest that participants conflated the greater attentional demands of negative items with that of the female voice, impairing source discrimination by increasing false alarms.

We note that the proposed effect of valence on attention and subsequent encoding and binding does not explain why we failed to observe any interaction of condition and valence on any of our item or source memory measures. Despite well-recognized effects of TSD on attention [76, 106, 107] and previous
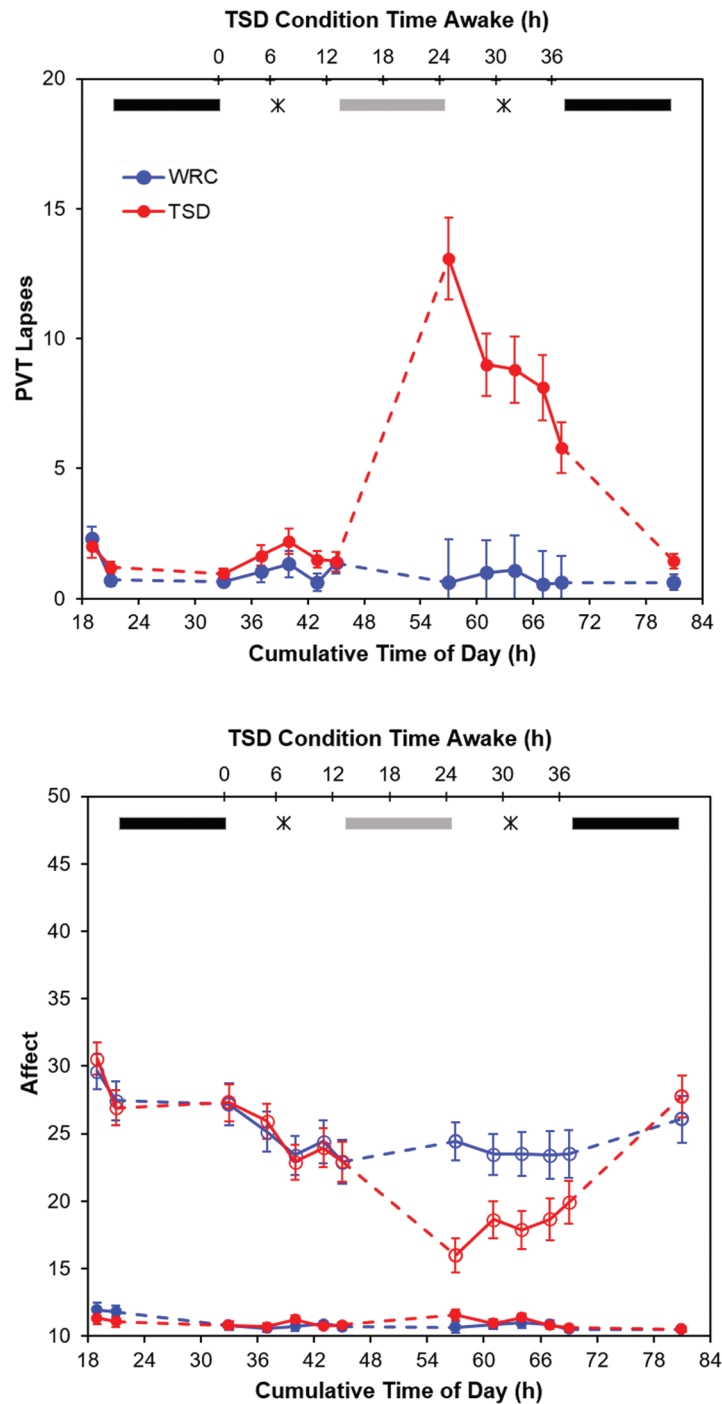
Figure 3. Number of attentional lapses on the psychomotor vigilance test (PVT) and positive and negative affect on the PANAS across the laboratory study. The PVT results are shown in the top panel; the positive affect (open circles) and negative affect (solid circles) are shown in the bottom panel. Each graph shows the means on each measure for each group (TSD or WRC); error bars represent standard error of the mean. On the top of each graph, black bars indicate scheduled sleep, gray bars indicate sleep for the WRC condition only, and the stars represent the two administrations of the AISM.

work indicating that memory for negative items is preserved under TSD [15, 55], we did not find evidence that TSD changed the pattern of how valenced items were encoded or bound with their sources. In two previous studies where an interaction between sleep condition and valence was observed on item memory [15, 55], memory was tested after two days of recovery sleep, whereas we tested memory immediately following study. It may be the case that the preferentially preserved encoding

of negative items during TSD is only apparent at longer delays, but including opportunities for sleep between encoding and test increases the chances that the effects of item valence are confounded by the effects of sleep-related consolidation [108, 109].

While our focus in this study was on how TSD affects the initial learning aspect of memory, our results regarding contextual memory are similar to those of previous work investigating the effects of insufficient sleep during consolidation. That is, like

Table 3. ANOVA table of the mixed-effects model results for lapses on the PVT and positive and negative affect

Measure	Effect	df	F	p
	Condition	1, 66	30.06	**<0.001**
Lapses	Bout	12, 792	20.98	**<0.001**
	Condition × Bout	12, 792	23.41	**<0.001**
	Condition	1, 66	1.06	0.307
PA	Bout	12, 792	41.86	**<0.001**
	Condition × Bout	12, 792	10.53	**<0.001**
	Condition	1, 66	0.03	0.873
NA	Bout	12, 792	4.68	**<0.001**
	Condition × Bout	12, 792	1.91	**0.030**

df, degrees of freedom (Satterthwaite approximation for degrees-of-freedom); d', discriminability index; PA, positive affect; NA, negative affect. The p values of significant effects ($p < 0.05$) are marked bold.

these previous studies [3–6], we also found that sleep loss degrades contextual memory performance. Yet it is unlikely that these comparable findings are due to TSD effects on the same underlying memory processes. The consolidation processes in question have been studied through testing after a sleep period (or matching non-sleep control period) and confirmed to be sleep-dependent [110, 111], whereas in our study testing occurred immediately following learning with no intervening sleep. With sleep-based consolidation thus ruled out in both the TSD and WRC conditions, our study design limited any meaningful consolidation from influencing the observed outcomes [112].

In principle, it is possible that our results reflect effects of TSD on encoding and/or retrieval, as TSD participants completed both study and test phases of the AISM task while sleep-deprived. However, any TSD-related deficits in retrieval were minimized in our study by use of a recognition test (rather than a recall test), which provided participants with the appropriate cues (male or female) and required them only to retrieve the association between the word and the speaker. Moreover, the few studies that have investigated the effects of a night of TSD before retrieving previously learned information, including on an associative memory task that would require binding [10, 11], have not found evidence of TSD-related deficits on retrieval. As such, consistent with the potential for sleep deprivation to affect encoding processes [7–9], the observed impairments to source memory are likely a TSD-induced issue not with retrieval but with initial encoding and binding.

Even so, failures in encoding or binding do not necessarily mean that the information is not available or has been lost forever. It is possible that TSD participants simply had weaker binding of items and sources, but if allowed to sleep and consolidate the information they might show some recovery of the information [113]. In considering how any problems with initial encoding or binding might manifest in later tests of memory, it would be important to bear in mind that the memory phases of encoding, consolidation, and retrieval are interdependent. The connections between memory stages can influence how well one remembers in any given scenario; for instance, matching the conditions at retrieval to those at encoding, such as one's environmental context or internal state [114, 115], can facilitate the retrieval of information from memory (i.e. encoding specificity and state-dependent learning [73, 116]). Because memory is a constructive and reconstructive process, presumably stable memories can become labile again, and previous knowledge or beliefs, like schemas, can distort newly formed memories [117, 118]. In addition, hippocampal binding can occur over shorter and longer time scales and in both short-term and long-term memory [16, 48, 119]. Binding processes can even take place post-encoding, when individuals are creating new representations in memory and long-term memories that have been reactivated by retrieval become malleable [117] and can be re-bound with new information [23, 118, 120, 121]. While we focused on binding very close in time to initial encoding, it remains to be investigated whether TSD similarly impacts binding at longer time scales or during post-encoding processes, such as reconsolidation.

Understanding how TSD affects source memory, which requires the binding of items to their contexts, is not only important in and of itself, but also serves as a critical first step in determining whether long-term binding is an underlying source of impairment during TSD on tasks that require people to form associations. Because we found that TSD impairs source memory, we encourage researchers to engage in functional task analyses to determine whether the TSD effects they observe may be downstream consequences of problems with binding. The relationship between binding and TSD may be more obvious on memory tasks that ask people to remember items in contexts or groups of items, but there are many tasks that assume that people are able to form associations during TSD, as noted earlier in the IGT example [22]. However, as mentioned previously, binding is critical not only for tasks that require long-term memory, but also for tasks that rely on short-term memory, such as change-detection tasks where participants must bind features like color or location to an object in order to respond correctly [122]. Further, binding is necessary for both basic perception, integrating features into a single, comprehensible object, and for behavior, linking stimuli to action plans [123, 124], although feature binding may rely on different neural circuits than binding in long-term memory [123, 125, 126]. Thus, even researchers not examining memory should be sensitive to the role binding may play in their tasks. Separating the effect of TSD on binding from its effects on other task components could provide clarity on which cognitive processes are in fact impacted by sleep loss. In this vein, we encourage work investigating whether the effects of TSD on complex cognitive tasks can be usefully organized by the tasks' need for participants to engage in binding.

In summary, the results from this study indicate that TSD disrupts the binding of items and their sources. Our findings highlight the impairing effect of sleep loss on memory systems and have critical implications for domains where accurate reporting of events and their contexts is vital, such as eyewitness testimony. The results also imply that some TSD deficits on tasks that require participants to make associations in order to perform well may be due to underlying problems with forming associations in memory. If true, employing methods to bolster memory for associations, such as unitization—that is, processing multiple items as a single unit (e.g. committing two words, *sweet* and *house*, to memory by processing them as a compound word, *sweethouse*) [127–129]—could lead to performance improvements during sleep loss on a wide range of tasks, including those relevant for real-world decision making.

Acknowledgments

Funding

Disclosure Statement

Financial disclosure: None declared.
Non-financial disclosure: None declared.

Data Availability

Deidentified participant data may be requested from the corresponding author.

References

1. Stickgold R. Parsing the role of sleep in memory processing. *Curr Opin Neurobiol*. 2013;**23**(5):847–853.
2. Paller KA, *et al*. Memory and sleep: how sleep cognition can change the waking mind for the better. *Annu Rev Psychol*. 2021;**72**:123–150.
3. Lewis PA, *et al*. The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia*. 2011;**49**(9):2619–2629.
4. van der Helm E, *et al*. Sleep-dependent facilitation of episodic memory details. *PLoS One*. 2011;**6**(11):e27421.
5. Mawdsley M, *et al*. The effect of sleep on item recognition and source memory recollection among shift-workers and permanent day-workers. *J Sleep Res*. 2014; **23** (5): 538–544.
6. Sopp MR, *et al*. Remembering specific features of emotional events across time: the role of REM sleep and prefrontal theta oscillations. *Cogn Affect Behav Neurosci*. 2017;**17**(6):1186–1209.
7. Drummond SP, *et al*. Altered brain response to verbal learning following sleep deprivation. *Nature*. 2000;**403**(6770):655–657.
8. Yoo SS, *et al*. A deficit in the ability to form new human memories without sleep. *Nat Neurosci*. 2007;**10**(3):385–392.
9. Poh JH, *et al*. Degradation of cortical representations during encoding following sleep deprivation. *Neuroimage*. 2017;**153**:131–138.
10. Schönauer M, *et al*. Evidence for two distinct sleep-related long-term memory consolidation processes. *Cortex*. 2015; **63**: 68–78.
11. Idzikowski C. Sleep and memory. *Br J Psychol*. 1984;**75** (Pt 4):439–449.
12. Johnson MK, *et al*. Source monitoring. *Psychol Bull*. 1993;**114**(1):3–28.
13. Harrison Y, *et al*. Sleep loss and temporal memory. *Q J Exp Psychol A*. 2000;**53**(1):271–279.
14. Kaida K, *et al*. Role of sleep for encoding of emotional memory. *Neurobiol Learn Mem*. 2015;**121**:72–79.
15. Tempesta D, *et al*. The effect of sleep deprivation on the encoding of contextual and non-contextual aspects of emotional memory. *Neurobiol Learn Mem*. 2016;**131**:9–17.
16. Zimmer HD, *et al*. Levels of binding: Types, mechanisms, and functions of binding in remembering. In: Zimmer HD, Mecklinger A, Lindenberger U, eds. *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience*. New York, NY: Oxford University Press; 2006: 3–22.
17. Harrison Y, *et al*. One night of sleep loss impairs innovative thinking and flexible decision making. *Organ Behav Hum Decis Process*. 1999;**78**(2):128–145.
18. Killgore WD, *et al*. Impaired decision making following 49 h of sleep deprivation. *J Sleep Res*. 2006;**15**(1):7–13.
19. McKenna BS, *et al*. The effects of one night of sleep deprivation on known-risk and ambiguous-risk decisions. *J Sleep Res*. 2007;**16**(3):245–252.
20. Whitney P, *et al*. Feedback blunting: total sleep deprivation impairs decision making that requires updating based on feedback. *Sleep*. 2015;**38**(5):745–754.
21. Honn KA, *et al*. Cognitive flexibility: a distinct element of performance impairment due to sleep deprivation. *Accid Anal Prev*. 2019;**126**:191–197.
22. Whitney P, *et al*. The role of source memory in gambling task decision making. *J Clin Exp Neuropsychol*. 2012;**34**(8):826–835.
23. Rubin RD, *et al*. The role of the hippocampus in flexible cognition and social behavior. *Front Hum Neurosci*. 2014;**8**:742.
24. Glisky EL, *et al*. Double dissociation between item and source memory. *Neuropsychology*. 1995; **9** (2): 229–235.
25. Mitchell KJ, *et al*. Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psychol Bull*. 2009;**135**(4):638–677.
26. Donaldson DI, *et al*. Remember the source: dissociating frontal and parietal contributions to episodic memory. *J Cogn Neurosci*. 2010;**22**(2):377–391.
27. Monge ZA, *et al*. Functional networks underlying item and source memory: shared and distinct network components and age-related differences. *Neurobiol Aging*. 2018;**69**:140–150.
28. Macklin CB, *et al*. The bizarreness effect: dissociation between item and source memory. *Memory*. 2005;**13**(7): 682–689.
29. Jurica PJ, *et al*. Monitoring item and source information: evidence for a negative generation effect in source memory. *Mem Cognit*. 1999;**27**(4):648–656.
30. Weis S, *et al*. Process dissociation between contextual retrieval and item recognition. *Neuroreport*. 2004;**15**(18):2729–2733.
31. Janowsky JS, *et al*. Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*. 1989;**27**(8):1043–1056.
32. Shimamura AP, *et al*. The relationship between fact and source memory: findings from amnesic patients and normal subjects. *Psychobiology*. 1991; **19**(1):1–10.
33. Mayes A, *et al*. Associative memory and the medial temporal lobes. *Trends Cogn Sci*. 2007;**11**(3):126–135.
34. Ranganath C. Binding items and contexts: the cognitive neuroscience of episodic memory. *Current Directions in Psychological Science*. 2010;**19**(3):131–137.
35. McCormick C, *et al*. Hippocampal-neocortical networks differ during encoding and retrieval of relational memory: functional and effective connectivity analyses. *Neuropsychologia*. 2010;**48**(11):3272–3281.
36. Wais PE. Hippocampal signals for strong memory when associative memory is available and when it is not. *Hippocampus*. 2011;**21**(1):9–21.
37. Van Der Werf YD, *et al*. Sleep benefits subsequent hippocampal functioning. *Nat Neurosci*. 2009;**12**(2):122–123.

38. Chai Y, *et al*. Two nights of recovery sleep restores hippocampal connectivity but not episodic memory after total sleep deprivation. *Sci Rep*. 2020;**10**(1):8774.

39. Brown MW, *et al*. Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nat Rev Neurosci*. 2001;**2**(1):51–61.

40. Davachi L, *et al*. Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci U S A*. 2003;**100**(4):2157–2162.

41. Davachi L. Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol*. 2006;**16**(6):693–700.

42. Atienza M, *et al*. Modulatory effects of emotion and sleep on recollection and familiarity. *J Sleep Res*. 2008;**17**(3):285–294.

43. Olsen RK, *et al*. The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front Hum Neurosci*. 2012;**6**:146.

44. Harrison Y, *et al*. Prefrontal neuropsychological effects of sleep deprivation in young adults–a model for healthy aging? *Sleep*. 2000;**23**(8):1067–1073.

45. Chuah YM, *et al*. The neural basis of interindividual variability in inhibitory efficiency after sleep deprivation. *J Neurosci*. 2006;**26**(27):7156–7162.

46. Chen J, *et al*. Sleep deprivation promotes habitual control over goal-directed control: behavioral and neuroimaging evidence. *J Neurosci*. 2017;**37**(49):11979–11992.

47. Woodcock EA, *et al*. The dorsal prefrontal and dorsal anterior cingulate cortices exert complementary network signatures during encoding and retrieval in associative memory. *Behav Brain Res*. 2015;**290**:152–160.

48. Rubin RD, *et al*. Dynamic hippocampal and prefrontal contributions to memory processes and representations blur the boundaries of traditional cognitive domains. *Brain Sciences*. 2017;**7**(7):82.

49. Senkfor AJ, *et al*. Who said what? An event-related potential investigation of source and item memory. *J Exp Psychol Learn Mem Cogn*. 1998;**24**(4):1005–1025.

50. Davidson PS, *et al*. Effects of emotion on item and source memory in young and older adults. *Cogn Affect Behav Neurosci*. 2006;**6**(4):306–322.

51. Keefe RS, *et al*. Source monitoring deficits in patients with schizophrenia; a multinomial modelling analysis. *Psychol Med*. 1999;**29**(4):903–914.

52. Yonelinas AP. The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *J Exp Psychol Learn Mem Cogn*. 1999;**25**(6):1415–1434.

53. Dodson CS, *et al*. Differential effects of cue dependency on item and source memory. *J Exp Psychol Learn Mem Cogn*. 2000;**26**(4):1023–1044.

54. Waters FA, *et al*. Memory for speech and voice identity in schizophrenia. *J Nerv Ment Dis*. 2009;**197**(12):887–891.

55. Walker MP, *et al*. Sleep, memory, and plasticity. *Annu Rev Psychol*. 2006;**57**:139–166.

56. Bradley MM, *et al*. Remembering pictures: pleasure and arousal in memory. *J Exp Psychol Learn Mem Cogn*. 1992;**18**(2):379–390.

57. Kensinger EA, *et al*. Memory enhancement for emotional words: are emotional words more vividly remembered than neutral words? *Mem Cognit*. 2003;**31**(8):1169–1180.

58. Talmi D, *et al*. The role of attention and relatedness in emotionally enhanced memory. *Emotion*. 2007;**7**(1):89–102.

59. Grider RC, *et al*. Discriminating between changes in bias and changes in accuracy for recognition memory of emotional stimuli. *Mem Cognit*. 2008;**36**(5):933–946.

60. Adelman JS, *et al*. Emotion and memory: a recognition advantage for positive and negative words independent of arousal. *Cognition*. 2013;**129**(3):530–535.

61. Doerksen S, *et al*. Source memory enhancement for emotional words. *Emotion*. 2001;**1**(1):5–11.

62. D'Argembeau A, *et al*. Influence of affective meaning on memory for contextual information. *Emotion*. 2004;**4**(2):173–188.

63. MacKay DG, *et al*. Emotion, memory, and attention in the taboo Stroop paradigm: an experimental analogue of flashbulb memories. *Psychol Sci*. 2005; **16** (1): 25–32.

64. Mather M, *et al*. Arousal-Enhanced Location Memory for Pictures. *J Mem Lang*. 2008;**58**(2):449–464.

65. Cook GI, *et al*. Source monitoring is not always enhanced for valenced material. *Mem Cognit*. 2007;**35**(2):222–230.

66. MacKenzie G, *et al*. Positive emotion can protect against source memory impairment. *Cogn Emot*. 2015;**29**(2):236–250.

67. Mao X, *et al*. Emotion impairs extrinsic source memory–An ERP study. *Biol Psychol*. 2015;**110**:182–189.

68. Earles JL, *et al*. Memory for positive, negative and neutral events in younger and older adults: does emotion influence binding in event memory? *Cogn Emot*. 2016;**30**(2):378–388.

69. Yoo S-S, *et al*. The human emotional brain without sleep — a prefrontal amygdala disconnect. *Curr Biol*. 2007;**17**(20):R877–R878.

70. Gujar N, *et al*. Sleep deprivation amplifies reactivity of brain reward networks, biasing the appraisal of positive emotional experiences. *J Neurosci*. 2011;**31**(12):4466–4474.

71. Pilcher JJ, *et al*. Sleep deprivation affects reactivity to positive but not negative stimuli. *J Psychosom Res*. 2015;**79**(6):657–662.

72. Dudai Y. The restless engram: consolidations never end. *Annu Rev Neurosci*. 2012;**35**:227–247.

73. Tulving E, *et al*. Encoding specificity and retrieval processes in episodic memory. *Psychol Rev*. 1973; **80** (5): 352–373.

74. Palmeri TJ, *et al*. Episodic encoding of voice attributes and recognition memory for spoken words. *J Exp Psychol Learn Mem Cogn*. 1993;**19**(2):309–328.

75. Campeanu S, *et al*. Voice congruency facilitates word recognition. *PLoS One*. 2013;**8**(3):e58778.

76. Lim J, *et al*. Sleep deprivation and vigilant attention. *Ann N Y Acad Sci*. 2008;**1129**:305–322.

77. Watson D, *et al*. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol*. 1988;**54**(6):1063–1070.

78. Bower GH. Mood and memory. *Am Psychol*. 1981;**36**(2):129–148.

79. Bradley MM, *et al*. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida; 1999.

80. Deese J. Serial organization in the recall of disconnected items. *Psychol Rep*. 1957; **3** (3): 577–582.

81. Jahnke JC. Primacy and recency effects in serial-position curves of immediate recall. *J Exp Psychol*. 1965;**70**:130–132.

82. Psychology Software Tools, Inc. E-Prime 2.0. Pittsburgh, PA: Psychology Software Tools, Inc.; 2012. https://www.pstnet.com.

83. Brysbaert M, *et al*. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav Res Methods*. 2009;**41**(4):977–990.

84. van Heuven WJ, *et al*. SUBTLEX-UK: a new and improved word frequency database for British English. *Q J Exp Psychol (Hove)*. 2014;**67**(6):1176–1190.

85. Doran SM, *et al*. Sustained attention performance during sleep deprivation: evidence of state instability. *Arch Ital Biol.* 2001;**139**(3):253–267.

86. Balkin TJ, *et al*. Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *J Sleep Res.* 2004;**13**(3):219–227.

87. Crawford JR, *et al*. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. *Br J Clin Psychol.* 2004;**43**(Pt 3):245–265.

88. Franzen PL, *et al*. Relationships between affect, vigilance, and sleepiness following sleep deprivation. *J Sleep Res.* 2008;**17**(1):34–41.

89. Riedy S, *et al*. Comparison of the SPANE and PANAS scales for measuring self-reported affect during total sleep deprivation. *Sleep-Wake Research in the Netherlands.* 2013;**24**: 102–105.

90. Schwarz J, *et al*. Mood impairment is stronger in young than in older adults after sleep deprivation. *J Sleep Res.* 2019;**28**(4):e12801.

91. Green DM, *et al*. *Signal Detection Theory and Psychophysics.* Oxford, UK: John Wiley; 1966.

92. Macmillan NA, *et al*. *Detection Theory: A User's Guide.* 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2005.

93. Slotnick SD, *et al*. An analysis of signal detection and threshold models of source memory. *J Exp Psychol Learn Mem Cogn.* 2000;**26**(6):1499–1517.

94. Qin J, *et al*. Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *J Exp Psychol Learn Mem Cogn.* 2001;**27**(4):1110–1115.

95. Koen JD, *et al*. The ROC Toolbox: a toolbox for analyzing receiver-operating characteristics derived from confidence ratings. *Behav Res Methods.* 2017;**49**(4):1399–1406.

96. Egan JP. *Signal Detection Theory and ROC-Analysis.* New York, NY: Academic Press; 1975.

97. Hittner JB, *et al*. A Monte Carlo evaluation of tests for comparing dependent correlations. *J Gen Psychol.* 2003;**130**(2):149–168.

98. Ratcliff R, *et al*. The effects of sleep deprivation on item and associative recognition memory. *J Exp Psychol Learn Mem Cogn.* 2018;**44**(2):193–208.

99. Mather M. Emotional arousal and memory binding: an object-based framework. *Perspect Psychol Sci.* 2007;**2** (1):33–52.

100. Phelps EA. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr Opin Neurobiol.* 2004;**14**(2):198–202.

101. Maratos EJ, *et al*. Recognition memory for emotionally negative and neutral words: an ERP study. *Neuropsychologia.* 2000;**38**(11):1452–1465.

102. Windmann S, *et al*. Electrophysiological correlates of emotion-induced recognition bias. *J Cogn Neurosci.* 2001;**13**(5):577–592.

103. McNeely HE, *et al*. ERP indices of emotionality and semantic cohesiveness during recognition judgments. *Psychophysiology.* 2004;**41**(1):117–129.

104. Dougal S, *et al*. "Remembering" emotional words is based on response bias, not recollection. *Psychon Bull Rev.* 2007;**14**(3):423–429.

105. Lattner S, *et al*. Voice perception: sex, pitch, and the right hemisphere. *Hum Brain Mapp.* 2005;**24**(1):11–20.

106. Lim J, *et al*. A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychol Bull.* 2010;**136**(3):375–389.

107. Hudson AN, *et al*. Sleep deprivation, vigilant attention, and brain function: a review. *Neuropsychopharmacology.* 2020;**45**(1):21–30.

108. Sterpenich V, *et al*. Sleep-related hippocampo-cortical interplay during emotional memory recollection. *PLoS Biol.* 2007;**5**(11):e282.

109. Payne JD, *et al*. Sleep's role in the consolidation of emotional episodic memories. *Current Directions in Psychological Science.* 2010;**19**(5):290–295.

110. Stickgold R, *et al*. Sleep and memory: the ongoing debate. *Sleep.* 2005;**28**(10):1225–1227.

111. Feld GB, *et al*. Sculpting memory during sleep: concurrent consolidation and forgetting. *Curr Opin Neurobiol.* 2017;**44**:20–27.

112. Frankland PW, *et al*. The organization of recent and remote memories. *Nat Rev Neurosci.* 2005;**6**(2):119–130.

113. Baena D, *et al*. Weakly encoded memories due to acute sleep restriction can be rescued after one night of recovery sleep. *Sci Rep.* 2020;**10**(1):1449.

114. Godden DR, *et al*. Context-dependent memory in two natural environments: on land and underwater. *Brit J Psychol.* 1975;**66**(3):325–331.

115. Eich E, *et al*. Mood dependent memory for internal versus external events. *J Exp Psychol.* 1989;**15**(3):443–455.

116. Goodwin DW, *et al*. Alcohol and recall: state-dependent effects in man. *Science.* 1969;**163**(3873):1358–1360.

117. Nader K. Memory traces unbound. *Trends Neurosci.* 2003;**26**(2):65–72.

118. Schacter DL. Constructive memory: past and future. *Dialogues Clin Neurosci.* 2012;**14**(1):7–18.

119. Yonelinas AP. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behav Brain Res.* 2013;**254**:34–44.

120. Bridge DJ, *et al*. Hippocampal binding of novel information with dominant memory traces can support both memory stability and change. *J Neurosci.* 2014;**34**(6):2203–2213.

121. Bridge DJ, *et al*. Active retrieval facilitates across-episode binding by modulating the content of memory. *Neuropsychologia.* 2014;**63**:154–164.

122. Wheeler ME, *et al*. Binding in short-term visual memory. *J Exp Psychol Gen.* 2002;**131**(1):48–64.

123. Treisman A. Feature binding, attention and object perception. *Philos Trans R Soc Lond B Biol Sci.* 1998;**353**(1373):1295–1306.

124. Hommel B. Event files: feature binding in and across perception and action. *Trends Cogn Sci.* 2004;**8**(11):494–500.

125. Wei P, *et al*. Neural correlates of binding features within- or cross-dimensions in visual conjunction search: an fMRI study. *Neuroimage.* 2011;**57**(1):235–241.

126. Parra MA, *et al*. Neural correlates of shape-color binding in visual working memory. *Neuropsychologia.* 2014;**52**: 27–36.

127. Quamme JR, *et al*. Effect of unitization on associative recognition in amnesia. *Hippocampus.* 2007;**17**(3):192–200.

128. Bastin C, *et al*. Associative memory in aging: the effect of unitization on source memory. *Psychol Aging.* 2013;**28**(1):275–283.

129. Ahmad FN, *et al*. Improving associative memory in older adults with unitization. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn.* 2015;**22**(4):452–472.