UNIVERSITÀ DEGLI STUDI FIRENZE

DOTTORATO DI RICERCA IN Neuroscienze

CICLO XXXIV

COORDINATORE Prof.ssa Maria Pia Amato

Sleep-related false memories production: the influence of impaired sleep quality and of characteristics of the memory task

Settore Scientifico Disciplinare M-PSI/01

Dottorando
Dott. *Malloggi Serena*

Tutore
Prof.ssa *Fiorenza Giganti*

Coordinatore
Prof. *Amato Maria Pia*

Anni 2018/2021

Index

Introduction

Human memory does not passively store our experiences. On the contrary, it represents an active and highly-constructive process (Bartlett, 1932) capable of transforming the record of past events into new representations that can substantially differ from what was originally encoded. In this perspective, studying false memories can offer an intriguing opportunity to understand how our memory is malleable and sometimes subjected to distortions.

A false memory occurs when someone erroneously recalls information or reports events that did not happen or remembers them differently from how they originally appeared (Roediger & McDermott, 1995). This error of commission happens involuntarily and can also closely resemble a person's actual experience (Brainerd et al., 2011).

The past decades have seen a growing scientific interest in investigating false memories and their influencing factors; however, among variables potentially affecting false memories production in the general population, such as emotions (Kaplan et al., 2016), stress (Payne et al., 2006) or age (Devitt & Shacter, 2016), sleep has only recently been addressed.

In this regard, some studies have pointed out the existence of a relationship between sleep and false memories through different experimental paradigms. As will be explained in **Chapter 1**, these researches yielded controversial results and produced different hypotheses on the mechanisms linking sleep and false memories production. In addition, this relationship has been investigated especially in healthy samples, neglecting populations of clinical interest whose study could be relevant to further understand how sleep quality affects the production of false memories. For example, up to now, no studies investigated the false memories phenomenon in relation to one of the most widespread sleep disorders in the general population, that is insomnia.

For its peculiarities, insomnia disorder might offer an interesting opportunity to study the mechanisms underlying false memories production. Specifically, its investigation may clarify whether chronically disturbed sleep influences false memories production by impacting on both the processes of sleep-dependent memory consolidation and memory retrieval.

Therefore, the first main aim of my research has been to investigate the influence of sleep quality on false memories production, by comparing individuals with insomnia symptoms and good sleepers in different conditions (**Chapter 2, 3, 4**).

Another object of my work concerned methodological aspects of research on sleep and false memories. Indeed, the above-mentioned mixed results emerging from the literature might also be explained in light of some methodological issues, regarding for example the nature of the task (recall vs recognition) and the presence of baseline measures of learning efficacy. Therefore, **Chapter 5** will describe two studies aimed to investigate the effect of sleep vs wake on false memories production, taking into account these variables.

References

Bartlett, S.F. (1932). Remembering. University Press: Oxford England.

Brainerd, C. J., Reyna, V. F., & Holliday, R. E. (2012). Development of recollection: A fuzzy-trace theory perspective. In S. Ghetti & P. J. Bauer (Eds.), Origins and development of recollection: Perspectives from psychology and neuroscience (pp. 101–143). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195340792.003.0005

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. Neuropsychologia, 91, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Kaplan, R. L., Van Damme, I., Levine, L. J., & Loftus, E. F. (2016). Emotion and False Memory. Emotion Review, 8(1), 8–13. https://doi.org/10.1177/1754073915601228

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. Journal of Memory and Language, 35(2), 261–285. https://doi.org/10.1006/jmla.1996.0015

Roediger H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of experimental psychology: Learning, Memory, and Cognition, 21(4), 803

Chapter 1

False memories and sleep

1.1. False memories: an overview

Memory traces are not stored as exact records of our experiences. Human memory is a constructive process (Bartlett, 1932) capable of transforming the record of past experiences into new representations that might substantially differ from what was originally encoded. In particular, its highly constructive nature can induce changes in the original memory traces (Schacter, 1995) by integrating experienced events with pre-existing information already stored in long-term memory (Loftus & Pickrell, 1995); moreover, it can also lead individuals to remember events that never actually occurred, as in the case of false memories (Loftus, 1993; Roediger & McDermott, 1995).

The term "false memories" refers to a wide range of phenomena, that can be observed in different settings, such as in laboratory, in psychotherapeutic settings, or in other areas of real life, in which people erroneously recall information or events that did not happen or claim to remember them differently from how they originally appeared (Roediger & McDermott, 1995). False memories are not simply false statements, because they involve claims about remembering. In fact, individuals who produce false memories sincerely believe to have experienced certain events or stimuli and feel that they recall concrete details about them, although they never experienced those stimuli (Payne et al., 1996).

A commonly used procedure for studying false memories in laboratory settings is the *Deese-Roediger-McDermott paradigm* (DRM; Deese, 1950; Roediger & McDermott, 1995; Stadler et al., 1999), a task developed by Roediger and McDermott from the original work of Deese (1959). This task requires participants to learn several lists of words (defined "*studied*" words) each semantically related to a critical word, named "*lure*", that does not appear on the list. Words can be presented orally or visually. In this "*learning phase*", participants are instructed to remember as many words as possible from the lists. For example, if the *lure* is "*sleep*", the words presented within the list could be "*bed*", "*rest*", "*nap*" and so on. This list would induce the study participants to involuntarily activate the semantically related word "*sleep*" even if it does not appear in the list. In the classical procedure (Roediger & McDermott, 1995), participants perform a free recall task immediately after the presentation of each list, as well as a recognition task after a short interval

(3 minutes) (*test phase*). In the free recall task, participants are required to write down all words they remember from the original list; in the recognition task, they are shown another list of words (containing *studied* words, *lure* words and unrelated distractors), and asked to identify those included in the studied lists.

In both tasks, participants tend to report that they remember hearing/seeing the *lure* word of each list, thereby producing a false memory. The DRM false memory illusion appears robust, since the critical *lure* words are falsely recalled with a relatively high probability (i.e., from 0.22 to 0.55), and at the recognition task, the *false recognition* rate further increases (i.e., from 0.56 to 0.85; for a review, see Gallo et al., 2001; for the normative data of the Italian DRM word lists, see Iacullo & Marucci, 2016). Overall, this result is imputable to the strong semantic association between the *studied* words and the critical *lure* word (Roediger & McDermott, 1995), whereas the structural differences characterizing the free recall and the recognition task can explain the dissimilar false memory rate observed in the two tasks.

Several theoretical frameworks have been proposed to explain why people produce false memories of the critical *lure* words in the DRM paradigm. In this regard, the most accounted are the *Activation-Monitoring Theory* (Roediger et al., 2001) and the *Fuzzy-Trace Theory* (Brainerd et al., 1995).

The *Activation-Monitoring Theory* (AMT) posits that the false memories derive from a combination of two synergic processes, i.e., on one hand, the activation of the theme semantically related to the critical *lure* word (*Activation*), and, on the other hand, a failure in the source monitoring process (*Monitoring*). Specifically, according to the AMT, the *lure* word of each list could be first internally generated at encoding (and, possibly, during retrieval at subsequent testing) as a result of a semantic spreading activation: during this process, the strong association between the *studied* words triggers the activation of semantic networks which, in turn, elicits the activation of the unstudied critical *lure* word. At subsequent testing (i.e. recall and/or recognition task), participants would produce the false memory due to a source-monitoring error. Specifically, they would mistakenly remember the critical *lure* as having been presented instead of internally generated.

According to the *Fuzzy Trace Theory* (FTT), instead, subjects simultaneously encode two independent traces for each word, respectively the "*verbatim trace*" and the "*gist" or "fuzzy*" trace. The first corresponds to the memory of the actual experience, is related to the contextual features of a word, and is especially linked to veridical memory (corresponding in the DRM paradigm to the "*studied*" words), whereas the "*gist" or "fuzzy*" trace represents the meaning of an item and is preferentially linked to false memories production. The FTT posits that both *verbatim* and *gist* memory influence false memories production albeit in a different manner (Reyna et al., 2016): the *verbatim* traces are generally used to *reject* the *lure* word (i.e. remembering "bed," "rest" etc. produces rejection of "sleep"), whereas the *gist* trace leads to *accepting* it (i.e. remembering that the theme of the list was "sleep" can promote remembering of the *lure* word "sleep" at subsequent testing). In some circumstances, the *gist* trace is more accessible than the *verbatim* one, consequently inducing subjects to produce a false memory. For example, when the semantic association between the *studied* words and the critical *lure* word is strong, or when the delay between the *learning* and the *test* phases is long, participants have easier access to the *gist* trace than the *verbatim* one, which, instead, shows a rapid deterioration in time (Brainerd et al., 1995).

Overall, the literature suggests that ATM and FTT are not mutually exclusive and that the above-mentioned processes could be both involved in false memories production (Steffens & Mecklenbräuker, 2007).

It should be clarified that the DRM paradigm is not the only instrument for studying false memories. We can produce false memories for different materials, therefore not only for semantically related words (Platt et al., 1998). For example, some people can have a distorted memory of events (i.e. false memories in eyewitness) or remember situations and details of their lives that did not happen (i.e. autobiographical false memories). Therefore, depending on the type of false memory, other tasks could be more adequate.

In this regard, an alternative to the DRM is the *misinformation paradigm* (Loftus et al., 1978). Here, participants witness some events and later are given some misinformation (i.e. inconsistent post-event information) about them. When tested on memory for event details, participants often incorporate pieces of misinformation into their memory of the original event and report those details as such (McCloskey & Zaragoza, 1985).

Another alternative to the DRM task is the *imagination inflation* paradigm, which addresses false memories production for autobiographical events (see, for example, Loftus & Pickrell, 1995; Garry et al., 1996). In this case, participants are given a list of events and are asked to rate their confidence on having experienced each event during childhood. Later, they are asked to imagine one of these events (the selected "critical" event) and, later, to fill out another identical life events inventory. This procedure generally induces participants to increase the confidence for the critical event from the first to the second inventory, therefore inducing a false memory for that event (Garry et al., 1996).

Despite the available alternatives, the DRM paradigm represents the most commonly adopted task investigating the false memories phenomenon.

The DRM task has been subjected to multiple modifications depending on the specific aims of researches (for a review, see Gallo, 2013). For example, in studies evaluating the effect of sleep on false memories production (Pardilla Del Gado & Payne, 2017a), the retention interval between the *learning phase* and the *test phase* is generally longer than that of the classical version, and it can be spent in (a) sleep, (b) day-time wake during the day, or (c) night-time wake (for studies based on sleep deprivation paradigms). Moreover, the memory of the word lists is generally tested (either through free recall or recognition tasks) directly after the retention period, without any immediate recall of words during the *learning phase*.

1.2. False memories and sleep

The effect of sleep on mnemonic processes has received interest for over a century (Ebbinghaus, 1885; Jenkins & Dallenbach, 1924; Diekelmann & Born, 2010). To date, it is well known that sleep plays a relevant role in memory processing. It is broadly accepted that a period of post-learning sleep promotes the consolidation of newly acquired memory traces to a greater extent than a wake interval of equal length. This phenomenon is known as "*sleep effect*" (Jenkins & Dallenbach, 1924) and it has been observed adopting numerous different cognitive tasks, both in the declarative and procedural domains (for a review, see Diekelmann & Born, 2010).

While several studies and literature reviews support the beneficial effect of sleep on veridical memory consolidation, the relationship between sleep and false memories is not as clear and requires further investigations. These challenges are mainly imputable to the controversial results obtained, to the methodological differences of available studies, as well as to the discrepancies relative to the mechanisms proposed to explain how sleep affects false memories production.

1.2.1. Evidence about the influence of sleep on false memories production

To date, available studies suggest that sleep can exert an influence on false memory production (for a review, see Newbury & Monaghan, 2019). In this regard, the first research question addressed by studies in the field concerned the existence of a *sleep effect* for false memories. To verify this hypothesis, several studies compared performance at a false memory task (mostly the DRM task) after a retention period spent asleep or awake (Payne et al., 2009; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017b). These studies obtained mixed findings, either concluding that sleep, compared to wake, *promotes* false memories production, or that it *reduces* false memories.

Sleep may potentially promote false memory

In a between-subjects design, Payne and colleagues (2009) reported that healthy individuals who slept after encoding of the DRM word lists, at morning testing falsely recalled more critical *lure* words, as well as *studied* words, than participants who spent the retention interval in daytime wakefulness. In the same study, the authors replicated this result with a nap paradigm, observing that a 90-min nap, compared to a 90-min wake period, increased *false recalls* but not *veridical recalls*. These data were later replicated through the same paradigm (Diekelmann et al., 2010; Pardilla Delgado & Payne, 2017b) as well as in a nap study (Shaw & Monaghan, 2017).

Moreover, the *sleep effect* for false memories has been observed also for emotionally negative words (McKeon et al., 2012) and appeared resistant across long delays (Pardilla Delgado & Payne, 2017b).

Interestingly, this effect has been mainly observed in studies adopting a free recall task to assess false memories production at the DRM paradigm (for a meta-analysis, see Newbury & Monaghan, 2019), whereas only one research found a *sleep effect* with a recognition task, in a nap study (Shaw & Monaghan, 2017).

Overall, the above-mentioned studies suggest that sleep *promotes* false memories production. This phenomenon can be explained by taking into account memory consolidation processes occurring during sleep. As previously mentioned, sleep facilitates the consolidation of newly acquired, and initially unstable, memories for long-term storage, protecting them against decay (for a review, see Born & Wilhelm, 2010). Moreover, both neurophysiological (for a review, see Rash & Born, 2013) and behavioural evidences (for a review, see Landmann et al., 2014; Conte & Ficca, 2013) agree that sleep *actively* consolidates memory traces. Specifically, during sleep, newly acquired memory traces seem to be strengthened in distinct neural circuits (*synaptic consolidation process*), then redistributed to other brain regions for long-term storage and finally integrated within pre-existing memories (*system consolidation process*). As shown by behavioural studies (reviewed in Conte & Ficca, 2013; Landmann et al., 2014), memory traces appear to undergo qualitative changes during this process. In other words, sleep not only sustains a *quantitative strengthening* of new memory traces but also promotes their *qualitative reorganization*. One of the possible consequences of this sleep-dependent memory "*reshaping*" (Conte & Ficca, 2013) is the emergence of new memory content that has not been directly learned (i.e. false memories). In this perspective, false memories represent a natural consequence of the active memory consolidation process occurring during sleep (Landmann et al., 2014; Conte & Ficca, 2013).

Sleep may potentially protect against false memories production

Some studies are not in accordance with the above-mentioned results. For example, Diekelmann and colleagues (2008), employing a recognition task, did not find a greater amount of false memories after sleep compared to a wakefulness retention interval in healthy subjects.
Moreover, Fenn and colleagues (2009) reported that sleep, compared to a wake condition, *reduced* the number of false recognitions of critical *lure* words. Similarly, Lo et al. (2014) found a *decrease* of false recognitions in the sleep condition in a sample of healthy older adults.

To explain their result, Fenn and colleagues (2009) argued that sleep might enhance the accuracy of the source monitoring process (although source monitoring ability was not directly investigated in the study) and advanced several hypotheses on the possible mechanisms involved. The first one assumes that sleep might have favoured the consolidation of contextual and item-specific details associated with DRM *studied* words. Alternatively, sleep might have weakened the association between contextual details and the *gist* trace*:* in fact, in both studies participants in the sleep condition tended to frequently attribute "*know*" responses[1] to the critical *lure* words, suggesting low confidence in their false memories. Another possibility is that sleep might have strengthened memory for internal cognitive processes that occurred at the time of encoding, promoting an efficient and accurate retrieval.

Overall, it is interesting to note that sleep appears to inhibit false memories production when performance is tested with a recognition task, whereas studies using free recall tasks (see, for example, Payne et al., 2009) usually obtained the opposite result. As will be discussed more in detail in **paragraph 1.4**, the presentation of words during the recognition task might reinstate the context of encoding and reactivate sensory details of the original list words, consequently facilitating the discrimination between *studied* and non-*studied* words and avoiding the false recognition of critical *lure* words (Curran et al., 2001).

Therefore, it could be the case that the free recall and the recognition tasks, in light of their structural differences, might differently influence the performance at the DRM paradigm and, consequently, the production of false memories in relation to sleep (Newbury & Monaghan, 2017).

[1] In the recognition task of the DRM paradigm, participants were shown a set of words corresponding to the *studied* words, the critical *lure* words, and unrelated distractors. Then, are asked to indicate whether each of these words is "*new*" or "*old*", referring to the lists previously learned. For the "*old*" response, they also have to specify the quality of their memory by selecting "*Remember*" (meaning that "*this word appeared in the study lists, and I can mentally relive the experience*") or "*Know*" (corresponding to "*I am confident that the item appeared on the list but unable to re-experience its occurrence"*).

Sleep loss increases false memories production

The relationship between sleep and false memories has also been studied in another perspective, i.e., by considering false memories as a consequence of sleep loss.

A robust literature indicates that sleep loss negatively affects various cognitive functions that are essential to avoid false memories production. In fact, sleep deprivation impairs cognitive functions that are strongly related to the integrity of prefrontal cortex (PFC), such as inhibitory control (Drummond et al., 2006), memory retrieval processes (Durmer & Dinges, 2005), and source monitoring (Harrison & Horne, 2000).

Notably, the PFC has been specifically implicated in false recognition. For example, empirical reports described an increase in false memory production in patients with frontal lobe damage compared to controls (Curran et al., 1997; Schacter et al., 1996). Neuroimaging and electrophysiological studies also corroborate the notion that prefrontal areas are involved in retrieval monitoring during performance at false memory paradigms (see, for review, Schacter & Slotnick, 2004).

Therefore, it is plausible that sleep loss, by negatively impacting memory retrieval, would make subjects susceptible to produce false memories. Among the first to address the relationship between sleep loss and false memories, Diekelmann and colleagues (2008) tested this hypothesis in a series of experiments. Specifically, they compared performance at the DRM task across three conditions in which the *learning* phase was followed by a retention interval including: a) two nights of undisturbed sleep (control condition), b) sleep deprivation on the first night, immediately following learning, in order to test the effect of sleep loss on the memory consolidation process, c) sleep deprivation on the second but not first night, to test the effect of sleep deprivation on the memory retrieval phase. At re-test (a recognition task, which followed the two nights in all conditions), only participants who had been sleep deprived on the second night, rather than on the first, falsely recognized more critical *lure* words compared to controls.

Other studies also reported an increase of false recognitions of critical *lure* words after sleep deprivation. For example, in a recent study by Verma & Kashap (2019), participants either slept or stayed awake on the experimental nights after the DRM *learning* session, and they were tested with a recognition task 36 hours after, including a night of recovery sleep. Again, it was found that sleep deprivation led to higher false memory production in comparison to undisturbed sleep.

Similar results were obtained with a free recall task, again using the DRM paradigm (Diekelmann et al., 2010). In this study, performance was assessed under three conditions: “Sleep”, “Night wake” and “Day wake” (**Figure 1a**).

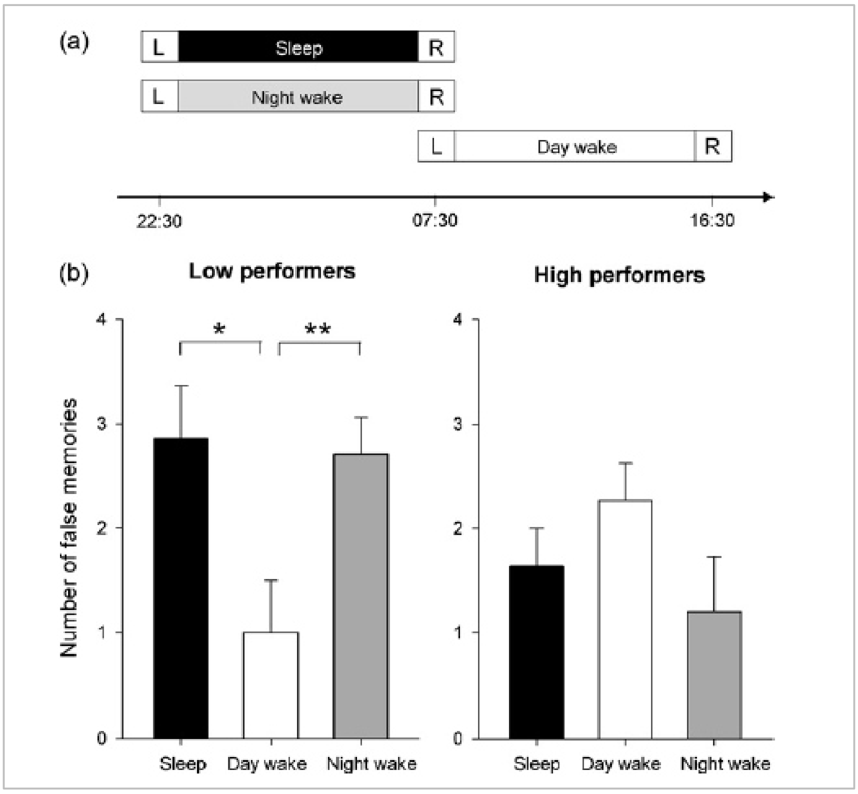

Figure 1. (a) Experimental design of the study. L = learning phase of DRM, R = testing phase of DRM with free recall task. (b) Between-groups differences in the number of false memories. The figure is from Diekelmann et al. (2010)

Results showed that both acute sleep deprivation and undisturbed sleep during the retention period significantly enhanced false recalls of critical *lure* words compared to daytime wakefulness. Additionally, the authors noticed that this effect was influenced by the individual level of memory performance: those who recalled less *studied* words on average (i.e., "low performers") were most likely to exhibit an increase in false memories production after sleep and sleep deprivation (**Figure 1b**).

Further confirmation came from studies employing a different false memories paradigm. For example, an increase in false memories production after sleep loss has also been observed with the misinformation paradigm (Frenda et al., 2014; Lo et al., 2016).

Only one study reported a reduction of false recognitions after both sleep deprivation and sleep restriction compared to undisturbed night-sleep through the DRM paradigm (Chatburn et al., 2017). However, this study did not include a long retention interval between the *DRM learning phase* and the *test phase*, and the DRM paradigm was entirely performed after the night-sleep (sleep condition) or the night-wake (sleep deprivation and restriction conditions).

Overall, available studies show that sleep deprivation can significantly impact false memories production. This effect has mostly been observed with the DRM paradigm and, specifically, with recognition tasks, whereas only one study replicated this effect with a free recall task. These data support the idea that sleep loss promotes false memories production by negatively impacting the frontally-mediated executive functions that ensure an efficient memory retrieval.

1.3. False memories and sleep quality

As previously described, the study of the relationship between false memories and sleep develops around two main lines of research: on one hand, investigating the effect of memory *consolidation* occurring during undisturbed sleep, compared to wakefulness, on false memories production; on the other hand, addressing whether sleep deprivation enhances susceptibility to produce false memories by negatively affecting the *retrieval* process.

In both lines of research, all of the available studies have been conducted on healthy subjects. However, it is reasonable to expect that when sleep patterns are chronically disturbed, such as in sleep-disordered populations, the ability to optimally *consolidate* or *recover* memory traces may be impaired, with consequences on false memories production. Therefore, a further interesting approach could be to investigate the role of *"sleep quality"* on this phenomenon. Indeed, to our knowledge, the influence of sleep quality on false memories production has not yet been systematically addressed.

In fact, only two studies evaluated false memories in samples with sleep disturbances, i.e., one with Obstructive Sleep Apnea Syndrome (Sarhane et al., 2014) and the second with sleep problems related to pregnancy (Berndt et al., 2014). In both studies, false memories production was higher in the sleep-disordered compared to the control group. However, in these studies sleep characteristics of the night preceding the experimental session were not controlled.

Surprisingly, the false memories phenomenon has not been addressed in insomnia disorder, representing one of the most prevalent health concerns in the general population and in clinical practice (Buysse, 2013; Ohayon et al., 2009; Aikens & Rouse, 2005).

Insomnia is a sleep disorder characterized by subjective complaints of non-restorative sleep and of difficulties in initiating and/or maintaining sleep, accompanied by decreased daytime functioning, which persists in time (American Psychiatric Association, 2013).

As for sleep characteristics, several studies documented objective sleep impairments in this population, such as changes in sleep architecture (i.e. reduction in slow-wave sleep and REM sleep duration) and more fragmented sleep compared to healthy subjects (Baglioni et al., 2013).

Insomnia disorder is also accompanied by several diurnal symptoms. Fatigue and mood disturbances (i.e. anxiety and negative mood) are the most commonly reported, but complaints related to altered cognitive functioning are also frequent (Fortier-Brochu et al., 2012; Shekleton et al., 2010). In this regard, it has been observed that individuals with insomnia perform more poorly than good sleepers on complex cognitive tasks depending on the efficiency of the prefrontal cortex: for example, deficits in working memory (e.g., retention and manipulation of previously acquired information), problem-solving, information processing and selective attention have been reported in this population (Ballesio et al., 2019; Wardle-Pinkston et al., 2019). Available studies in people with insomnia also shed light on deficits in performing declarative memory tasks including free recall (Bonnet & Arand, 1995) or recognition (Rosa & Bonnet, 2000) of unrelated words and in a verbal fluency task (Mendelson et al., 1984), although results are not always in accordance (Vignola et al., 2000; Orff et al., 2007).

For its peculiarities relative to the cognitive functioning and to the characteristics of the sleep episode, insomnia disorder might be a suitable way to study mechanisms underlying the production of false memories. In particular, focusing on insomnia allows us to study the influence of sleep quality on false memories assuming these different perspectives:

(a) On one hand, assessing whether chronic poor sleep, by impacting on frontally-mediated executive functioning, could increase the subjects' susceptibility to commit source monitoring errors and produce false memories;
(b) On the other hand, investigating whether chronic poor sleep affects the process of memory consolidation and reorganization occurring during sleep, with consequences on false memories production.

To date, it is unknown whether false memories represent one of the possible "costs" of the chronic poor sleep condition affecting individuals with insomnia. Moreover, the few studies addressing sleep-dependent memory consolidation in insomnia (Cellini, 2017) suggest that it may be compromised in this population. However, these studies were conducted using traditional declarative and procedural tasks (Cellini, 2017). Therefore, the effect of disturbed sleep on more complex memory *reshaping* processes (Conte & Ficca, 2013; Landmann et al., 2014), such as those inducing false memories, remains unexplored.

The studies presented in the next chapters aim to clarify these issues. Specifically, they address how poor sleep in insomnia influences false memories production by impacting on the efficiency of the memory retrieval processes (**Chapters 2 and 3**) and on the memory consolidation and reshaping processes occurring during sleep (**Chapter 4**).

1.4. Methodological aspects of studying the relationship between sleep and false memories production

The previously highlighted discrepancies in data on the relationship between sleep and false memories may be linked to some methodological concerns.

First of all, the effect of sleep on false memories production, evaluated through the DRM paradigm, appears to differ depending on the task adopted, namely free recall or recognition tasks.

In a recent meta-analysis, Newbury and Monaghan (2019) investigated factors potentially influencing false memories production in relation to sleep. To run their analyses, they included only 9 studies, selected for the adoption of the DRM paradigm and for the presence of both a sleep and a wake condition. The meta-analysis revealed that the effect of sleep on false memories production appears to be mainly moderated by the task adopted in the *test* phase (i.e. free recall vs recognition). In fact, they reported that this effect is generally greater with free recall than recognition tasks. Specifically, the studies adopting a free recall task globally reported an increase of the number of critical *lure* words falsely recalled after sleep compared to diurnal wake. Instead, research conducted with recognition tasks reported non-significant effect sizes, indicating that the amount of false memories does not differ between the sleep and wake conditions.

The difference between the DRM free recall and recognition task actually remains one of the main critical issues when examining the effect of sleep on false memories production. In fact, if sleep leads to a greater spread of semantic activation of previously *studied* words (see, for example, Landamann et al., 2014), then we should expect to see an increase both in *false recall* and *recognition* of critical *lure* words after sleep in comparison to wakefulness.

Possible explanations to controversial findings have been suggested (Diekelmann et al., 2010). The first concerns the different cognitive operations elicited by the recognition and the free recall tasks at memory retrieval (Tulving & Madigan, 1970). Specifically, in the recognition task, subjects are explicitly asked to discriminate between "old" and "new" items from a lists of words presented by the experimenter. In this case, the words represent external memory cues to choose from. Instead, in the free recall test subjects are requested to write down all words they remember from the original lists, in total absence of external memory cues that could aid retrieval.

In this perspective, it is plausible that sleep's reorganizing role on memory traces may be more easily highlighted through a free recall rather than a recognition task. In fact, in a DRM paradigm, post-learning sleep, compared to wake, would benefit the consolidation of veridical information (Diekelmann & Born, 2010; Payne et al., 2009) and, in parallel, it would favour the extraction of the general meaning of the set of words ("*gist*"; Landmann et al., 2014; Payne et al., 2009). Later, in the free recall task, the process of "self-cueing" would facilitate access to the *gist* trace activated during sleep, therefore highlighting the effects of sleep on the memory trace especially in terms of "*reshaping*" (Conte & Ficca, 2013). Instead, the presentation of words in a recognition procedure would presumably reinstate the context of encoding and favour the retrieval of the details of the *studied* words consolidated during sleep. Therefore, this process would simplify the discrimination between *studied* and *lure* words, preventing false recognitions.

Another possible explanation refers to previous data suggesting that sleep, compared to wake, is more beneficial when task difficulty increases (Stickgold &Walker, 2004; Sio et al., 2013). In this regard, it could be argued that a free recall task cognitively engages subjects to a great extent than a recognition task: while in the recognition task the procedure is guided by the experimenter, who provides participants with external cues to choose from, in the free recall task subjects are requested to generate their own cues to retrieve the original memory traces. In this perspective, the promoting role of sleep on false memories production might more clearly emerge with the free recall rather than the recognition task.

To date, only one study (Pardilla Delgado & Payne, 2017b) investigated the sleep effect on false memories adopting both a free recall and a recognition task in a DRM paradigm. The authors found a selective effect of sleep on the free recall task, whereas no differences between the sleep and the wake conditions emerged at the recognition task. However, this negative result could depend on the temporal order of tasks administration rather than on the tasks *per se*. In fact, in this study participants performed the free recall task always first. As the same authors suggest, the previous free recall could have increased recognition performance to a ceiling effect, resulting in non-significant differences between conditions.

Another critical aspect when addressing the relationship between sleep and false memories concerns the evaluation of the baseline measures of memory performance and its influence on subsequent, delayed, testing (after the retention period, which usually spans over several hours in a sleep-memory study). Most of the available studies did not assess baseline DRM performance (i.e., through an immediate recall task) but directly tested it after the retention interval spent in sleep or wakefulness (**paragraph 1.2.1**). However, as in classical sleep-memory studies, the evaluation of baseline performance is critical in order to exclude that performance observed at delayed testing depends on the effectiveness of the encoding phase rather than specifically on consolidation processes occurred over the retention interval (see Conte & Ficca, 2013). This reasoning applies all the more to the study of false memories. In fact, in absence of baseline measures of false memories production, it is impossible to determine whether those produced after sleep or wake reflect consolidation of false memories produced at encoding or the emergence of "new" false memories occurred over the retention interval. In addition, false memories performance at immediate testing may represent a useful index of individual propensity to produce false memories.

Payne and colleagues (2009) attempted to address this issue through a between-subjects design. In a DRM paradigm, an immediate free recall test was administered to a control group, while the Sleep and Wake groups performed the free recall test directly after the retention interval spent asleep or awake, respectively, as in most studies of the field. The authors observed that, compared to the control participants, while veridical recall deteriorated in both experimental groups, false recalls were reduced only in the Wake group whereas they were preserved in the Sleep group. Further studies using within-subjects design are warranted to corroborate this interesting finding.

Furthermore, investigating the effect of immediate on delayed testing, e.g., by comparing delayed performance on stimuli that have already undergone immediate testing vs stimuli that have not, may be interesting *per se*. In fact, it could be the case that sleep differently consolidates memory traces that have undergone different depths of encoding, with different effects also on false memories production.

In conclusion, the available literature on sleep and false memories should be interpreted with caution due to several methodological issues that require further clarification. The studies that will be presented in **Chapter 5** have been conducted to address some of these concerns, namely:

a) to compare sleep-related false memories production between a free recall and a recognition task;
b) to explore whether an immediate testing phase, aimed to collect baseline performance measures, affects subsequent memory re-processing of false memory traces occurring during sleep or wake.

References

Aikens, J. E., & Rouse, M. E. (2005). Help-seeking for insomnia among adult patients in primary care. The Journal of the American Board of Family Practice, 18(4), 257–261. https://doi.org/10.3122/jabfm.18.4.257

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders, 5th ed. American Psychiatric Association.

Baglioni, C., Regen, W., Teghen, A., Spiegelhalder, K., Feige, B., Nissen, C., & Riemann, D. (2013). Sleep changes in the disorder of insomnia: A meta-analysis of polysomnographic studies. Sleep Medicine Reviews, 18, 195–213. https://doi.org/10.1016/j.smrv.2013.04.001

Ballesio, A., Aquino, M., Kyle, S. D., Ferlazzo, F., & Lombardo, C. (2019). Executive Functions in Insomnia Disorder: A Systematic Review and Exploratory Meta-Analysis. Frontiers in Psychology, 10, 101. https://doi.org/10.3389/fpsyg.2019.00101

Bartlett, S.F. (1932). Remembering. University Press: Oxford England.

Berndt, C., Diekelmann, S., Alexander, N., Pustal, A., & Kirschbaum, C. (2014). Sleep fragmentation and false memories during pregnancy and motherhood. Behavioural Brain research, 266, 52–57. https://doi.org/10.1016/j.bbr.2014.02.030

Bonnet, M.H., & Arand, D.L. (1995). 24-hour metabolic rate in insomniacs and matched normal sleepers. Sleep, 18, 581-588.

Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. Journal of Memory and Language, 34(2), 157–185. https://doi.org/10.1006/jmla.1995.1008

Buysse, D. J. (2013). Insomnia. JAMA, 309(7), 706–716. https://doi.org/10.1001/jama.2013.193

Cellini, N. (2017). Memory consolidation in sleep disorders. Sleep medicine reviews, 35, 101–112. https://doi.org/10.1016/j.smrv.2016.09.003

Chatburn, A., Kohler, M. J., Payne, J. D., & Drummond, S. (2017). The effects of sleep restriction and sleep deprivation in producing false memories. Neurobiology of learning and memory, 137, 107–113. https://doi.org/10.1016/j.nlm.2016.11.017

Conte, F., & Ficca, G. (2013). Caveats on psychological models of sleep and memory: A compass in an overgrown scenario. Sleep Medicine Reviews, 17, 105–121. https://doi.org/10.1016/j.smrv.2012.04.001

Curran, T., Schacter, D. L., Norman, K. A., & Galluccio, L. (1997). False recognition after a right frontal lobe infarction: Memory for general and specific information. Neuropsychologia, 35(7), 1035–1049.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58, 17-22.

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. Neuropsychologia, 91, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Diekelmann, S., & Born, J. (2010). The memory function of sleep. Nature Reviews Neuroscience, 11(2), 114–126.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. Behavioural Brain Research, 208(2), 425–429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. PLoS ONE, 3(10), e35

Drummond, S. P., Paulus, M. P., & Tapert, S. F. (2006). Effects of two nights sleep deprivation and two nights recovery sleep on response inhibition. Journal of Sleep Research, 15, 261–265. https://doi.org/10.1111/j.1365-2869.2006.00535.x

Durmer, J.S., & Dinges, D.F. (2005). Neurocognitive consequences of sleep deprivation. Seminars Neurology, 25, 117–129

Ebbinghaus, H. (1885). Memory. Columbia University, New York.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. Learning & Memory, 16(9), 509–513.

Fortier-Brochu, E., Beaulieu-Bonneau, S., Ivers, H., & Morin, C. M. (2012). Insomnia and daytime cognitive performance: A meta-analysis. Sleep Medicine Reviews, 16, 83-94.

Frenda, S. J., Patihis, L., Loftus, E. F., Lewis, H. C., & Fenn, K. M. (2014). Sleep Deprivation and False Memories. Psychological Science, 25(9), 1674–1681. https://doi.org/10.1177/0956797614534694

Gallo, D. (2013). Associative illusions of memory: False memory research in DRM and related tasks. Psychology Press.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. Psychonomic Bulletin & Review, 3(2), 208–214.

Harrison, Y., & Horne, J. A. (2000). The impact of sleep deprivation on decision making: a review. Journal of experimental psychology. Applied, 6(3), 236–249. https://doi.org/10.1037//1076-898x.6.3.236

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. The American Journal of Psychology, 605–612.

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. Sleep medicine reviews, 18(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Lo, J. C., Chong, P. L., Ganesan, S., Leong, R. L., & Chee, M. W. (2016). Sleep deprivation increases formation of false memory. Journal of Sleep Research, 25(6), 673–682. https://doi.org/10.1111/jsr.12436

Lo, J. C., Sim, S. K., & Chee, M. W. (2014). Sleep reduces false memory in healthy older adults. Sleep, 37(4), U247–U665

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. Journal of Experimental Psychology: Human Learning and Memory, 4, 19–31.

Loftus, E.F. & Pickrell, J.E. (1995). The formation of false memories. Psychiatric Annals, 25, 720-725.

Loftus, E.F. (1993). The reality of repressed memories. American Psychologist, 48, 518–537.

McCloskey, M., & Zaragoza, M. (1985). Misleading post-event information and memory for events: Arguments and evidence against memory impairment hypotheses. Journal of Experimental Psychology: General, 114(1), 1–16. doi:10.1037/0096-3445.114.1.1

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. (2012). Interaction of sleep and emotional content on the production of false memories. PLoS ONE, 7(11), e49353.

Mendelson, W.B., Garnett, D., Gillin, J.C., & Weingartner, H. (1984). The experience of insomnia and daytime and nighttime functioning. Psychiatry Research, 12, 235 – 50.

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. Psychonomic bulletin & review, 26(2), 387–400. https://doi.org/10.3758/s13423-018-1528-4

Ohayon, M. M., & Reynolds, C. F., 3rd (2009). Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). Sleep Medicine, 10(9), 952–960. https://doi.org/10.1016/j.sleep.2009.07.008

Orff, H.J., Drummond, S.P, Nowakowski, S., & Perils, M.L. (2007). Discrepancy between subjective symptomatology and objective neuropsychological performance in insomnia. *Sleep,* 30, 1205 -11.

Pardilla Delgado, E., & Payne, J. D. (2017a). The Deese-Roediger-McDermott (DRM) Task: A Simple Cognitive Paradigm to Investigate False Memories in the Laboratory. Journal of visualized experiments: JoVE, (119), 54793. https://doi.org/10.3791/54793

Pardilla-Delgado, E., & Payne, J. D. (2017b). The impact of sleep on true and false memory across long delays. Neurobiology of Learning and Memory, 137, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. Journal of Memory and Language, 35(2), 261–285. https://doi.org/10.1006/jmla.1996.0015

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., ... Stickgold, R. (2009). The role of sleep in false memory formation. Neurobiology of Learning and Memory, 92(3), 327–334.

Platt, R. D., Lacey, S. C., Iobst, A. D., & Finkelman, D. (1998). Absorption, dissociation, fantasy-proneness as predictors of memory distortion in autobiographical and laboratory-generated memories. Applied Cognitive Psychology, 12, S77–S89. https://doi.org/10.1002/(SICI)1099-0720(199812)12:7<S77:AID-ACP601>3.0.CO;2-C

Rasch, B., & Born, J. (2013). About sleep's role in memory. Physiological reviews, 93(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How Fuzzy-Trace Theory Predicts True and False Memories for Words, Sentences, and Narratives. Journal of applied research in memory and cognition, 5(1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Roediger H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of experimental psychology: Learning, Memory, and Cognition, 21(4), 803

Roediger, H. L., 3rd, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. Psychonomic bulletin & review, 8(3), 385–407. https://doi.org/10.3758/bf03196177

Rosa, R.R., & Bonnet, M.H. (2000). Reported chronic insomnia is independent of poor sleep as measured by electroencephalography. Psychosomatic Medicine, 62, 474-82.

Sarhane, M., Etcheverry, I., Tiberge, M., & Daurat, A. (2014) False memories in patients suffering from obstructive sleep apnoea syndrome (OSAS), Journal of Cognitive Psychology, 26:6, 662-669, DOI: 10.1080/20445911.2014.937812

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. Neuron, 44(1), 149–160. https://doi.org/10.1016/j.neuron.2004.08.017

Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. P., & Bates, J. F. (1996). False recognition and the right frontal lobe: A case study. *Neuropsychologia*, *34*(8), 793–808.

Shaw, J. J., & Monaghan, P. (2017). Lateralised sleep spindles relate to false memory generation. Neuropsychologia, 107, 60–67. https://doi.org/10.1016/j.neuropsychologia.2017.11.002

Shekleton, J. A., Rogers, N. L., & Rajaratnama, S. M. (2010). Searching for the daytime impairments of primary insomnia. Sleep Medicine Reviews, 14, 47-60.

Sio, U., Monaghan, P., & Ormerod, T. (2013). Sleep on it, but only if it is difficult: Effects of sleep on problem solving. Memory & Cognition, 41, 159–166. https://doi.org/10.3758/s13421-012-0256-7

Stadler, M. A., Roediger, H. L. III, & McDermott, K. B. (1999). Norms for word lists that create false memories. Memory & Cognition, 27(3), 494–500. https://doi.org/10.3758/BF03211543

Steffens, M. C., & Mecklenbräuker, S. (2007). False memories: Phenomena, theories, and implications. Zeitschrift für Psychologie/Journal of Psychology, 215(1), 12–24. https://doi.org/10.1027/0044-3409.215.1.12

Stickgold, R., & Walker, M. (2004). To sleep, perchance to gain creative insight? Trends in Cognitive Sciences, 8, 191–192. https://doi.org/10.1016/j.tics.2004.03.003

Tulving, E., & Madigan, S.A. (1970). Memory and verbal learning. Annual Review of Psychology, 21,437–8.

Verma, K., & Kashyap, N. (2019). Sleep Deprivation Enhances False Memory on the Deese-Roediger-McDermott (DRM) Task. Psychological Thought, 12(1), 120-130. doi:http://dx.doi.org/10.5964/psyct.v12i1.339

Vignola, A., Lamoureux, C., Bastien, C. H., & Morin, C. M. (2000). Effects of chronic insomnia and use of benzodiazepines on daytime performance in older adults. The journals of gerontology. Series B, Psychological sciences and social sciences, 55(1), P54–P62. https://doi.org/10.1093/geronb/55.1.p54

Wardle-Pinkston, S., Slavish, D. C., & Taylor, D. J. (2019). Insomnia and cognitive performance: A systematic review and meta-analysis. Sleep Medicine Reviews, 48, 101205. https://doi.org/10.1016/j.smrv.2019.07.008

Chapter 2

Insomnia symptoms affect false memories production[2]

2.1. Introduction

Individuals suffering from Insomnia (American Psychiatric Association, 2013) complain about non-restorative sleep and difficulties in initiating and/or maintaining sleep, along with impairments in daytime functioning persisting in time. As for diurnal symptoms, fatigue and mood disturbances are the most commonly reported but complaints about altered cognitive functioning are also frequent in this population (Riedel et al., 2000; Fortier-Brochu et al., 2012).

In this regard, available research reports that people with insomnia symptoms show deficits in several cognitive domains (for a review, see Fortier-Brochu et al., 2012), especially in those strongly depending on the efficiency of the prefrontal cortex (for a review, see Ballesio et al., 2019; Wardle-Pinkston et al., 2019). For example, it has been observed that individuals suffering from insomnia perform more poorly than good sleepers in tasks assessing working memory (e.g., retention and manipulation of previously acquired information), problem-solving, information processing, and selective attention (Ballesio et al., 2019; Wardle-Pinkston et al., 2019).

An efficient frontally-mediated executive functioning is crucial for a proper memory retrieval process, to correctly judge whether or not a certain event has actually occurred, as well as to avoid false memories production (Durmer & Dinges, 2005; Frenda & Fenn, 2016).

False memories represent distorted memories or recollections of events that did not actually happen (Roediger & McDermott, 1995) and their development has been explained through different theoretical frameworks (Nichols & Loftus, 2016). Among them, the Activation-Monitoring Theory (ATM; Roediger & McDermott, 1995; Roediger et al., Gallo, 2001) highlights that the memory retrieval phase is the crucial step in the process leading to false memory formation. False memories would result from a combination of two synergic processes: on one hand, an activation of the semantic theme associated to the encoded stimuli (*Activation*) and, on the other hand, a failure in the source monitoring process (*Monitoring*). In this perspective, false memories arise from an error of commission leading individuals, at retrieval, to erroneously discriminate between items that were

2 Malloggi S., Conte F., De Rosa O., Cellini N., Di Iorio I., Ficca G., Giganti F. Insomnia symptoms affect false memories production. Accepted as Poster presentation at World Sleep 2022 in Rome, Italy (March 11-16, 2022).

studied and those that were not. Here, frontally-mediated executive functions are essential to avoid this memory error and to guarantee accurate retrieval (Johnson et al., 2012).

As previously mentioned, individuals with insomnia display diurnal impairments in the cognitive functions involved in rejecting false memories and ensuring accurate recall, i.e. retention and manipulation of information in working memory, inhibitory control, and cognitive flexibility (Ballesio et al., 2019; Wardle-Pinkston et al., 2019). However, false memories production in this population has not yet been addressed.

Here we aim to investigate the influence of poor sleep quality on false memories production by comparing performance at a Deese-Roediger-McDermott task (DRM; Deese, 1959; Roediger & McDermott, 1995) between a sample of individuals with insomnia symptoms and one of good sleepers. The DRM paradigm is most commonly adopted, in laboratory settings, to study false memories formation, since it reliably induces high rates of false memories. Moreover, in light of previous studies suggesting a relationship between frontally-mediated cognitive functioning and false memories production (see, for example, Peters et al., 2007; Leding, 2012), we also assess between-groups differences in executive functioning and source monitoring abilities.

We hypothesize that, compared to good sleepers, individuals with insomnia symptoms will be more susceptible to produce false memories as a consequence of their poor sleep. Furthermore, as for executive functioning and source monitoring ability, we expect an association between performance at these tasks and performance at the DRM task, as well as reduced performance at the former tasks in participants with insomnia symptoms compared to good sleepers.

2.2. Materials and method

2.2.1. Participants

One hundred and ten potential participants were asked to fill out a web-based set of screening questionnaires through the Google Form platform: the Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), the Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016), the Sleep Disorder Questionnaire (SDQ; Violani et al., 2004), the Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007), and the Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007) (described in details below). In addition, they were administered online ad hoc questions in order to assess general medical conditions and health habits, presence of psychiatric disorders and of sleep disorders.

Based on scores at the screening instruments, 68 university students were recruited for the study and included in either the "good sleep group" (**GS group**, n = 35) or the "insomnia symptoms group" (**IN group**, n = 33). Inclusion criteria common to both groups were: absence of any relevant somatic or psychiatric disorder; absence of clinically significant depression and anxiety symptoms (BDI-II score ≤ 29; BAI score ≤ 25); no history of drug or alcohol abuse; absence of sleep disorders (other than insomnia for the IN group) and of any sleep apnea or respiratory disorder symptom; having a regular sleep-wake pattern (e.g., individuals with irregular study or working habits such as shift-working were excluded); no use of psychoactive medication or alcohol at bedtime. In addition, for inclusion in the IN group, participants had to score ≥ 5 at the PSQI, ≥ 8 at the ISI, and to be classified as presenting "clinically significant insomnia" at the SDQ. In this group, participants did not receive a proper diagnosis of insomnia based on a clinical interview.

Instead, inclusion criteria in the GS group were: PSQI score < 5, ISI score < 8, being classified as "good sleeper" at the SDQ.

As reported in **Table 1**, the two groups did not differ in terms of age, gender distribution, circadian preference (measured through the reduced version of the Morningness-Eveningness Questionnaire; Italian version from Natale et al., 2006). Instead, as expected, significant between-group differences emerged in habitual sleep features assessed through the PSQI, specifically sleep duration and bedtime (**Table 1**).

Table 1. Age, gender distribution, circadian preference and habitual sleep features in the IN and GS groups. Significant *p*-values are in bold.

	IN group	GS group	Statistical test
Age	24.4 ± 3.98	24.9 ± 4.35	t = 0.51, p = .611
Gender	12 M, 21 F	10 M, 25 F	χ^2 = 0 .47, p = .492
MEQr score	12.87 ± 3.45	13.22 ± 3.01	t = 0.45, p = .668
Habitual sleep duration	07:33 ± 01:10	08:14 ± 00:56	t = 2.63, p = **.011**
Habitual bedtime	00:37 ± 01:30	23:57 ± 00:53	t = - 2.19, p = **.032**
Habitual rise time	08:22 ± 01:45	08:12 ± 01:14	t = - 0.48, p = .635

Notes. MEQr: Morningness-Eveningness Questionnaire (reduced version); IN group: insomnia group; GS group: good sleep group; M: Males; F: Females. For the variable "age", mean and standard deviation are reported. *t* Student's test is reported for between-groups comparisons for all variables except gender. Habitual sleep duration (hh:mm), bedtime (hh:mm) and rise time (hh:mm) were collected through the PSQI. Results of the chi-squared test are reported for differences in gender distribution.

2.2.2. Procedure

Due to the Covid-19 pandemic, data collection has been entirely performed online. Each subject participated individually in two video-call sessions with the experimenter, who was blind to the study groups. In one session ("*DRM session*"), participants were administered the DRM paradigm (Roediger & McDermott, 1995), whereas in the other one ("*cognitive testing session*") they performed a set of cognitive tests to evaluate executive functioning and source monitoring ability. The two sessions were scheduled in balanced order between-subjects. All sessions were performed in the morning, between 11:00 and 13:00. Participants were requested to complete a sleep diary on the day of the two sessions, in order to control that they performed the tasks after a night of sleep that was representative of their habitual sleep.

The local Ethical Committee approved the research protocol and all participants signed a consent form. There was no money or credit compensation for participating in the study.

2.2.3. Instruments

Screening instruments

(1) Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), a self-report questionnaire evaluating subjective sleep quality in the past month. It is composed of 19-items grouped into 7 subscales: Subjective Sleep Quality, Sleep Latency, Sleep Duration, Habitual Sleep Efficiency, Sleep Disturbances, Use of Sleep Medication and Daytime Dysfunctions due to sleepiness. The PSQI total score ranges from 0 to 21, with higher scores indicating sleep difficulties and lower sleep quality. The cut-off score ≥ 5 is adopted to discriminate between good and bad sleepers.

(2) Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016) assessing the severity of insomnia symptoms during the previous 2 weeks. It permits to classify subjects into four categories: (a) no clinically significant insomnia (score 0 – 7); (b) sub-threshold insomnia (8 – 14); (c) clinical insomnia – moderate severity (score 15 – 21); (d) clinical insomnia – severe (22 – 28).

(3) Sleep Disorder Questionnaire (SDQ; Violani et al., 2004) is a self-rating questionnaire with 27 items evaluating the presence of different sleep problems in the last month. The first three questions concern symptoms of insomnia, while the others investigate the presence of other sleep disorders symptoms. A subsequent set of questions investigates the duration, frequency and consequences of the sleep problem, and is used for the evaluation of the

severity of the sleep disturbances reported. The SDQ permits the classification of subjects into three main categories: subjects who do not complain of any sleep disorder; subjects who report the occurrence of subthreshold insomnia and subjects with clinically significant insomnia.

(4) Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007) is a 21 items self-report instrument assessing the severity of depressive symptoms. Its total score ranges from 0 to 63, with higher scores indicating more severe depressive symptoms. Particularly, scores 0 – 13 represent minimal depression, scores 14 – 19 mild depression, scores 20 – 28 moderate depression, and scores 29 – 63 severe depression symptoms.

(5) Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007) is a 21 items self-report instrument assessing the presence and severity of anxiety symptoms in the past week. The score range is 0 – 63, with higher scores indicating more severe anxiety symptoms: specifically, a total score of 0 – 7 is considered to index minimal severity, 8 – 15 mild, 16 – 25 moderate and 26 – 63 severe.

False memories task

The Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger and McDermott, 1995) has been employed to investigate false memories production for semantically-related words. In our study, participants learned 16 lists of 15 words taken from Iacullo & Marucci (2016) (see **Appendix**). Each list contained words semantically associated with a word not present in the list, defined "*lure*" (e.g., "ink", "paper", "school", all related to "pen"). The words in each list were presented by the experimenter in order of associative strength with the unpresented *lure* (from strongest to weakest) and were read aloud by the experimenter approximately at a 1.5-s rate.

Participants performed this paradigm on their personal computer, while in video-call with the experimenter. They were instructed to listen carefully to the word lists read by the experimenter and were informed that the memory for these words would be tested. After each list was presented, an immediate free recall task was administered, through a module of *Google Form* platform. Participants were allotted 1.5 min to type on the screen all the words they could remember from the list, avoiding to guess.

A final overall recognition test was administered 5 min after recall of the last list, again through a module of *Google Form* platform. The recognition task consisted of a list of 96 words, made up of studied items from the lists (n = 48; three from each list), all the critical *lure* words (n = 16), and unrelated distractors (n = 36), presented in randomized order between subjects. Participants were asked to indicate whether each word was "*old*" (had been presented in the learning phase) or "*new*". If an item was judged "*old*", subjects had to further specify the quality of their memory, distinguishing between "*remembering*" and "*knowing*" judgments. Specifically, subjects were told that a "*remember*" judgment should be made for items for which they had a vivid memory of the actual presentation (for example, if they remembered the speaker's voice while saying the word or what they were thinking when they heard the word), whereas "*know*" judgments were reserved for items that they were sure had been presented but for which they lacked the feeling of remembering the actual occurrence of the words. Detailed instructions on the remember/know distinction were given following the instructions of Roediger & McDermott (1995).

Two experimenters where simultaneously connected in video-call with each participant, to ensure accurate data collection.

Executive functioning tasks

For the assessment of executive functioning we employed classical tasks which measure the main executive components (e.g., Denckla, 1994; Myake, 2000): specifically, working memory was evaluated through the Working Memory subtests of the WAIS-IV (Wechsler, 2008) and inhibitory control was tested through the Stroop task (Stroop, 1935). In addition, we created an ad hoc task aimed to evaluate source monitoring ability, which is deemed to be specifically linked to false memories formation (Mitchell & Johnson, 2000). Below, a brief description of the tasks:

1) Working memory subtests of the WAIS-IV (Wechsler, 2008), including the Digit Span sub-test (made up of three increasingly difficult tasks: digit span forwards, backward, and sequencing) and the Arithmetic subtest (requiring to perform mental arithmetic problems): taken together, performance at these tasks provides the Working Memory Index (WMI), a global measure of the ability to attend to information presented verbally, manipulate it in short-term memory and then formulate a response. The tests were administered by the use of a video call, according to the standard procedure reported in the WAIS-IV manual.

2) Stroop Color and Word Test (Stroop, 1935): here we adopted a computerized version of the task developed on the Open Sesame software (version 3.3.8). The stimuli consisted of the words "red," "green," "yellow" and "blue" presented at the center of a black computer screen in one of the four colors. The color of the word displayed corresponded to its meaning in 50% of the trials (congruent condition), whereas in the remaining 50% of the

trials word color and meaning were different (incongruent condition) (based on Gajewski et al., 2020). Subjects had to indicate, as soon as possible, the color of the text by pressing a key on the keyboard corresponding to the effective color of the text. Subjects performed a short training phase consisting of 24 trials in order to familiarize themselves with the task and afterward they performed the task including 240 trials. Participants performed the task on *MindProbe*, a server to host online experiments. Two experimenters were simultaneously connected via video call to administer the task and to ascertain participant's procedure adherence.

3) Source Monitoring task: a computerized Source Monitoring Task was included to evaluate the ability to discriminate between internal and external sources of information (i.e. *Reality Monitoring* ability; Johnson et al., 1993). We developed this task from Nienhow and Docherty (2014), who originally evaluated *Internal Source Monitoring* ability, which is the capacity to discriminate between two internal sources of information.

In this task, 32 incomplete statements were presented to participants, one at a time, at the centre of the computer screen (e.g., "*the color of grass is* _____."). After each statement, one of two possible messages could appear in the centre of the screen: (a) "*response*" or (b) the word that logically completes the sentence. On 16 statements ("*response*" instruction), subjects were required to vocalize the word (corresponding to the internal source of information), while on the other 16 statements they had to pay attention to the word appearing on the screen, as external source (**Figure 1**). The type of response for each statement were counterbalanced between subjects. Once all statements had been presented, the participant completed a source-recognition task containing 32 target words and 16 new distractor words. We asked to identify the source of each word as "*response*", "*experimenter's response*" or "*new*".

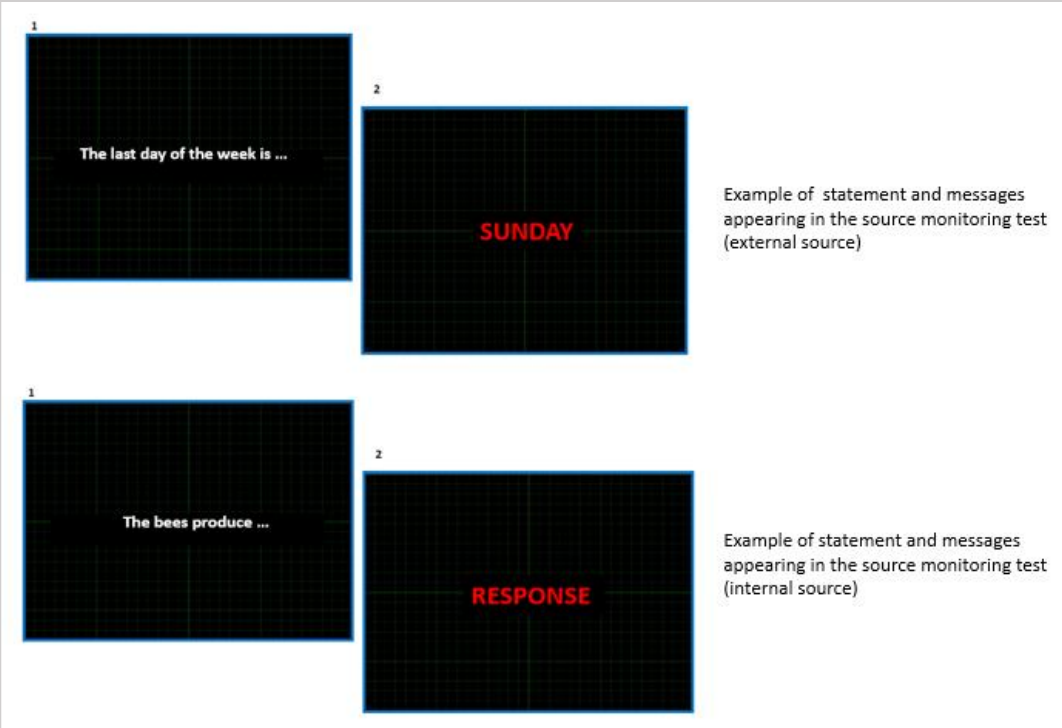

Figure 1. Example of statements and messages appearing during the source monitoring task.

Each statement was first piloted on 50 college students to ensure that they used the same target word to complete each statement. An inter-subject agreement of 99% for each word had been required.

The presentation of the statements was performed through the *MindProbe* platform and the source-recognition task through the *Google Form* platform. Two experimenters were simultaneously connected via video call to administer the task and to ascertain participants' adherence to the procedure.

2.2.4. Data analysis

Outcome measures of the DRM task were:

- At the free recall test, the number of *false recalls* (i.e. total number of falsely recalled critical *lure* words), the number of *veridical recalls* (i.e. corresponding to the total number of words correctly recalled from the original word lists), and the number of *recalled intrusions* (representing the total number of recalled words not corresponding to studied items nor to the critical *lure* words).

- At the recognition test, the number of *false recognitions* (i.e. ‘‘old” responses given to critical *lure* words), the number of *hits* (i.e. ‘‘old” responses given to *studied* words), and the number of *false alarms* (i.e. ‘‘old” responses given to unrelated distractors); the proportion of “*Remember*” judgments attributed to *false recognitions* (“*FR-R”*), *hits* (“*Hits-R”*) and *false alarms* (“*FA-R”*); the proportion of “*Know*” judgments attributed to *false recognitions* (“*FR-K”*), *hits* (“*Hits-K”*) and *false alarms* (“*FA-K”*).

Concerning executive functioning, outcome measures were: digit span scores, arithmetic scores and the working memory index (WMI) obtained from the WAIS-IV subtests; number of correct responses, number of errors and mean response times (ms) for the Stroop task.

Finally, outcome variables considered for the source monitoring task were: the number of words that were correctly attributed to the internal and the external sources of information (“*SM* - *correct”*); the proportion of words correctly identified as from internal source out of the total number of words correctly recognized as “old” (*SM – Index 1*); the proportion of words correctly identified as from external source out of the total number of words correctly recognized as “old” (“*SM* - *Index 2*”).

The Shapiro-Wilk test showed that the majority of variables were normally distributed, therefore we adopted parametric statistics. A *t-Student* test for independent samples was adopted to assess between-groups differences in DRM performance, in executive functioning, and source monitoring variables. A Pearson correlation analysis was adopted to test the association between DRM performance, cognitive testing, and source monitoring ability in the whole sample.

Analyses were performed using SPSS (version 27) and the significance level was set at $p \leq .05$.

2.3. Results

False memory task

At the immediate free recall test, the IN group produced a higher number of *false recalls* (*t* (66) = - 3.23, p = .002; **Figure 2**) and *recalled intrusions* (*t* (66) = - 2.97, p = .004) compared to the GS group, whereas no differences between groups were observed for *veridical recalls* (*t* (66) = - .889, p = .377). **Table 2** displays between-groups comparisons of performance at the free recall task.

At the recognition task, no between-groups differences emerged (**Table 3**).

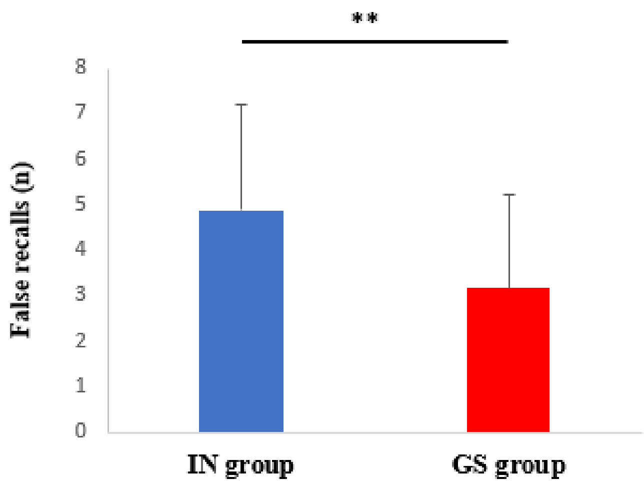

Figure 2. Between-groups comparisons in the number of false recalls. **= p <.01. Error bars represent standard deviations.

Table 2. Comparisons between IN group and GS group in DRM the free recall variables.

	IN group	GS group	*t*	*p*	*Cohen's d*
False recalls	4.87 ± 2.20	3.17 ± 2.05	-3.23	**.002**	-0.78
Veridical recalla	133.42 ± 19.11	137.86 ± 21.84	0.89	.842	0.22
Intrusions	1.03 ± 1.16	.37 ± .59	-2.97	**.004**	-0.72

Notes. Mean and standard deviation, as well as Cohen's *d* are reported. Significant p-values are in bold.

Table 3. Comparisons between IN group and GS group in DRM the recognition variables.

	IN group	GS group	*t*	*p*	*Cohen's d*
False recognitions	11.18 ± 3.37	10.22 ± 3.74	-1.13	.269	-0.27
Hits	36.78 ± 4.09	36.54 ± 5.81	-0.20	.842	-0.05
False alarms	4.87 ± 2.99	4.25 ± 4.67	-0.65	.528	-0.16
FR - R	0.64 ± 0.24	0.61 ± 0.25	-0.65	.520	-0.16
FR - K	0.33 ± 0.22	0.38 ± 0.26	1.09	.282	0.27
Hits –R	0.74 ± 0.19	0.76 ± 0.19	0.37	.710	0.09
Hits - K	0.22 ± 0.15	0.20 ± 0.14	-0.64	.528	-0.16
FA - R	0.58 ± 0.64	0.36 ± 0.27	1.81	.075	0.45
FA - K	0.61 ± 0.29	0.59 ± 0.78	-0.14	.889	-0.04

Notes. Mean, standard deviation and Cohen's *d* are reported. Significant p-values are in bold.

Executive functioning tasks

The IN group showed lower digit span scores (p = .029) and WMI than the GS group (p = .008). No other significant between-groups differences emerged at the Stroop task or source monitoring task (see **Table 4**).

Table 4. Between groups comparisons in WAIS's scores, Stroop's and source-monitoring task's variables.

	IN group	GS group	*t*	*p*	*Cohen's d*
Digit span scores	8.27 ± 2.98	9.77 ± 2.56	22.29	**.029**	0.54
Arithmetic scores	4.39 ± 2.83	5.46 ± 2.86	15.39	.128	0.37
WMI	78.73 ± 12.42	86.00 ± 11.84	27.24	**.008**	0.66
Stroop - *correct responses*	230.88 ± 21.40	236.86 ± 3.62	16.28	.108	0.39
Stroop - *errors*	8.52 ± 21.41	3.43 ± 2.00	-13.77	.173	-0.33
Stroop - *response times*	926.78 ± 197.03	907.36 ± 211.03	- 0.38	.701	-0.09
SM - *corrects*	35.73 ± 5.25	35.91 ± 4.55	0.16	.875	0.03
SM - Index 1	50.24 ± 10.75	46.62 ± 9.70	-14.59	.152	-0.04
SM - Index 2	49.76 ± 10.75	53.38 ± 9.70	14.59	.149	0.35

Notes. Mean and standard deviation, as well as Cohen's *d* are reported. Significant p-values are in bold. WMI = Working Memory Index; SM = Source Monitoring task.

Correlational analysis revealed several significant associations between DRM performance, executive functioning and source monitoring ability in the whole sample.

Specifically, the number of *false recalls* negatively correlated with WAIS digit span scores (r = - 0.26, p = .031) and WMI (r = - 0.28, p = .022). The number of *false recognitions* negatively correlated with WAIS digit span scores (r = - 0.39, p = .001), WAIS arithmetic scores (r = - 0.28, p =.023) and WMI (r = 0.43, p < .001); the same variable also showed a trend to a significant negative correlation with the number of correct responses at the source monitoring task (r = - 0.24, p = .051).

As for *veridical recalls*, we observed a negative correlation with reaction times at the Stroop task (r = - 0.29, p = .021) and a positive correlation with the number of correct responses at the source monitoring task (r = 0.31, p = .011); moreover, the same DRM variable was positively correlated with WAIS digit span scores (r = 0.29, p = .013) and WMI (r = - 0.32, p = .008). As for the recognition task, no significant correlations emerged between the number of *hits*, executive functioning and source monitoring variables.

The number of *recalled intrusions* positively correlated with Stroop reaction times (r = 0.29, p = .022) and negatively with the number of correct responses at the source monitoring task (r = - 0.32, p = .008). No significant correlations emerged between the number of *false alarms*, executive functioning and source monitoring variables.

Both the number of *false recalls* and of *recalled intrusions* showed a positive correlation with PSQI scores (*false recalls*: r = 0.29, p = .017; *recalled intrusions*: r = 0.23, p = .056) and ISI scores (*false recalls*: r = 0.29, p = .014; *recalled intrusions*: r = 0.27, p = .029).

2.4. Discussion

To the best of our knowledge, the present study is the first to address false memories production for semantically related words in individuals with insomnia symptoms, by using the classical Deese - Roediger – McDermott paradigm (DRM; Deese, 1959; Roediger and McDermott, 1995).

In line with our working hypothesis, the first main result is the higher production of false memories observed in the IN group compared to the GS group. Indeed, the IN group not only reported more *false recalls* but also more frequently recalled words that were unrelated to the critical *lure* words (i.e. *recalled intrusions*) compared to the GS group. This result is in accordance with previous studies reporting that poor sleep in people suffering from insomnia is accompanied by diurnal impairments in the same frontally-mediated cognitive functions that are essential to ensure an efficient memory retrieval process and to avoid false memories production (i.e. working memory, inhibitory control, and cognitive flexibility) (Ballesio et al., 2019; Wardle-Pinkston et al., 2019). Our correlational analysis also corroborates this result by showing that the degree of sleep impairment is associated with accuracy of memory retrieval. In fact, we observed a positive correlation of the ISI and PSQI scores with both the number of *false recalls* and *recalled intrusions*.

The difference between the IN and the GS group in the number of *false recalls* can be explained in light of the Activation-Monitoring Theory (AMT; Roediger & McDermott, 1995; Roediger et al., 2001). This theory argues that false memories arise from two main processes: the first, related to the activation of the critical *lure* word at memory encoding and/or retrieval (*Activation*); the second, linked to the subsequent failure in the source monitoring process at retrieval (*Monitoring*). This latter process brings subjects to mistakenly remember the critical *lure* word (internally generated) as having been presented, instead of merely cognitively activated. Given the above-mentioned deficits in frontally-mediated executive functioning documented in insomnia disorder, it is plausible that IN group subjects have been more susceptible than controls to this type of source monitoring error and, consequently, to produce false memories. In this regard, our correlation analysis sheds light on two relevant issues. First, it reveals a relationship between executive functioning and DRM performance, i.e., a negative correlation between WMI and the number of *false recalls*, as well as of *false recognitions*. This result is in agreement with previous studies (Peters et al., 2007; Leding, 2012) and suggests the importance of an efficient executive functioning for ensuring proper memory retrieval and avoiding false memories production. Second, our results highlight that source monitoring ability is related to DRM measures. Specifically, we observed that the number of correct responses at the source monitoring task positively correlates with the number of *veridical recalls* and negatively with the number of *false recognitions*. This result suggests that the more a subject is able to distinguish between memories for thoughts (internal) and perceived (external) events, the less he will tend to produce false memories at the DRM paradigm.

Our result on *recalled intrusions* is in line with previous studies conducted on people suffering from insomnia. In fact, deficits in performing declarative memory tasks including free recall (Bonnet & Arand, 1995) or recognition (Rosa & Bonnet, 2000) of unrelated *studied* words, as well as in a verbal fluency task (Mendelson et al., 1984), have already been observed in this population.

The difficulty in memory retrieval does not seem to have affected *veridical recall* in the IN group. We did not find between-group differences in the number of *studied* words correctly recalled. This could be the consequence of the semantic association between the words within each DRM list.

In this regard, previous studies argued that the strong semantic association between stimuli generally facilitates their recall (Aka et al., 2020; Silberman et al., 2005). In other words, the nature of the DRM stimuli may have allowed the participants with insomnia symptoms to recall the *studied* words as efficiently as good sleepers.

Contrary to the DRM free recall task, no between-groups differences emerged at the recognition task in the number of *false recognitions*, *hits*, and *false alarms*. This negative finding is probably related to the characteristics of the task, that strongly differ from those of the free recall procedure (Cabeza et al., 2001; Tulving & Madigan, 1970). Specifically, in the recognition task subjects are explicitly asked to discriminate between "old" and "new" words by the experimenter, who provides a set of external cues to choose from. Conversely, the free recall task does not provide any external memory cues at retrieval, therefore subjects have to generate their own cues to reinstate the original memory traces. This process of "self-cueing" is highly cognitively demanding and is also at risk of inducing memory errors. Therefore, it is plausible that between-group differences in the ability to correctly retrieve previously encoded information might be better detected through a more cognitively demanding task, such as the free recall.
To note that the proportion of *false recognitions*, *hits*, and *false alarms* in the IN group not only is comparable to the GS group, but also to that of previous research conducted on healthy samples (Gallo et al., 2010). Again, this data suggests that difficulties in memory retrieval at the DRM paradigm might more easily emerge with a more difficult task.

As for executive functioning, we observed lower digit span scores and a poorer working memory index in the IN group compared to the GS group. This result is in line with previous studies showing impairments in people suffering from insomnia in performing working memory tasks (for a review, see Ballesio et al., 2019). Instead, the two groups performed similarly at the Stroop task. In this regard, other studies observed comparable performance between individuals with insomnia and good sleepers when evaluating inhibitory control with the same task (Orff et al., 2007) and with a similar color-word interference test (Siversten et al., 2013). Nevertheless, we can explain this result in light of the two main limitations of our study. First, our sample included subjects reporting insomnia symptoms with no diagnosis of clinically relevant insomnia. Therefore, between-group differences in performing this classical neuropsychological task could emerge by recruiting individuals with the ascertained diagnosis. In addition, due to the Covid-19 pandemic, the task was administered online. Although we tried to ascertain the presence of an efficient internet connection and the workstation adequacy for each subject, the procedure of task administration might have undergone some bias.

As for the source monitoring task, no differences emerged between the IN group and the GS group. In our study, we investigated participants' ability to discriminate between information coming from an internal and an external source. In particular, we required them to indicate whether an item (presented during a *learning* phase) had been generated by themselves or by an external source. The task we developed focused on a specific type of source monitoring, defined "*Reality Monitoring*" (Johnson et al., 1993). However, two other source monitoring categories have been described in literature, the "*Internal*" and the "*External*" ones respectively (Johnson et al., 1993). The "*Internal*" source monitoring represents the ability to discriminate between two similar internally-derived sources of information, whereas the "*External*" one allows for discrimination between two externally-derived sources. According to Johnson and colleagues (1993), the comparison between stimuli coming from two qualitatively different sources (internal vs external), as in *Reality Monitoring*, might be easier than comparing information from similar sources (both internal or external). Therefore, we hypothesize that differences between IN group and GS group might emerge by adopting the more difficult *Internal* and *External* Source Monitoring tasks rather than the *Reality Monitoring*.

In conclusion, our results showed that the IN group falsely recalled more critical *lure* words and, in parallel, reported a lower working memory efficiency than the GS group. Together with correlational data highlighting a relationship between sleep quality, executive functioning and DRM performance, the results confirm our initial hypothesis that false memories formation can be a consequence of chronic poor sleep. Furthermore, the observed association of ISI and PSQI scores with *false recalls* suggests that the process of false memories production is modulated by the degree of sleep impairment.

References

Aka, A., Phan, T. D., & Kahana, M. J. (2020). Predicting recall of words and lists. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication. https://doi.org/10.1037/xlm0000964

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders, 5th ed. American Psychiatric Association.

Ballesio, A., Aquino, M., Kyle, S. D., Ferlazzo, F., & Lombardo, C. (2019). Executive Functions in Insomnia Disorder: A Systematic Review and Exploratory Meta-Analysis. Frontiers in psychology, 10, 101. https://doi.org/10.3389/fpsyg.2019.00101

Bonnet, M.H., & Arand, D.L. (1995). 24-hour metabolic rate in insomniacs and matched normal sleepers. Sleep, 18, 581-588.

Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R., & Schacter, D.L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. PNAS, 98, 4805–10.

Castronovo, V., Galbiati, A., Marelli, S., Brombin, C., Cugnata, F., Giarolli, L., Anelli, M. M., Rinaldi, F., & Ferini-Strambi, L. (2016). Validation study of the Italian version of the Insomnia Severity Index (ISI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 37(9), 1517–1524. https://doi.org/10.1007/s10072-016-2620-z

Curcio, G., Tempesta, D., Scarlata, S., Marzano, C., Moroni, F., Rossini, P. M., Ferrara, M., & De Gennaro, L. (2013). Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 34, 511–519. https://doi.org/10.1007/s10072-012-1085-y

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58, 17-22.

Denckla, M.B. (1994). Measurement of executive function. In G.R. Lyon (Ed.), Frames of reference for the assessment of learning disabilities: New views on measurement issues (pp. 117–142). Paul H. Brookes Publishing Co.

Durmer, J.S., & Dinges, D.F. (2005). Neurocognitive consequences of sleep deprivation. Seminars in Neurology, 25, 117–129.

Fortier-Brochu, E., Beaulieu-Bonneau, S., Ivers, H., & Morin, C.M. (2012). Insomnia and daytime cognitive performance: A meta-analysis. Sleep Medicine Reviews, 16, 83–94. https://doi.org/10.1016/j.smrv.2011.03.008

Frenda, S.J., & Fenn, K.M. (2016). Sleep less, think worse: The effect of sleep deprivation on working memory. Journal of Applied Research in Memory and Cognition, 5, 463–469. https://doi.org/10.1016/j.jarmac.2016.10.001

Gajewski, P. D., Falkenstein, M., Thönes, S., & Wascher, E. (2020). Stroop task performance across the lifespan: High cognitive reserve in older age is associated with enhanced proactive and reactive interference control. NeuroImage, 207, 116430. https://doi.org/10.1016/j.neuroimage.2019.116430

Gallo D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. Memory & cognition, 38(7), 833–848. https://doi.org/10.3758/MC.38.7.833

Iacullo, V.M., & Marucci, F.S. (2016). Normative data for Italian Deese/Roediger-McDermott lists. Behavior Research Methods, 48, 381–389. https://doi.org/10.3758/s1342 8-015-0582-3

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. Psychological Bulletin, 114, 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M.K., Raye, C.L., Mitchell, K.J., & Ankudowich, E. (2012). The cognitive neuroscience of true and false memories. Nebraska Symposium on Motivation. Nebraska Symposium on Motivation, 58, 15–52. https://doi.org/10.1007/978-1-4614-1195-6_2

Leding J. K. (2012). Working memory predicts the rejection of false memories. Memory (Hove, England), 20(3), 217–223. https://doi.org/10.1080/09658211.2011.653373

Mendelson WB, Garnett D, Gillin C, Weinberger H. (1984). The experience of insomnia and daytime and night-time functioning. Psychiatry Research. 12:235–50.

Mitchell, K.J., & Johnson, M.K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving, & F.I.M. Craik (Eds.), The Oxford handbook of memory (pp. 179–195). Oxford University Press.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. Cognitive Psychology, 41, 49–100. https://doi.org/10.1006/cogp.1999.0734

Natale, V., Esposito, M., Martoni, M., & Fabbri, M. (2006). Validity of the reduced version of the Morningness-Eveningness Questionnaire. Sleep Biological Rhythms, 4, 72–74. https://doi.org/10.1111/j.1479-8425.2006.00192.x

Nichols, R. M., & Loftus, E. F. (2019). Who is susceptible in three false memory tasks?. Memory, 27(7), 962–984. https://doi.org/10.1080/09658211.2019.1611862

Nienow, M.S., & Docherty, N. M. (2004). Internal Source Monitoring and Thought Disorder in Schizophrenia. The Journal of Nervous and Mental Disease, 192, 696–700.

Orff HJ, Drummond SPA, Nowakowski S, Perlis ML (2007). Discrepancy between subjective symptomatology and objective neuropsychological performance in insomnia. Sleep, 30, 1205–11.

Peters, M. J. V., Jelicic, M., Verbeek, H., & Merckelbach, H. L. G. J. (2007). Poor working memory predicts false memories. European Journal of Cognitive Psychology, 19, 213-232. https://doi.org/10.1080/09541440600760396

Riedel, B. W., & Lichstein, K. L. (2000). Insomnia and daytime functioning. Sleep medicine reviews, 4(3), 277–298. https://doi.org/10.1053/smrv.1999.0074

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. Psychonomic bulletin & review, 8(3), 385–407. https://doi.org/10.3758/bf03196177

Roediger, H.L, & McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. Journal of Experimental Psychology, 21, 803 – 14

Rosa, R.R., & Bonnet, M.H. (2000). Reported chronic insomnia is independent of poor sleep as measured by electroencephalography. Psychosomatic Medicine, 62, 474-82.

Sica, C. & Ghisi, M. (2007). The Italian versions of the Beck Anxiety Inventory and the Beck Depression Inventory-II: Psychometric properties and discriminant power. In M. A. Lange (Ed.), Leading-edge psychological tests and testing research (p. 27–50). Nova Science Publishers

Silberman, Y., Miikkulainen, R., & Bentin, S (2005). Associating unseen events: Semantically mediated formation of episodic associations. Psychological Science, 16 161– 166.

Sivertsen, B., Hysing, M., Wehling, E., Pallesen, S., Nordhus, I. H., Espeseth, T., & Lundervold, A. J. (2013). Neuropsychological performance in older insomniacs. Neuropsychology, development, and cognition. Section B, Aging, neuropsychology and cognition, 20(1), 34–48. https://doi.org/10.1080/13825585.2012.672947

Stroop, J. (1935). Studies of interference in serial verbal learning reactions. Journal of Experimental Psychology, 18, 643–662.

Tulving, E., & Madigan, S.A. (1970). Memory and verbal learning. Annual Review of Psychology, 21, 437–84.

Violani, C., Devoto, A., Lucidi, F., Lombardo, C., Russo, P.M. (2004). Validity of a short insomnia questionnaire: the SDQ. Brain Research Bulletin, 63, 5, 415-421

Wardle-Pinkston, S., Slavish, D. C., & Taylor, D. J. (2019). Insomnia and cognitive performance: A systematic review and meta-analysis. Sleep medicine reviews, 48, 101205. https://doi.org/10.1016/j.smrv.2019.07.008

Wechsler, D. (2008). Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV). NCS Pearson.

Appendix

LIST 1	LIST 2	LIST 3	LIST 4	LIST 5	LIST 6
Alto	**Bandiera**	**Dolce**	**Fiume**	**Ladro**	**Sonno**
Basso	Italia	Caramella	Foce	Furfante	Riposo
Magro	Rossa	Amaro	Affluente	Truffatore	stanchezza
Palazzo	Tricolore	Pasticcino	Scorrere	Rubare	Notte
Adige	Vento	Tenero	Diga	Rapinatore	Sogno
Dritto	Nazione	Zucchero	Pescare	Scippo	Profondo
Fragile	Bianca	Miele	Ponte	Delinquente	Sbadiglio
Medio	Alza	Goloso	Torrente	Scassinare	Tranquillità
Gigante	Asta	Crostata	Corrente	Bandito	Sera
Muro	Colore	Acre	Flusso	Rapina	Rilassante
Grattacielo	Italiana	Sapore	Pesce	Assassino	Buio
Scaffale	Strisce	Torta	Arno	Prigione	Occhi
Metro	Cuore	Crema	Lungo	Sbirro	Pesante
Grado	Mondiali	Panna	Valle	Pistola	Fatica
Palo	Innalzare	Frutto	Detriti	Carcere	Incubo
Torre	Inno	Salato	Tevere	Crimine	Leggero

LIST 7	LIST 8	LIST 9	LIST 10	LIST 11	LIST 12
Lampada	**Musica**	**Sedia**	**Letto**	**Freddo**	**Re**
Luce	Concerto	Tavolo	Dormire	Ghiaccio	Potere
Lume	Pentagramma	Legno	cuscino	Guanti	Regina
Sfregare	Radio	Dondolo	coperta	Vento	Corona
Accesa	Ritmo	Gambe	Pigiama	Coperta	Cavallo
Spina	Melodia	Rotelle	comodità	Metallico	Monarca
Voce	Sassofono	Sedere	Amore	Montagna	Cuori
Volta	Strumento	Paglia	Caldo	Neve	Sudditi
Desiderio	Nota	Poltrona	Relax	Pioggia	Monarchia
Ragione	Sinfonia	Riposo	Piumone	Monte	Sole
Calore	Arpa	Scomoda	camera	Tiepido	Trono
Alogena	Disco	Sdraio	Lenzuola	Coperto	Nudo
Lampadina	Pianoforte	Sgabello	Casa	Maglione	Elogio
Aladino	Suono	Studio	Morbido	Naso	Impero
Neon	Chitarra	Comoda	Libro	Corrente	Matto
Genio	Orchestra	Cucina	Materasso	Sciare	Oro

LIST 13	LIST 14	LIST 15	LIST 16
Ago	**Uomo**	**Penna**	**Fumo**
Filo	Donna	Inchiostro	Pipa
Puntura	Barba	Scrivere	Sigaro
Pagliaio	Forte	Matita	Sigaretta
Punta	Genere	Foglio	Tabacco
Cucito	Persona	Biro	Cenere
Bilancia	Adamo	Calamaio	Fumare
Cucire	Onore	Nera	Nicotina
Dolore	Bello	Piuma	Polmoni
Abito	Capelli	Scrittura	Camino
Cammello	Forza	Scuola	Inquinamento
Siringa	Virile	Sfera	Cancro
Ditale	Cravatta	Stilo	Vapore
Iniezione	Bambino	Astuccio	Puzza
Pungere	Grande	Grafia	Accendere
Sarta	Maschile	Oca	Arrabbiato

Chapter 3

False memories formation is increased in individuals with insomnia[3]

3.1. Introduction

Previous studies report that false memories can be influenced by sleep (see for a review, Conte & Ficca, 2013; Landmann et al., 2014). Among the first to investigate the relationship between sleep and false memories, Diekelmann and colleagues (2008; 2010) showed that sleep deprived individuals produce a higher number of false memories at morning retest compared to participants in an undisturbed sleep condition. The authors specified that this effect could be mainly linked to an impaired memory retrieval process. In fact, acute sleep loss can affect several cognitive functions related to prefrontal activity that are essential to accurate recall from long term memory (Durmer & Dinges, 2005; Frenda et al., 2016).

Most of the available studies on sleep and false memories have been conducted on healthy subjects exposed to experimental sleep deprivation (see Diekelmann et al., 2008, 2010) or restriction (e.g., Lo et al., 2016), while clinical samples have been neglected. Since there is growing acceptance that the nature and severity of the cognitive consequences of these experimental sleep interventions differ from those reported in chronic sleep disorders (Shekleton et al., 2010), we intended to assess the effect of chronically disturbed sleep on false memories production.

Insomnia is a sleep disorder characterized by subjective complaints of non-restorative sleep and of difficulties in initiating and/or maintaining sleep, accompanied by decreased daytime functioning, which persist in time (American Psychiatric Association, 2013). Objective sleep impairments are also often reported in this population, such as changes in sleep architecture (i.e. reduction in slow wave sleep and REM sleep duration) compared to healthy subjects (Baglioni et al., 2014). Moreover, individuals suffering from insomnia show several cognitive impairments that could contribute to false memories generation. It has been observed that they perform more poorly than good sleepers on complex cognitive tasks depending on the efficiency of the prefrontal cortex, for example in tests assessing working memory (e.g., retention and manipulation

[3] Malloggi, S., Conte, F., De Rosa, O., Righi, S., Gronchi, G., Ficca, G., & Giganti, F. (2021). False memories formation is increased in individuals with insomnia. Journal of Sleep Research, 00, e13527. https://doi.org/10.1111/jsr.13527

of previously acquired information), problem solving, information processing and selective attention (for a review, see Fortier-Brochu et al., 2012).

Other factors are likely to modulate the effects of disordered sleep on false memories production. For instance, an important one could be the nature of the administered stimuli. Indeed, several studies report that individuals with insomnia preferentially focus their attention on stimuli that are related to sleep, which appear to them more salient than neutral ones (Espie et al., 2006; Harvey, 2002). This phenomenon is known as "*attentional bias*" and has been previously observed in this population through specific cognitive tasks, such as the Stroop task (Spiegelhalder et al., 2008; Zhou et al., 2018), the dot probe task (MacMahon et al., 2006), priming tasks (Giganti et al., 2017) and eye-tracking paradigms (Woods et al., 2013). Overall, these studies suggest that sleep-related stimuli induce a higher activation in individuals with insomnia relative to good sleepers, leading the former to respond differently to these stimuli. For instance, it has been observed that individuals with insomnia, compared to good sleepers, show slower reaction times for sleep-related stimuli at the Stroop task (Spiegelhalder et al., 2008; Zhou et al., 2018) and that they recognize these stimuli at lower spatial frequencies in a priming task (Giganti et al., 2017). Instead, the effect of stimulus category has not yet been investigated in tasks based on semantically associated items, such as the Deese-Roediger-McDermott paradigm (DRM; Roediger & McDermott, 1995).

A complementary hypothesis has been sometimes put forward (e.g., Williams et al., 1996), that the attentional bias may reflect the way in which "experts" react to their expertise-related stimuli when performing specific tasks. A few studies actually show that experts generally produce higher rates of false memories for words that are related to the domain of their expertise compared to non-experts (Baird, 2003; Castel et al., 2007). This finding is attributed to the stronger semantic activation occurring in experts: in the case of DRM word lists, expertise would increase the number and strength of associations between expertise-related terms and enhance the spreading activation to include the non-presented critical words.

In light of this literature, the investigation of the effects of stimulus category in a DRM task in individuals with insomnia, who may be considered "experts" and strongly activated by the theme of sleep (Espie et al., 2006; Harvey, 2002), appears particularly interesting,

The first aim of this study is to assess whether chronic poor sleep quality in subjects with insomnia affects false memories production. To this end, we compare performance at the DRM paradigm (Roediger & McDermott, 1995) between a group of individuals showing insomnia symptoms and one of good sleepers. In light of literature on the attentional bias described in people suffering from insomnia (Giganti et al., 2017; Harris et al., 2015), we also evaluate the possible effect of stimulus category by comparing performance at neutral and sleep-related word lists included in the DRM task. Finally, considering the association between false memories production and executive functioning (e.g., Peters et al., 2007; Leding, 2012), as well as the observed impairments of these functions in insomnia (see for example: Haimov et al., 2008; Joo et al., 2013), we assess in both groups working memory, inhibitory control and source monitoring ability, the latter being considered as especially linked to false memories formation (Mitchell & Johnson, 2000).

3.2. Materials and method

3.2.1 Participants

Eighty potential participants were approached at university sites (i.e. lecture halls, library, etc.) and asked to fill out a set of screening questionnaires: the Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), the Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016), the Sleep Disorder Questionnaire (SDQ; Violani et al., 2004), the Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007), and the Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007) (described in details below). In addition, they were administered a semi-structured interview at the sleep laboratory, conducted by a licensed psychologist who had received specific training, in order to assess general medical condition and health habits, presence of psychiatric disorders and of sleep disorders. The presence of clinical insomnia was specifically addressed by means of the semi-structured interview (Morin, 1993).

Based on scores at the screening instruments and on the interview, 53 university students were recruited for the study and included in either the “good sleep group” (**GS group**, n = 28) or the “insomnia group” (**IN group**, n = 25). Inclusion criteria common to both groups were: absence of any relevant somatic or psychiatric disorder; absence of clinically significant depression and anxiety symptoms (BDI-II score ≤ 29; BAI score ≤ 25); no history of drug or alcohol abuse; absence of sleep disorders (other than insomnia for the IN group) and of any sleep apnea or respiratory disorder symptom; having a regular sleep-wake pattern (e.g., individuals with irregular study or working habits such as shift-working were excluded); no use of psychoactive medication or alcohol at bedtime. In addition, for inclusion in the IN group, participants had to score ≥ 5 at the PSQI, ≥ 8 at the ISI and to be classified as presenting “clinically significant insomnia” at the SDQ; further, they had to fully meet DSM-5 criteria for Insomnia Disorder, as verified through the interview.

Finally, inclusion in the GS group was based on: PSQI score < 5, ISI score < 8, being classified as "good sleeper" at the SDQ, and absence of any sleep disorder as also verified through the interview.

The two groups did not differ in terms of age, gender distribution, circadian preference (measured through the reduced version of the Morningness-Eveningness Questionnaire; Italian version from Natale et al., 2006) and daytime sleepiness (measured through the Epworth Sleepiness Scale; Italian version from Vignatelli et al., 2003). Instead, as expected, significant between-groups differences emerged in several habitual sleep features assessed through the PSQI, such as bedtime, sleep duration and sleep onset latency, as well as in PSQI global scores (**Table 1**).

Table 1. Age, gender distribution, circadian preference, daytime sleepiness, habitual sleep features and sleep quality in the IN and GS groups.

	IN group	**GS group**	**Statistical test**
Age	25.16 ± 4.34	24.10 ± 3.17	U = 278.50, p = .190
Gender	13 M, 12 F	11 M, 17 F	χ^2 = 0.86, p = .862
MEQr score	13.00 ± 2.61	14.54 ± 3.18	U = 245.00, p = .060
ESS score	7.64 ± 2.82	6.32 ± 3.28	U = 253.00, p = .110
Habitual Bedtime *(hh:mm)*	00:32 ± 00:52	23:32 ± 00:59	U = 91.50, p = .001
Habitual Rise time *(hh:mm)*	08:08 ± 01:04	07:32 ± 01:23	U = 182.50, p = .233
Habitual Sleep duration *(hh:mm)*	06:27 ± 00:59	07:41 ± 00:48	U = 78.50, p < .001
Habitual Sleep onset latency *(hh:mm)*	00:25 ± 00:13	00:11 ± 00:05	U = 116.00, p <. 001
PSQI global score	8.16 ± 2.26	3.21 ± 0.87	U = 0.00, p < .001

Notes. MEQr: Morningness-Eveningness Questionnaire (reduced version); ESS: Epworth Sleepiness Scale; IN group: insomnia group; GS group: good sleep group; PSQI: Pittsburgh Sleep Quality Index; M: Males; F: Females. Habitual bedtime, rise time, sleep duration and Sleep onset latency were collected through the PSQI. Mann Whitney's U is reported for between-groups comparisons for all variables except gender. Results of the chi squared test are reported for differences in gender distribution

Participants were requested to complete a sleep diary on the day of the DRM's testing session, in order to control that they performed the task after a night of sleep that was representative of their habitual sleep. In **Table 2** we report the sleep measures of the night before the administration of the memory task in both groups.

Table 2. Sleep features of the night preceding the DRM task session in the IN and GS group.

	IN group	GS group	Statistical test
Bedtime *(hh:mm)*	00:42 ± 00:47	23:46 ± 00:55	$U = 85.50$, $p = .003$
Rise time *(hh:mm)*	07:50 ± 01:01	07:48 ± 00:51	$U = 219.50$, $p = .980$
Sleep duration *(hh:mm)*	06:47 ± 01:01	07:52 ± 01:06	$U = 114.0$, $p = .004$
Sleep onset latency*	3 (i.e., *"≥ 15 min"*)	2 (i.e., "*10 min*")	$U = 202.00$, $p = .010$
Number of Awakenings	1.2 ± 1.18	0.53 ± 0.83	$U = 235.50$, $p = .025$
Rise time latency*	2 (i.e., "*10 min*")	1 (i.e., "*5 min*")	$U = 206.00$, $p = .019$

Notes. IN group: insomnia group; GS group: good sleep group. Sleep features were collected through sleep logs. Mann Whitney's *U* is reported for between-groups comparisons for all variables. An asterisk (*) indicates median values. Prior night's sleep onset latency was obtained through the question: "How long did it take you to fall asleep last night?" ("< 5 minutes", "5 minutes", "10 minutes", "≥ 15 minutes"). Rise time latency was obtained through the question: "How long did it take you to rise from bed after this morning's awakening?" ("<5 minutes", "5 minutes", "10 minutes", "≥15 minutes").

3.2.2. Procedure

All selected participants were individually invited to the sleep laboratory, where they were administered the DRM paradigm (Roediger & McDermott, 1995).

On separate days, a subsample of 17 participants from the IN group (8 M, 9 F; mean age = 24.5 ± 2.2) and 21 from the GS group (8 M, 13 F; mean age = 24.1 ± 2.2) were again invited individually to the sleep lab where they were administered a set of cognitive tests to evaluate executive functioning and source monitoring ability. **Table 3** displays demographic characteristics, circadian preference, daytime sleepiness and habitual sleep features of the subsample. All testing sessions (both the DRM and the executive functioning tasks) were performed in the morning, between 11:00 and 13:00, by an experimenter who was blind to the study group.

There was no money or credit compensation for participating in the study.

The study design was submitted to the Ethical Committee of the Department of Psychology, University of Campania "L. Vanvitelli", which approved the research (code 22/2020) and certified that the involvement of human participants was performed according to acceptable standards.

Table 3. Age, gender distribution, circadian preference, daytime sleepiness, habitual sleep features and sleep quality in participants of the subsample.

	IN group	GS group	Statistical test
Age	24.53 ± 2.18	24.10 ± 3.56	U = 143.00, p = .310
Gender	8 M, 9 F	8 M, 13 F	χ^2 = .310, p = .578
MEQr score	12.82 ± 2.51	14.52 ± 3.61	U = 127.50, p = 1.360
ESS score	8.35 ± 3.58	6.57 ± 3.52	U = 119.00, p = .080
Habitual Bedtime *(hh:mm)*	00:36 ± 00:49	23:21 ± 00:40	U = 27.50, p <.001
Habitual Rise time *(hh:mm)*	07:51 ± 01:14	07:23 ± 00:51	U = 87.00, p = .305
Habitual Sleep duration *(hh:mm)*	06:22 ± 00:59	07:39 ± 00:48	U = 33.50, p = .001
Habitual Sleep onset latency *(hh:mm)*	00:27 ± 00:14	00:10 ± 00:04	U = 41.00, p < .001
PSQI global score	8.35 ± 3.58	3.19 ± 0.98	U = 0.000, p = <.001

Notes. MEQr: Morningness-Eveningness Questionnaire (reduced version); ESS: Epworth Sleepiness Scale; IN group: insomnia group; GS group: good sleep group; PSQI: Pittsburgh Sleep Quality Index; M: Males; F: Females. Habitual bedtime, rise time, sleep duration and Sleep onset latency were collected through the PSQI. Mann Whitney's *U* is reported for between-groups comparisons for all variables except gender. Results of the chi squared test are reported for differences in gender distribution.

3.2.3. Instruments

Screening instruments

(1) Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), a self-report questionnaire evaluating subjective sleep quality in the past month. It is composed of 19-items grouped into 7 subscales: Subjective Sleep Quality, Sleep Latency, Sleep Duration, Habitual Sleep Efficiency, Sleep Disturbances, Use of Sleep Medication and Daytime Dysfunctions due to sleepiness. The PSQI total score ranges from 0 to 21, with higher scores indicating sleep difficulties and lower sleep quality. The cut-off score ≥ 5 is adopted to discriminate between good and bad sleepers.

(2) Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016), assessing the severity of insomnia symptoms during the previous 2 weeks. Based on the scores, subjects are classified into four categories: (a) no clinically significant insomnia (score 0 – 7); (b) subthreshold insomnia (8 – 14); (c) clinical insomnia – moderate severity (score 15 – 21); (d) clinical insomnia – severe (22 – 28).

(3) Sleep Disorder Questionnaire (SDQ; Violani et al., 2004) is a self-rating questionnaire with 27 items evaluating the presence of different sleep problems in the last month. The first three questions concern symptoms of insomnia, while the others investigate the presence of excessive sleepiness, sleep apnea, parasomnias and snoring. A subsequent set of questions investigates the duration, frequency and consequences of the sleep problem, and is used for the evaluation of the severity of the sleep disturbances reported. The SDQ permits the classification of subjects into three main categories: subjects who do not complain of any sleep disorder; subjects who report the occurrence of subthreshold insomnia and subjects with clinically significant insomnia.

(4) Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007), assessing the severity of depressive symptoms. It consists in 21 items and the total score ranges from 0 to 63, with higher scores indicating more severe depressive symptoms. Particularly, scores 0 – 13 represent minimal depression, scores 14 – 19 mild depression, scores 20 – 28 moderate depression, and scores 29 – 63 severe depression symptoms.

(5) Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007), a self-report instrument assessing the presence and severity of anxiety symptoms in the past week. It is made up on 21 items measuring the intensity of common somatic and cognitive symptoms of anxiety through a Likert scale ranging from 0 (*Not at all*) to 3 (*Severely – it bothered me a lot*). The score range is 0 – 63, with higher scores indicating more severe anxiety symptoms: specifically, a total score of 0 – 7 is considered to index minimal severity, 8 – 15 mild, 16 – 25 moderate and 26 – 63 severe.

False memories task

In the classical DRM paradigm (Roediger & McDermott, 1995), an immediate free recall test is administered on a list of words that are semantically associated to an unstudied critical *lure* word (e.g., “ink”, “paper”, “school”, all related to “pen”). This task reliably produces high rates of false memories for unstudied critical lures (Roediger & McDermott, 1995).

In order to highlight the possible effect of stimulus category, here we adopted, as in Baird (2003), a reduced version of the DRM paradigm (Roediger & McDermott, 1995) consisting in the presentation of four word lists made up of 15 words each (see **Appendix**). Indeed, as observed in a recent meta-analysis (Newbury & Monaghan, 2019), the length of the lists rather than their number appears to significantly affect false recall rates (with longer lists producing greater false recall).

The sleep-related list used in this study (i.e., the one corresponding to the unpresented *lure* "sleep") was created ad hoc in Italian following the method used by Iacullo and Marucci (2016), due to the absence of any such list standardized in Italian. Thus created, the list was preliminarily presented to 30 university students (21 F, 9 M; mean age = 24.10 ± 4.06), not enrolled in the present study, to assess the false recall rate for it (which was 27%). Then, we selected from Iacullo and Marucci's set (2016) the three lists showing the most similar false recall rates. The selected lists, corresponding to the lures "flag", "pen", and "river", all showed a false recall rate of 21%. Furthermore, in order to test possible differences between our sleep-related list and the neutral one, 25 individuals (16 F, 9 M, mean age = 25.88 ± 3.23), who were not enrolled in the main study, were asked to rate on a 1-5 Likert scale the sleep-relatedness, familiarity, activation and valence of each word belonging to the two lists, as well as their respective critical lures. Comparisons between the two lists revealed no significant difference for familiarity ($t = -1.00$, $p = .33$), activation ($t = 0.57$, $p = .58$) or valence (which was judged as neutral for both lists; $t = -1.31$, $p = .20$), whereas a significant difference emerged for sleep-relatedness ($t = -17.6$, $p < .001$).

As in Roediger and McDermott (1995) and Iacullo and Marucci (2016), the words in each list were presented in order of associative strength with the unpresented *lure* (from strongest to weakest).

As for task administration, the experimenter read the lists aloud with an interval of 20 seconds between lists. Participants were instructed to memorize the words as accurately as possible and were informed that they would be tested on them later. The "flag" and "river" lists (List 1 and 4, respectively), here used to control for primacy and recency effects as in Baird (2003), were presented to all participants as the first and last list of the set, respectively. The order of presentation of the "pen" and "sleep" lists (List 2 and 3, respectively), instead, was balanced between subjects (Baird, 2003).

After the "river" list was presented, participants performed the free recall test. Specifically, they were requested to write down on a blank piece of paper as many words as possible from all the presented lists. Five minutes were allotted for recall. In order to hold recall time constant between subjects, participants were instructed to continue thinking about the words for the whole allotted time.

Executive functioning tasks

For the assessment of executive functioning we employed classical tasks which measure the main executive components (e.g., Denckla, 1994; Myake, 2000): specifically, working memory was evaluated through the Working memory subtests of the WAIS-IV (Wechsler, 2008) and inhibitory control was tested through the Stroop task (Stroop, 1935). In addition, we created an ad hoc task aimed to evaluate source monitoring ability, which is deemed to be specifically linked to false memories formation (Mitchell & Johnson, 2000). Below, a brief description of the tasks:

1) Working memory subtests of the WAIS-IV (Wechsler, 2008), including the Digit Span subtest (made up of three increasingly difficult tasks: digit span forwards, backwards and sequencing) and the Arithmetic subtest (requiring to perform mental arithmetic problems): taken together, performance at these tasks provides the Working Memory Index (WMI), a global measure of the ability to attend to information presented verbally, manipulate it in short-term memory and then formulate a response. The tests were administered according to the standard procedure reported in the WAIS-IV manual.

2) Stroop Color and Word Test (Stroop, 1935): here we adopted a computerized version of the task developed on the *Open Sesame* software (version 3.3.8). The stimuli consisted of the words "red," "green," "yellow" and "blue" presented at the center of a black computer screen in one of the four colors. The color of the word displayed corresponded to its meaning in 50% of the trials (congruent condition), whereas in the remaining 50% of the trials word color and meaning were different (incongruent condition). Subjects had to indicate, as soon as possible, the color of the text by pressing a key on the keyboard corresponding to the effective color of the text. Subjects performed a short training phase consisting of 24 trials in order to familiarize with the task and afterwards they performed the task including 240 trials.

3) Source Monitoring task: a computerized Source Monitoring Task (see **Supplementary Material**) was included to evaluate the ability to discriminate between different sources of information. We developed this task from Nienhow and Docherty (2014), who originally evaluated internal source monitoring ability, that is the ability to discriminate between two internal sources of information. According to the classification of source monitoring's types proposed by Johnson and colleagues (1993), we extended the original task in order to test External Source Monitoring (i.e. the ability to discriminate between two externally derived sources of information) and Reality Monitoring ability (i.e. the ability to discriminate between internal and external information sources). Therefore, our task included three different sub-tests: *Internal Source Monitoring* (I-SM), *External Source Monitoring* (E-SM) and *Reality Monitoring* (RM-SM). Tasks presentation was counterbalanced between subjects.

3.2.4. Data analysis

Outcome measures of the DRM task were: number of *false recalls*, i.e. total number of falsely recalled critical *lure* words; number of *veridical recalls*, corresponding to the total number of words correctly recalled from the original word lists; number of *intrusions*, representing the total number of recalled words not corresponding to studied items or to the critical *lure* words. Only performance at the 2 experimental lists ("pen" and "sleep" lists) were included in data analysis, since the first and last lists ("flag" and "river" lists) were used to control for primacy and recency effects (as in Baird, 2003). *Veridical recalls* were significantly more numerous for Lists 1 + 4 compared to Lists 2 + 3 (Lists 1 + 4: 11.64 ± 2.58 vs. Lists 2 + 3: 9.98 ± 3.18; Wilcoxon's Z: 878.00; $p<.001$; *ES*: .624), confirming the presence of primacy and recency effects.

Concerning executive functioning, outcome measures were: digit span scores, arithmetic scores and the working memory index obtained from the WAIS-IV subtests, as well as number of correct responses, number of errors and response times (ms) for the Stroop task. Finally, **Table 4** displays outcome variables considered for the Source Monitoring task.

Table 4. Outcome variables of the Source Monitoring task.

	Source Monitoring task	
Subtest	**Variables**	**Description**
I - SM	**I - correct**	the number of words correctly attributed to the internal sources of information
	I - Index 1	the proportion of words correctly identified as "said" out of the total number of words correctly recognized as "old"
	I - Index 2	the proportion of words correctly identified as "thought" out of the total number of words correctly recognized as "old"
E - SM	**E - correct**	the number of words correctly attributed to the external sources of information
	E - Index 1	the proportion of words correctly identified as from "man" source out of the total number of words correctly recognized as "old"
	E - Index 2	the proportion of words correctly identified as from "women" source out of the total number of words correctly recognized as "old"
RM - SM	**RM - correct**	the number of words that were correctly attributed to the internal and the external sources of information
	RM - Index 1	the proportion of words correctly identified as from internal source out of the total number of words correctly recognized as "old"
	RM - Index 2	the proportion of words correctly identified as from external source out of the total number of words correctly recognized as "old"

Notes. I-SM: Internal Source Monitoring sub-task; E-SM: External Source Monitoring sub-task; RM-SM: Reality Monitoring sub-task.

Due to non-normal distribution of the data, we employed non parametric statistics. Cardinal variables were compared between IN and GS groups through the Mann-Whitney *U*-test. Within-subject comparisons were performed through the Wilcoxon Signed-Rank test. Finally, Spearman and point-biserial correlation analysis were performed to test the association memory performance, cognitive testing and source monitoring ability in the whole sample. Spearman correlation was also performed to assess associations between sleep features of the night before the DRM session and DRM performance in the whole sample. Significance level was set at $p \leq .05$.

In order to test the interaction between groups and stimulus type, we analyzed the data with a mixed model logistic regression using the statistical software R (version 4.0.3) and the package "lme4". In this analysis, we considered the total number of false memories as dependent variable, the group (GS group and IN group) as fixed effect. The random effects were the type of list and participant ID.

We also performed a mediation analysis (using PROCESS macro; SPSS version 27; Hayes, 2018) to test the role of source monitoring ability as mediator of the relationship between sleep quality (i.e., IN group and GS group) and the total number of false memories produced. We considered Group as independent variable and the total number of false memories as dependent variable. The considered mediator was the source monitoring ability, calculated by summing up all the correct responses to the three Source Monitoring subtests (i.e. internal, external source monitoring and reality monitoring). We calculated the indirect effect of "*Group*" on false memories production, through source monitoring ability, quantified as the product of the ordinary least squares (OLS) regression coefficient estimating source monitoring ability from "*Group*" and the OLS regression coefficient estimating false memories production from source monitoring ability controlling for "*Group*". A bootstrapping procedure (with 5000 bootstrap samples) to estimate 95% confidence intervals was used. According to Preacher and Hayes (2008), a 95% CI that does not include zero provides evidence of a significant indirect effect.

An a-priori power analysis was conducted. Taking into account the sample size of the study and an alpha-level of .05, a power analysis based on Mann-Whitney *U*-test testified that we were able to detect an effect size equal to $p = .717$ (i.e., P represents the effect size index (Trumble et al., 2020), in particular, P ($X < Y$) where X represents random draws from the first probability distribution and Y represents random draws from the other distribution) with a power equal to .80.

3.3. Results

False Memories task

The IN group globally produced more *false recalls* (*U* = 247.00; *p* = .04; *ES* = - 0.27) than the GS group (**Figure 1**). No differences emerged between groups in the total number of *veridical recalls* (IN group = 9.44 ± 3.33 vs GS group = 10.36 ± 2.93; *U* = 284.00, *p* = .24) or in the number of *intrusions* (IN group = 1.24 ± 1.09 vs GS group = 1.18 ± 1.02; *U* = 340.50, *p* = .86).

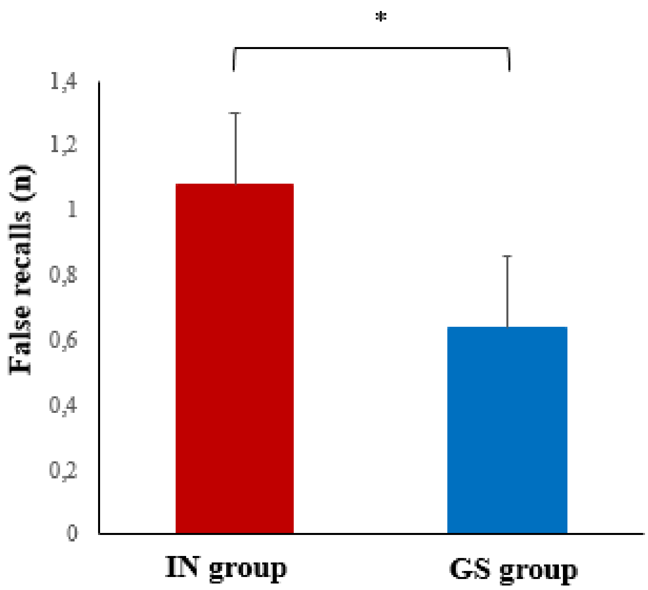

Figure 1. Comparison between IN group and GS group in the total number of *false recalls*. **p* ≤ .05. Error bars represent standard deviations

Moreover, IN group generated more *false recalls* (*U* = 254.00; *p* = .04; *ES* = - 0.28; **Figure 2**) and less *veridical recalls* (*U* = 242.500; *p* = .05; *ES* = - 0.27; **Figure 3**) at the sleep-related list compared to the GS group. Instead, the two groups did not differ neither in the number of *false recalls* (IN group = 0.52 ± 0.51 vs GS group = 0.36 ± 0.48, *U* = 239.00, *p* = .24) nor of veridical *recalls* (IN group = 5.52 ± 2.48 vs GS group = 5.50 ± 1.67, *U* = 349.5, *p* = .99) at the neutral list.

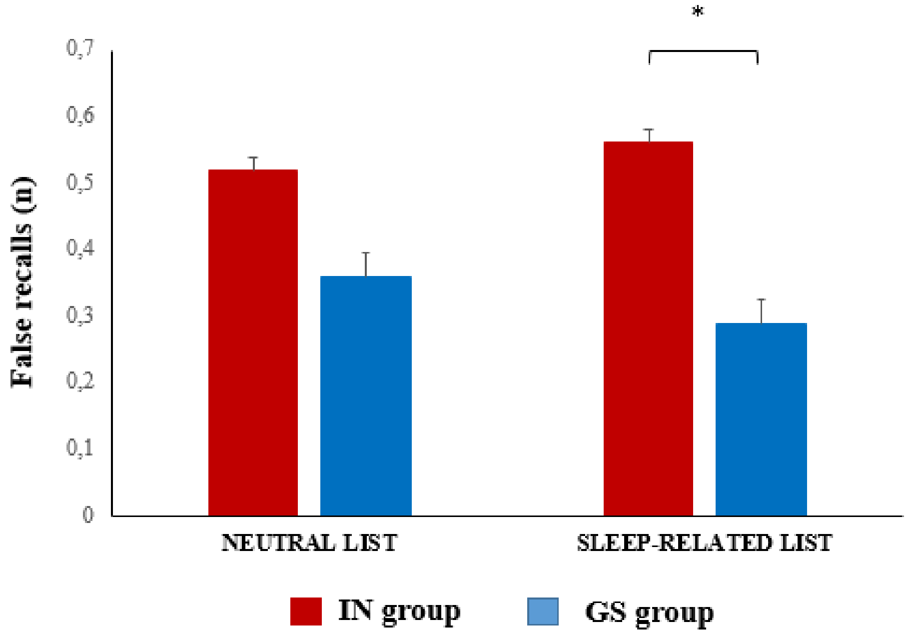

Figure 2. Comparison between IN group and GS group in the number of *false recalls* for neutral list and sleep-related list. *p ≤ .05. Error bars represent standard deviations.

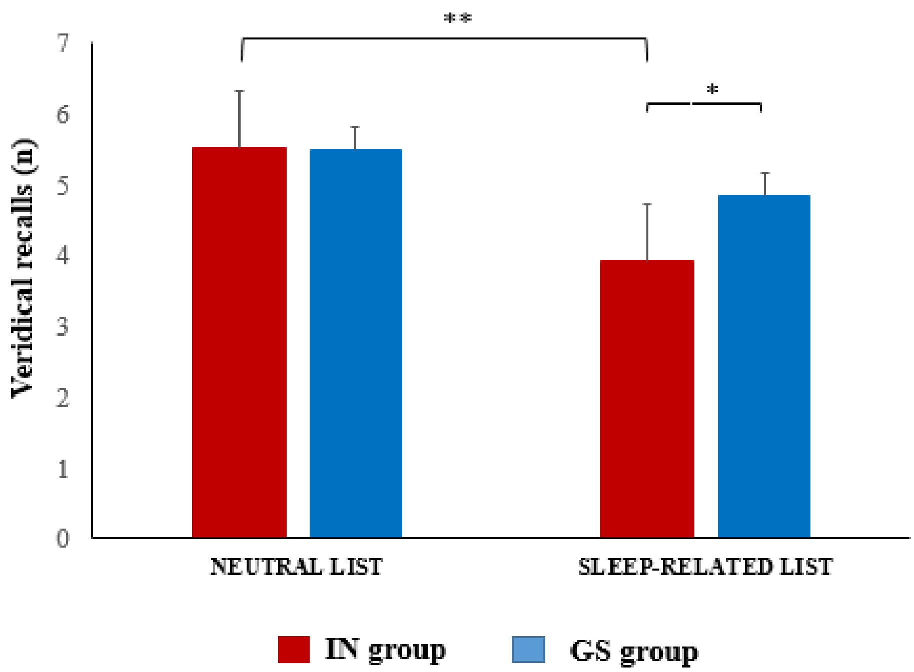

Figure 3. Number of *veridical recalls* for neutral list and sleep-related list in IN group and GS group. *$p \leq .05$; ** $p \leq .01$. Error bars represent standard deviations

As for within-subject comparisons, IN group produced less *veridical recalls* for the sleep-related list compared to the neutral list (Z = -2.587, p = .01; *ES* = -0.52; **Figure 3**), while no differences between lists emerged in the number of *false recalls* (Z = -0.378, p = .71; **Figure 2**). GS group did not show differences in the number of *false recalls* (Z = -0.471, p = .63; **Figure 2**) or *veridical recalls* (Z = -1.81, p = .10; **Figure 3**). As for the linear regression model, we observed a significant main effect of *Group* ($F_{1,51}$ = 5.00, p = .03), whereas no significant main effect of *list type* ($F_{1,51}$ = 0.04, p = .84) nor interaction effect ($F_{1,51}$ = 0.37, p = .55) emerged.

Executive functioning tasks

No between-groups differences emerged at the working memory and Stroop tasks (see **Table 5**).

Table 5. Comparison between IN group and GS group in Stroop task and WAIS-IV performance.

Tasks	*Variables*	**IN group**	**GS group**	*U*	*p*
Stroop	**Stroop – correct responses**	230.35 ± 29.54	232.20 ± 25.52	306.00	.84
	Stroop - errors	9.65 ± 29.53	7.80 ± 25.52	363.00	.85
	Stroop - response times	872.78 ± 163.26	851.07 ± 177.47	364.00	.70
WAIS	**Digit span subtest**	9.71 ± 3.02	9.57 ± 2.38	322.50	.87
	Arithmetic subtest	5.00 ± 3.06	5.29 ± 3.28	324.00	.78
	WMI	84.82 ± 13.72	85.38 ± 13.09	325.50	.90

Notes. Mean and standard deviation are reported. WMI = Working Memory Index

As for source monitoring ability, the IN group showed lower scores in Index 1 at the reality monitoring subtest compared to the GS group (*ES*= -0.29; **Table 6**), suggesting difficulties in correctly discriminating between internal and external sources of information. No other between-groups differences were observed.

Table 6. Comparison between IN group and GS group in Source Monitoring task.

Subtest	*Variables*	**IN group**	**GS group**	***U***	***p***
I - SM	**I - correct**	18.64 ± 3.84	19.19 ± 2.73	166.00	.61
	I - Index 1	56.18 ± 5.87	55.87 ± 8.07	177.00	.83
	I - Index 2	43.82 ± 5.87	44.13 ± 8.07	177.00	.82
E - SM	**E - correct**	18.11 ± 3.99	18.38 ± 2.94	169.50	.82
	E - Index 1	55.10 ± 14.23	56.14 ±13.56	167.00	.84
	E - Index 2	44.91 ± 14.24	43.85 ± 13.56	167.00	.88
RM - SM	**RM - correct**	20.52 ± 2.34	19.00 ± 3.11	125.50	.43
	RM - Index 1	47.33 ± 6.22	54.16 ± 13.52	107.50	.03
	RM - Index 2	51.28 ± 6.34	45.65 ± 13.46	114.50	.21

Notes. Mean and standard deviation are reported. I – SM = Internal Source Monitoring task; E- SM = External Source Monitoring task; RM – SM = Reality Monitoring task.

The number of *false recalls* at the sleep-related list showed a negative correlation with the number of correct responses to the Stroop task (r = -0.35, p = .03) and a positive correlation with the number of errors (r = 0.36, p =.03), while the number of *veridical recalls* for the same list positively correlates with the digit span score (r = 0.37, p = .02) and the Working Memory Index of the WAIS-IV (r = 0.34, p = .03). Also, the total number of *veridical recalls* showed a positive correlation with the Working Memory Index of the WAIS-IV (r = 0.35, p =.03) and a negative correlation with the total number of errors at the internal source monitoring task (r = - 0.32, p = .05). As for the relationship between sleep measures of the night preceding the DRM session and subsequent DRM performance, we observed a positive correlation between the total number of *false recalls* and the number of night awakenings (r = 0.27, p = .05), whereas the total number of *veridical recalls* showed a trend to a significant positive correlation with sleep duration (r = .27, p = .09). No other significant correlations emerged.

The results of the mediation analysis revealed a non significant indirect effect of sleep quality on false memories production through source monitoring ability (point estimate = 0.04, 95% CI = [-.085, .086]).

3.4. Discussion

In this study we investigated false memories production in individuals with insomnia and good sleepers, assuming that poor sleep quality and its cognitive consequences (for a review, see Fortier-Brochu et al., 2012) can render the former more prone to this phenomenon.

As a main result, we observed that the IN group globally produced more *false recalls* compared to the GS group, thus supporting an association between sleep quality and false memories production. In light of the literature on cognitive functioning in insomnia disorder, this result is of particular interest. According to the Activation-Monitoring theory (Roediger & McDermott, 1995; Roediger et al., 2001), during the retrieval phase participants generally rely on a source monitoring process to separate items that were studied from those that were not: in this phase, frontally-mediated executive functions are essential to ensure efficient source monitoring and memory accuracy (Johnson et al., 2012). To this regard, it has been observed that false memories production is increased in healthy subjects after sleep deprivation (Diekelmann et al., 2008; 2010), a procedure that strongly affects prefrontal functioning (Durmer & Dinges, 2005). In subjects suffering from insomnia, previous studies documented diurnal impairment in the same cognitive functions that may help to reject false memories and ensure efficient memory recall, i.e., retention and manipulation of information in working memory, inhibitory control and cognitive flexibility (Fortier-Brochu et al., 2012).

In our study we did not observe significant between-groups differences in most executive tasks. However, it would be hazardous to rule out the presence of executive impairments in insomnia. It might be that the changes in cognitive performance reported in this population are of a subtler and more situational kind (Fortier-Brochu et al., 2012) and therefore went partially undetected in the classical neuropsychological tasks adopted here. Indeed, previous studies also did not report executive functions impairments in insomniacs (Fortier-Brochu et al., 2012) or even found that this population performs better on these tasks than sleep-deprived good sleepers (Ballesio et al. 2018). Anyway, the suggested relationship between executive functioning and performance at the DRM paradigm (see, e.g., Leding et al., 2012; Peters et al., 2007) seems to be supported by the correlational analysis. In fact, we observed that the number of *false recalls* is negatively associated with accuracy at the Stroop task, and, conversely, that the number of veridical memories correlates both positively with the Working Memory Index and negatively with accuracy at the source monitoring task.

Additionally, an interesting result comes from the reality monitoring subtest of the source monitoring task, at which IN group were less accurate than GS group. Importantly, this bias was limited to the reality monitoring subtest, which requires participants to discriminate between internally and externally generated stimuli, i.e., the same ability required by the DRM task (while it did not extend to the ability to discriminate between two internal or two external sources). Together with our observation of higher *false recall* in the IN group, this finding lends support to the *SMF* (*Source Monitoring Framework*) in the explanation of false memories formation (Mitchell & Johnson, 2000) and to the hypothesis that the protective role of executive functioning against false memory is weakened in subjects suffering from insomnia. In fact, according to the *SMF*, false memories arise from an error of commission, that is when thoughts or images coming from one source (e.g., an external one) are erroneously attributed to another one (e.g., an internal one) (Mitchell & Johnson, 2000). The ability to correctly discriminate between two sources of information is linked to the efficiency of executive functioning and especially of memory retrieval processes (Johnson et al., 2012): the latter are strongly modulated by prefrontal functioning and are affected by acute (Mitchell & Johnson, 2000; Durmer & Dinges, 2005; Frenda et al., 2016) and chronic sleep loss, as in the case of individuals with insomnia (Fortier-Brochu et al., 2012). Therefore, in line with the *SMF*, we may explain our results by assuming that the IN group produced more numerous *false recalls* than the GS group because they are more susceptible to errors of commission as a consequence of their chronic sleep loss.

Data described so far should be still cautiously interpreted for the methodological limitation represented by the limited sample size, possibly accounting for the low magnitude of the finding and the negative results of our mediation analysis. However, taken overall, they encourage to thoroughly consider and further experimentally explore the hypothesis that the efficiency of executive functions, including the crucial source monitoring ability, promotes accurate retrieval and prevents false memories formation (Diekelmann et al., 2008; Peters et al., 2007).

Another interesting finding concerns the influence of stimulus type on DRM performance in the IN group. In line with literature on the attentional bias for sleep-related stimuli in individuals with insomnia (Giganti et al., 2017; Harris et al., 2015), we observed greater *false recalls* at the “sleep” list in the IN group compared to the GS group. Indeed, it is known that individuals with insomnia preferentially focus their attention on sleep-related items, considering them more salient and “threatening” than neutral ones (Espie et al., 2006; Harvey, 2002). This phenomenon has been previously observed through specific cognitive tests (see, e.g., Giganti et al., 2017; MacMahon et al., 2006), but had not yet been investigated in a task based on strong semantic associations, such as the DRM paradigm. Previous studies on the DRM task show that, relative to neutral word lists, arousing and negatively valenced lists can promote stronger associative connections between their items, so that the non-presented *lures* undergo greater activation and their false recall is facilitated at subsequent recovery (Howe et al., 2009; Otgaar et al., 2016). Therefore, we can explain our result by assuming that, for our IN group, items of the "sleep" list were more arousing and negatively valenced compared to the neutral list. Although in our study we did not directly ascertain whether IN group actually judged the sleep-related words as more negative and arousing than neutral ones, it has been previously shown that emotional valence and arousing capacity of sleep-related stimuli strongly differ between subjects with insomnia and good sleepers (Baglioni et al., 2010; Zhou et al., 2018).

Here again, given that the regression model did not show a significant interaction between group and stimulus type, we should take into account, beyond the small sample size, two further limits of our study: a) we did not include a specific measure of attentional bias which would have enabled us to exclude that between-groups differences are due to factors other than stimulus type; b) we could not analyse fine-grained differences in characteristics of our DRM lists, such as semantic relatedness and forward-backward associative strength between words, that might have played a role.

Surprisingly, we observed that *veridical recall* at the "sleep" list was impaired in the IN group, both relative to their own performance at the neutral list and to the GS group. This result suggests that, in the IN group, the attentional bias for the sleep-related list has different consequences on false and veridical memories, with an enhancement of *false recall* paralleling an impoverished *veridical recall*. Assuming in this group a higher activation driven by the salience of the sleep-list, a better *veridical recall* (relative both to the neutral list and to the GS group) for this list could be expected. In fact, previous studies showed that, in subjects with insomnia, the attentional bias generally enhances performance on sleep-related stimuli by locating greater attentional resources on them (see, for example, Giganti et al., 2017; MacMahon et al., 2006). Moreover, studies adopting the DRM paradigm on healthy subjects showed that salient stimuli generally enhance both false memories and veridical recollection of stimuli (Baird et al., 2003; Castel et al., 2007). However, some authors pointed out that certain stimuli not only promote false memories production but also, in parallel, reduce veridical retrieval (Brainerd et al., 2010). To this regard, adopting the DRM paradigm, Brainerd and colleagues (2010) observed that stimuli with negative valence generally increase false memories production but, at the same time, can also suppress true memory recollection, explaining this result in light of the *Fuzzy-Trace Theory* (Brainerd & Reyna, 2002). According to this theory, subjects simultaneously encode two independent traces for each word, respectively the "*verbatim* trace" (i.e. the trace related to the contextual features of a word and especially linked to veridical memory, corresponding in the DRM paradigm to the "*studied* words") and the "*gist*" or "*fuzzy*" trace (i.e. the trace representing the meaning of an item, preferentially linked to false memories production). The presentation of arousing and negatively valenced stimuli generally leads to strong *gist* traces but, at the same time, could also interfere with simultaneous processing of verbatim traces, causing lower subsequent *hit* rates for negative targets (Brainerd et al., 2010). In our study, the sleep-related word list might have performed in this way. In other words, supposing a high activation driven by the sleep-related list in our IN group, the triggering of the *gist* trace "sleep" in this group could have: on one hand, promoted *false recalls* at the sleep-related list; on the other hand, interfered with the processing of *verbatim* sleep-related traces and consequently impacted the *veridical recall* of words semantically associated to the *gist* trace.

Concerning the neutral word list, we did not observe between-groups differences either in the number of *false* or *veridical recalls*. This result seems to suggest that, in absence of interference such as that linked to sleep-related stimuli, individuals with insomnia have an efficient declarative memory system for words that are semantically related. In fact, as further evidence of this cognitive efficiency, we did not detect between-groups differences in the number of *intrusions* (i.e. words not belonging to the original word lists and also not semantically related to the critical *lure* words). It could be the case that the well-documented declarative memory deficits in people suffering from insomnia (Fortier-Brochu et al., 2012) specifically emerge in tasks assessing memory retrieval of semantically *unrelated* words. In other words, the semantic association between stimuli, which generally facilitates their recall (Aka et al., 2020; Silberman et al., 2007), would allow sleep-impaired individuals to achieve at the DRM task the same performance as good sleepers.

Because of the limited statistical power and the small effect size, our results need to be carefully interpreted and require further replications in larger samples. Indeed, as pointed out by Fortier-Brochu and colleagues (2012), small sample size and low statistical power are a common issue in studies comparing cognitive performance between people suffering from insomnia and good sleepers and may prevent the detection of small group differences. Nevertheless, our results added to previous literature on the attentional bias in subjects with insomnia and open to new research question.

In conclusion, our data show that individuals with insomnia symptoms produce more false memories than good sleepers and point to a relevant role of the attentional bias for sleep-related stimuli in the DRM task in this clinical sample. Although we cannot assert that the increase in false memories production in people suffering from insomnia is due to a widespread impairment of executive functioning, our results highlight in this population a notable bias in source monitoring ability that could have contributed to their false memories production.

References

Aka, A., Phan, T. D., & Kahana, M. J. (2020). Predicting recall of words and lists. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication. https://doi.org/10.1037/xlm0000964

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Association.

Ballesio, A., Cerolini, S., Ferlazzo, F., Cellini, N., & Lombardo, C. (2018). The effects of one night of partial sleep deprivation on executive functions in individuals reporting chronic insomnia and good sleepers. Journal of behavior therapy and experimental psychiatry, 60, 42–45. https://doi.org/10.1016/j.jbtep.2018.02.002

Baglioni, C., Lombardo, C., Bux, E., Hansen, S., Salveta, C., Biello, S., Violani, C., & Espie, C. A. (2010). Psychophysiological reactivity to sleep-related emotional stimuli in primary insomnia. Behaviour Research and Therapy, 48(6), 467–475. https://doi.org/10.1016/j.brat.2010.01.008

Baglioni, C., Regen, W., Teghen, A., Spiegelhalder, K., Feige, B., Nissen, C., Riemann, D. (2014). Sleep changes in the disorder of insomnia: A meta-analysis of polysomnographic studies. Sleep Medicine Reviews, 18, 195-213.

Baird, R.R. (2003). Experts sometimes show more false recall than novices: A cost of knowing too much. Learning and Individual Differences, 13, 349–355. 10.1016/S1041-6080(03)00018-9

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-Trace Theory and False Memory. Current Directions in Psychological Science, 11, 164–169. https://doi.org/10.1111/1467-8721.00192

Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Toglia, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. Journal of Experimental Child Psychology, 107(2), 137–154.

Castel, A. D., McCabe, D. P., Roediger, H. L. III, & Heitman, J. L. (2007). The Dark Side of Expertise: Domain-Specific Memory Errors. Psychological Science, 18(1), 3–5. https://doi.org/10.1111/j.1467-9280.2007.01838.x

Castronovo, V., Galbiati, A., Marelli, S., Brombin, C., Cugnata, F., Giarolli, L., Anelli, M. M., Rinaldi, F., & Ferini-Strambi, L. (2016). Validation study of the Italian version of the Insomnia Severity Index (ISI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 37, 1517–1524. https://doi.org/10.1007/s10072-016-2620-z

Conte, F., & Ficca, G. (2013). Caveats on psychological models of sleep and memory: a compass in an overgrown scenario. Sleep Medicine Review, 17, 105 – 121.

Curcio, G., Tempesta, D., Scarlata, S., Marzano, C., Moroni, F., Rossini, P. M., Ferrara, M., & De Gennaro, L. (2013). Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 34, 511–519. https://doi.org/10.1007/s10072-012-1085-y

Denckla, M. B. (1994). Measurement of executive function. In G. R. Lyon (Ed.), Frames of reference for the assessment of learning disabilities: New views on measurement issues (pp. 117–142). Paul H. Brookes Publishing Co.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. Behavioural Brain Research, 208 (2), 425 - 429.

Diekelmann, S., Landolt, H., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. PLoS ONE, 3, 3512–3512.

Durmer, J. S., & Dinges, D.F. (2005). Neurocognitive consequences of sleep deprivation. Seminars in Neurology, 25, 117-29.

Espie, C. A., Broomfield, N. M., MacMahon, K. M., Macphee, L. M., & Taylor, L. M. (2006). The attention-intention-effort pathway in the development of psychophysiologic insomnia: a theoretical review. Sleep Medicine Reviews, 10, 215–245. https://doi.org/10.1016/j.smrv.2006.03.002

Fortier-Brochu, E., Beaulieu-Bonneau, S., Ivers, H., & Morin, C. M. (2012). Insomnia and daytime cognitive performance: A meta-analysis. Sleep Medicine Reviews, 16, 83-94.

Frenda, S.J., Fenn, K.M. (2016). Sleep Less, Think Worse: The Effect of Sleep Deprivation on Working Memory. Journal of Applied Research in Memory and Cognition, 5, 463-469, https://doi.org/10.1016/j.jarmac.2016.10.001.

Giganti, F., Aisa, B., Arzilli, C., Viggiano, M. P., Cerasuolo, M., Conte, F., & Ficca, G. (2017). Priming recognition in good sleepers and in insomniacs. Journal of Sleep Research, 26, 345-352.

Haimov, I., Hanuka, E., & Horowitz, Y. (2008). Chronic insomnia and cognitive functioning among older adults. Behavioural Sleep Medicine. 6, 32–54. doi: 10.1080/15402000701796080

Harris, K., Spiegelhalder, K., Espie, C. A., MacMahon, K. M., Woods, H. C., & Kyle, S. D. (2015). Sleep-related attentional bias in insomnia: A state-of-the-science review. Clinical Psychology Review, 42, 16–27. https://doi.org/10.1016/j.cpr.2015.08.001

Harvey A.G. (2002). A cognitive model of insomnia. Behaviour Research and Therapy, 40, 869–893. https://doi.org/10.1016/s0005-7967(01)00061-4

Hayes, A. F. (2018). Partial, conditional, and moderated mediation: Quantification, inference, and interpretation. Communication Monographs, 85, 4-40.

Howe, M.L., M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. Journal of Memory and Language, 60, 229-251. https://doi.org/10.1016/j.jml.2008.10.002.

Iacullo, V. M., & Marucci, F. S. (2016). Normative data for Italian Deese/Roediger-McDermott lists. Behavior Research Methods, 48, 381–389. https://doi.org/10.3758/s13428-015-0582-3

Johnson, M. K., Raye, C. L., Mitchell, K. J., & Ankudowich, E. (2012). The cognitive neuroscience of true and false memories. Nebraska Symposium on Motivation. Nebraska Symposium on Motivation, 58, 15–52. https://doi.org/10.1007/978-1-4614-1195-6_2

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. Psychological Bulletin, 114, 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Joo, E. Y., Noh, H. J., Kim, J. S., Koo, D. L., Kim, D., Hwang, K. J., Kim, J. Y., Kim, S. T., Kim, M. R., & Hong, S. B. (2013). Brain Gray Matter Deficits in Patients with Chronic Primary Insomnia. Sleep, 36(7), 999–1007. https://doi.org/10.5665/sleep.2796

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. Sleep medicine reviews, 18(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Leding, J. K. (2012). Working memory predicts the rejection of false memories. Memory (Hove, England), 20, 217–223. https://doi.org/10.1080/09658211.2011.653373

Lo, J. C., Chong, P. L., Ganesan, S., Leong, R. L., & Chee, M. W. (2016). Sleep deprivation increases formation of false memory. Journal of Sleep Research, 25, 673–682. https://doi.org/10.1111/jsr.12436

MacMahon, M. A., Broomfield, N., & Espie, C. (2006). Attention bias for sleep-related stimuli in primary insomnia and delayed sleep phase syndrome using dot-probe task. Sleep, 29, 1420-7.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), The Oxford handbook of memory (pp. 179–195). Oxford University Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cognitive Psychology, 41, 49–100. doi: 10.1006/cogp.1999.0734

Morin CM (1993). Insomnia, psychological assessment and management. Guilford Press.

Natale, V., Esposito, M., Martoni, M., & Fabbri, M. (2006). Validity of the reduced version of the Morningness–Eveningness Questionnaire. Sleep Biological Rhythms, 4, 72–74. DOI: 10.1111/j.1479-8425.2006.00192.x

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. Psychonomic Bulletin & Review, 26(2), 387–400. https://doi.org/10.3758/s13423-018-1528-4.

Nienow, M.S., & Docherty, N. M. (2004). Internal Source Monitoring and Thought Disorder in Schizophrenia. The Journal of Nervous and Mental Disease, 192, 696–700.

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. Journal of Experimental Psychology: General, 145, 31-55. http://dx.doi.org/10.1037/xge000012

Peters, M. J. V., Jelicic, M., Verbeek, H., & Merckelbach, H. L. G. J. (2007). Poor working memory predicts false memories. European Journal of Cognitive Psychology, 19, 213-232. https://doi.org/10.1080/09541440600760396

Preacher, K., & Hayes, A. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40, 879–891. doi:10.3758/BRM.40.3.879

Roediger, H. L., 3rd, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. Psychonomic Bulletin & Review, 8, 385–407. https://doi.org/10.3758/bf03196177

Roediger, H.L, & McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. Journal of Experimental Psychology, 21, 803 – 14

Shekleton, J. A., Rogers, N. L., & Rajaratnama, S. M. (2010). Searching for the daytime impairments of primary insomnia. Sleep Medicine Reviews, 14, 47-60.

Sica, C. & Ghisi, M. (2007). The Italian versions of the Beck Anxiety Inventory and the Beck Depression Inventory-II: Psychometric properties and discriminant power. In M. A. Lange (Ed.), Leading-edge psychological tests and testing research (p. 27–50). Nova Science Publishers

Silberman, Y., Miikkulainen, R., & Bentin, S. (2005). Associating unseen events: Semantically mediated formation of episodic associations. Psychological Science, 16 161– 166

Spiegelhalder, K., Espie, C., Nissen, C., & Riemann, D. (2008). Sleep-related attentional bias in patients with primary insomnia compared with sleep experts and healthy controls. Journal of Sleep Research, 17, 191–196. https://doi.org/10.1111/j.1365-2869.2008.00641.x

Stroop, J. (1935). Studies of interference in serial verbal learning reactions. Journal of Experimental Psychology, 18, 643–662.

Trumble, I., Ferrer, O., Bay, C., Mollan, K. (2020). Precise and Accurate Power of the Wilcoxon-Mann-Whitney Rank-Sum Test for a Continuous Variable. R package version 3.0.2.

Vignatelli, L., Plazzi, G., Barbato, A., Ferini-Strambi, L., Manni, R., Pompei, F., D'Alessandro, R., & GINSEN (Gruppo Italiano Narcolessia Studio Epidemiologico Nazionale) (2003). Italian version of the Epworth sleepiness scale: external validity. Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 23(6), 295–300. https://doi.org/10.1007/s100720300004.

Violani, C., Devoto, A., Lucidi, F., Lombardo, C., & Russo, P.M. (2004). Validity of a short insomnia questionnaire: the SDQ. Brain Research Bulletin, 63, 5, 415-421

Wechsler, D. (2008). Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV). San Antonio, TX: NCS Pearson.

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. Psychological Bulletin, 120, 3–24.

Woods, H. C., Scheepers, C., Ross, K. A., Espie, C. A., & Biello, S.M. (2013). What are you looking at? Moving toward an attentional timeline in insomnia: A novel semantic eye tracking study. Sleep, 36, 1491–1499. doi: 10.5665/sleep.3042

Zhou, N., Zhao, C., Yang, T., Du, S., Yu, M., & Shen, H. (2018). Attentional bias towards sleep-related stimuli in insomnia disorder: a behavioural and ERP study. Journal of Sleep Research, 27, e12652. https://doi.org/10.1111/jsr.12652

Appendix

LIST 1	LIST 2	LIST 3	LIST 4
Bandiera	**Penna**	**Sonno**	**Fiume**
Italia	Inchiostro	Riposo	Foce
Rossa	Scrivere	Stanco	Affluente
Tricolore	Matita	Letto	Scorrere
Vento	Foglio	Notte	Diga
Nazione	Biro	Sogno	Pescare
Bianca	Calamaio	Dormire	Ponte
Alza	Nera	Profondo	Torrente
Asta	Piuma	Sbadiglio	Corrente
Colore	Scrittura	Sera	Flusso
Italiana	Scuola	Occhi	Pesce
Strisce	Sfera	Cuscino	Arno
Cuore	Stilo	Divano	Lungo
Mondiali	Astuccio	Pigiama	Valle
Innalzare	Grafia	Pisolino	Detriti
Inno	Oca	Veglia	Tevere

Supplementary Material

Source Monitoring task: description

The source monitoring task we proposed included three different sub-tests: Internal Source Monitoring (I-SM), External Source Monitoring (E-SM) and Reality Monitoring (RM-SM). A detailed description of these sub-tests is presented below.

Internal Source Monitoring sub-task

In the Internal Source Monitoring sub-task (I-SM), participants have to generate one-word responses to 16 very simple and incomplete statements, each of them presented on a computer screen (e.g., “the color of grass is ______.”). After each statement, one of two different instructions can appear in the centre of the screen, namely “respond” or “think”: on half of the trials, the instruction “respond” prompts the participants to vocalize the word that most logically completes the statement, whereas on the other half to just think of that word (**Figure a**). At the end of the task, participants generated eight vocalized and eight thought responses, in a counterbalanced order between subjects. Immediately after all statements had been presented, the participant completed a source recognition task containing 16 target words and eight new distractor words. We asked to identify the source of each word as “said”, “thought” or “new”.

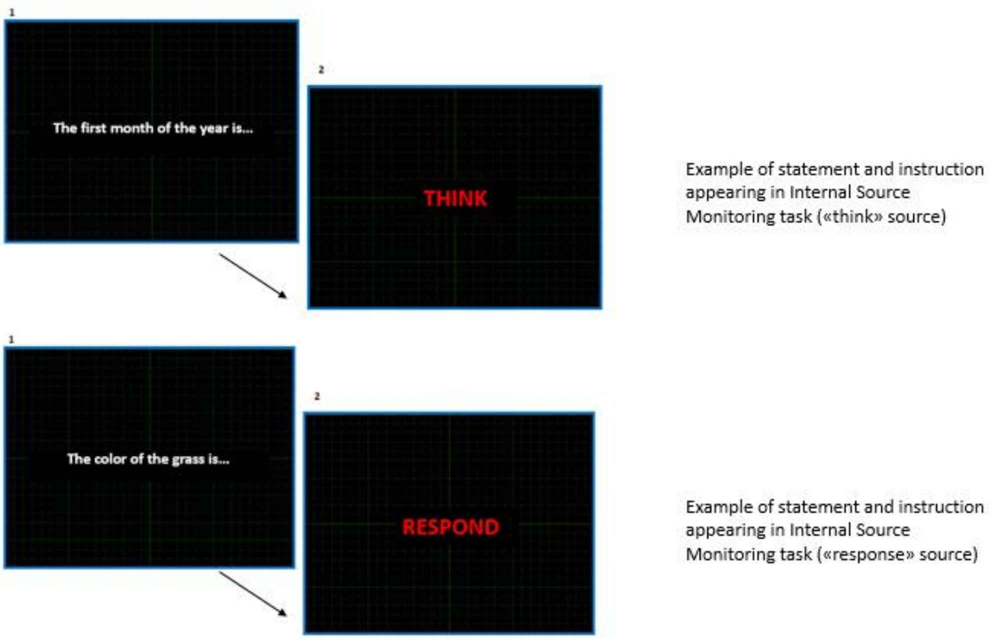

Figure a. Example of statements and instructions appearing during Internal Source Monitoring sub-task (I-SM).

External Source Monitoring sub-task

In the External Source Monitoring sub-task (E-SM) participants have to pay attention to one-word responses generated to 16 incomplete statements by two different external sources of information. An incomplete sentence appears in the centre of the screen and two different subjects (stylized drawn figures, either man or woman) can fill the blank with the word that logically complete the statement instead of participants. Particularly, after each incomplete statement, on half of the trials (8 statements) the stylized man and the target word appear on the screen (the first in the top and the second in the centre of the screen), indicating that the man has completed the statement; conversely, on the other half of the trial, both the stylized woman and the target word appear on the screen (8 statements), indicating that the woman has completed the statement (see **figure b**). When the target word appears in the screen, it is not accompanied by vocalization. The two source types of response were counterbalanced between subjects. Once all statements have been presented, participants completed a source recognition task containing 16 target words and eight new distractor words. We asked to identify the source of each word as "man", "woman" or "new".

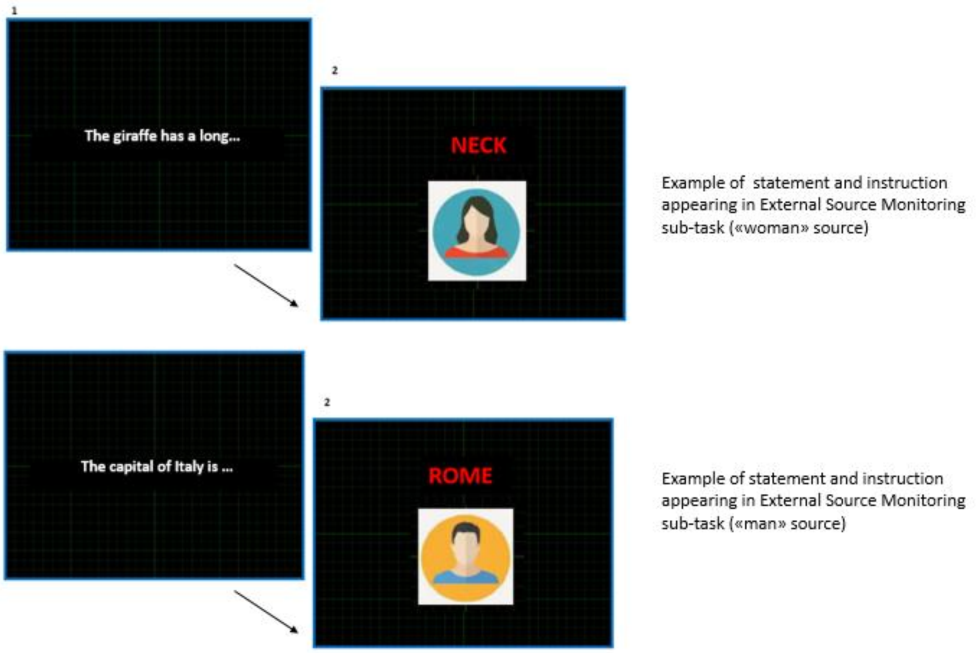

Figure b. Example of statements and instructions appearing during External Source Monitoring task (E-SM).

Reality Monitoring sub-task

In the Reality Monitoring sub-task (RM-SM), we asked participants to generate one-word responses to incomplete statements or to pay attention to one-word responses given by the external source. As in previous sub-tasks, statements were presented, one at a time, at the center of the computer screen followed by one of two different messages: "respons" or the word that logically completes the sentence. On eight statements ("response" instruction), subjects were required to choose between thinking or vocalizing the word (corresponding to the internal source of information), while on other eight statements they had to pay attention to the word appearing on the screen, as external source (**Figure c**). The types of response for each statement were counterbalanced between subjects. Once all statements had been presented, the participant completed a source recognition task containing 16 target words and eight new distractor words. We asked to identify the source of each word as "response", "experimenter's response" or "new".

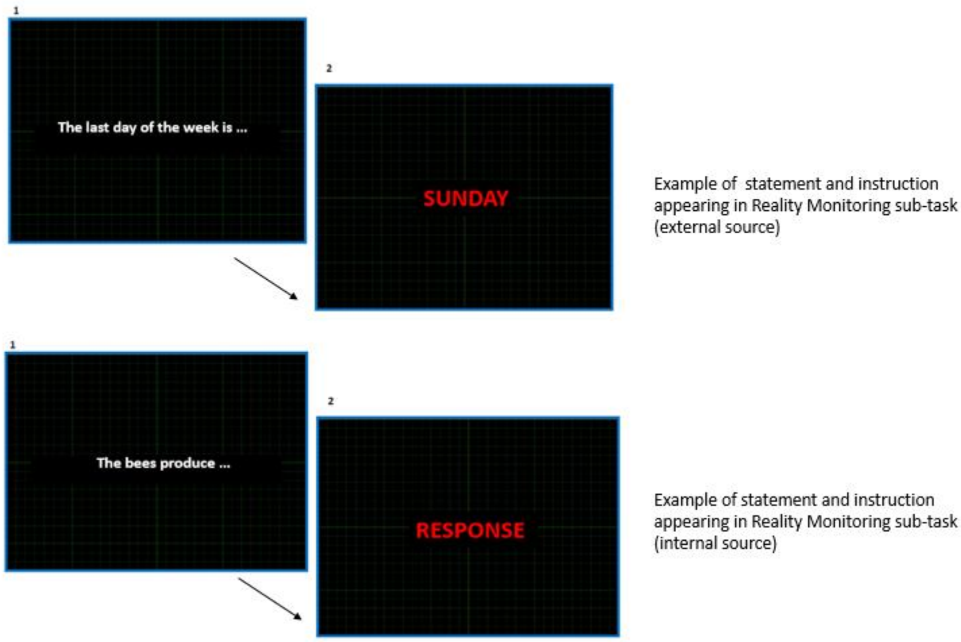

Figure c. Example of statements and instructions appearing during Reality Monitoring sub-task (RM-SM).

Overall, 48 statements were presented to participants during the whole task. In order to avoid interference between the three source monitoring sub-tasks, we selected target words with no semantical and phonological similarities; sub-task's statements are also semantically unrelated.

Each statement was first piloted on 50 college students to ensure that they used the same target word to complete each statement. An inter-subject agreement of 99% for each word had been required.

Chapter 4

False memories production in individuals with insomnia and good sleepers: a polysomnographic study

4.1 Introduction

Research over the past years showed that sleep plays a relevant role in memory consolidation. Sleep after learning not only produces a strengthening of newly-acquired memory traces but also promotes qualitative changes in memory representations thanks to an active system consolidation process (Rash & Born, 2013). This process leads to a qualitative transformation of newly encoded memories and brings to re-organize mnemonic traces in several ways (Landmann et al., 2014). For example, behavioural findings showed that sleep plays an important role in the formation of associative schemata, facilitates the integration between new and old memory traces, and promotes the extraction of hidden rules from sets of information (Landmann et al., 2014; Conte & Ficca, 2013). One of the possible by-products of this memory "*reshaping*" (Conte & Ficca, 2013) is the abstraction of the "*gist*" from the newly encoded memories and the emergence of new memory contents that have not been directly learned, i.e. false memories.

False memories represent memory distortions in which people erroneously recall information or events that did not happen or claim to remember them differently from how they originally appeared (Roediger & McDermott, 1995). Previous studies support the role of sleep in the development of false memories. For example, Payne and colleagues (2009) found that sleep increases false memories production compared to an equivalent period of daytime wakefulness, adopting the Deese-Roediger-McDemott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995). Later, other studies replicated these findings (Diekelmann et al., 2010; Pardilla Delgado & Payne, 2017). Moreover, the sleep effect for false memories has also been observed for emotionally negative words (McKeon et al., 2012) and appeared resistant across long delays (Pardilla Delgado & Payne, 2017).

To date, the effect of sleep on false memories production has been investigated only in healthy subjects, neglecting individuals with chronically disturbed sleep. In this regard, it is reasonable to expect that when sleep patterns are chronically altered, such as in the case of insomnia disorder, the ability to efficiently consolidate memory traces may be impaired (Cipolli et al., 2013; Cellini, 2017), with consequences also on false memories production, considered as the by-product of a memory reorganization.

The disorder of insomnia is a common complaint in clinical practice (Aikens & Rouse, 2005) and represents a highly prevalent condition (Morin et al., 2006; Ohayon et al., 2008). According to the criteria of DSM 5 (American Psychiatric Association, 2013), insomnia is a sleep disorder characterized by subjective complaints of non-restorative sleep with difficulties in initiating and/or maintaining sleep, accompanied by decreased daytime functioning (American Psychiatric Association, 2013). As for sleep characteristics, objective impairments in sleep continuity and architecture have been documented in individuals suffering from insomnia (for a meta-analysis, see Baglioni et al., 2013). In particular, the disorder appears to be accompanied by a reduction of total sleep time and an increase in sleep onset latency, in the number of night-awakenings, and in the duration of wake time after sleep onset, i.e. all sleep continuity alterations resulting in a diminished sleep efficiency compared to good sleep. In parallel, a reduced percentage of Slow Wave Sleep (SWS) and REM have been documented in people with insomnia compared to good sleepers.

The above-mentioned sleep impairments could negatively affect the active system memory consolidation process occurring during sleep in individuals with insomnia, with a consequent impact on false memories production. Surprisingly, sleep-dependent memory consolidation process remains scarcely investigated in insomnia (Cellini, 2017) and there is a lack of studies specifically addressing the impact of the disorder on the qualitative reorganization of memories occurring during sleep (Landmann et al., 2014). Therefore, it still remains unknown whether individuals with insomnia produce false memories during sleep and how their poor sleep quality affects this phenomenon.

In this study, we aimed to cover this gap by investigating the effect of a retention period spend asleep or awake on false memories production at the DRM task in individuals suffering from insomnia and in a control group of good sleepers.

4.2. Materials and method

4.2.1. Participants

Potential participants to the study were approached at university sites and asked to fill out a set of screening questionnaires including: the Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), the Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016), the Sleep Disorder Questionnaire (SDQ; Violani et al., 2004), the Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007), and the Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007). In addition, they were administered a semi-structured interview at the sleep laboratory, conducted by a licensed psychologist who had received specific training, in order to assess general medical condition and health habits, presence of psychiatric disorders and of sleep disorders. The presence of clinical insomnia was specifically addressed by means of the semi-structured interview (Morin, 1993).

Based on scores at the screening instruments and on the interview, 35 university students were recruited for the study and included in either the "insomnia group" (**IN group**, n = 17) or in the "good sleep group" (**GS group**, n = 18).

Inclusion criteria common to both groups were: absence of any relevant somatic or psychiatric disorder; absence of clinically significant depression and anxiety symptoms (BDI-II score ≤ 29; BAI score ≤ 25); no history of drug or alcohol abuse; absence of sleep disorders (other than insomnia for the IN group) and of any sleep apnea or respiratory disorder symptom; having a regular sleep-wake pattern (e.g., individuals with irregular study or working habits such as shift-working were excluded); no use of psychoactive medication or alcohol at bed-time.

In addition, for inclusion in the IN group, participants had to score ≥ 5 at the PSQI, ≥ 8 at the ISI and to be classified as presenting "clinically significant insomnia" at the SDQ; further, they had to fully meet DSM-5 criteria for Insomnia Disorder, as verified through the interview. In this regard, the IN group included participants reporting at the interview: a) difficulty in initiating and/or maintaining sleep and/or early morning awakening, b) clinically significant distress or impairments in relevant areas of functioning due to the sleep disturbance, c) sleep difficulties occurring at least three nights per week, for at least 3 months, 4) no medical/psychiatric condition, substance abuse, and/or other sleep disorder (American Psychiatric Association, 2013).

The inclusion in the GS group was based on: PSQI score < 5, ISI score < 8, being classified as "good sleeper" at the SDQ, and absence of any sleep disorder as also verified through the interview.

The local Ethical Committee approved the research protocol and all participants signed a consent form. There was no money or credit compensation for participating in the study.

4.2.2. Procedure

A mixed design has been adopted to evaluate false memories production in IN group and GS group using the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995) in two different conditions, namely after a night spent in sleep ("*Sleep*" condition) and after an interval spent in daytime wakefulness ("*Wake*" condition). The two conditions were administered in balanced order across participants. For each participant, the two conditions were performed with an interval of one week from each other. During the 5 days preceding the study conditions, subjects were requested to keep sleep-wake schedules and daily activities as habitual as possible. To control for these factors, participants completed a sleep diary.

In Sleep, subjects' nocturnal sleep was polysomnographically recorded; an adaptation night was administered 2 days before the experimental night to avoid a first night effect.

On the days scheduled for the Sleep condition, the experimenter arrived at the subject's house approximately 1 hour and a half before usual bedtime and proceeded to electrodes set up. Then, subjects performed the first phase of the DRM paradigm just before bedtime (i.e. *DRM learning phase*) and were re-tested in the morning (i.e. *DRM test phase*), approximately 30 minutes after final awakening to allow sleep inertia dissipation (for a detailed description of the DRM paradigm, see the next paragraph). Bedtime and awakening time were not predetermined: participants were asked to maintain their regular sleep-wake habits in both conditions.

In Wake, participants spent awake the retention period (i.e., the time interval between the *learning phase* and the *test phase* of the DRM paradigm), which corresponded to the duration of the subject's habitual sleep duration, obtained from sleep diaries. During the retention interval, subjects were allowed to engage in non-cognitively demanding activities, which they later reported in a questionnaire, and were instructed to avoid naps. Both the testing sessions were performed at the sleep laboratory.

To control for subjective sleepiness levels, participants completed the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990) before performing the *learning* phase and the *test* phase of the DRM paradigm in both conditions. In addition, to control for psychological factors possibly affecting memory performance, participants also rated their concentration and motivation on 5-point Likert-scales (1 indicating "*not at all*" and 5 indicating "*very much*"), as in Diekelmann and colleagues (2008).

4.2.3. Instruments

False memory task

The Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995) was adopted to investigate false memories production in IN group and GS group.

In our study, participants individually learned 16 DRM word lists, 8 in Sleep condition and 8 in Wake condition respectively (see **Appendix**). Each list consisted of 15 words (named "*studied*" words) semantically related to a critical word, defined "*lure*", that was not presented by the experimenter (e.g. "*bed*", "*rest*", "*nap*" with "*sleep*" as *lure* word). As in Roediger and McDermott (1995) and Iacullo and Marucci (2016), the words in each list were presented in order of associative strength with the unpresented *lure* word (from strongest to weakest).

Lists were selected from those already translated in Italian by Iacullo & Marcucci (2016) and did not contain words juxtaposition. Different lists were employed in Sleep and Wake condition. Furthermore, in order to control for possible differences between the lists presented in the two experimental conditions, 15 individuals (10 F and 5 M; mean age = 27.45 ± 2.35), who were not enrolled in the main study, were asked to rate on a 1 – 5 Likert scale the familiarity, activation and valence of each word belonging to the two sets of lists, as well as their respective critical *lures*. Comparisons revealed no significant difference for familiarity ($t = -2.10$, $p = .54$), activation ($t = 1.02$, $p = .43$) or valence ($t = 1.15$, $p = .22$) between the two set of word lists.

As for task administration, the experimenter read the lists aloud with an interval of 30 s between them. In this "*learning phase*", participants were instructed to memorize the *studied* words as accurately as possible and were informed that they would be tested on them later. The *test phase* has been performed after an interval spent in sleep (Sleep condition) and in diurnal wakefulness (Wake condition).

The *test phase* involved two separate tasks: a free recall task and a recognition task. All participants first performed the free recall task and the recognition task immediately after.

For the free recall task, participants were given a blank piece of paper and are asked to write any words they remembered from the previously heard lists. They had 10 min to complete this task and were notified when they had 2 min remaining. As for the recognition task, participants were shown another list of 56 words on a paper, containing *lure* words (n = 8), original *studied* words (n = 24, taken from serial positions 1, 8, and 10 in the DRM studied lists, as in Roediger & McDermott, 1995), and unrelated distractors (n = 24), all presented in the same randomized order between subjects. They were asked to identify those from the previous lists ("old" words), as well as new words. In addition, they were requested to rate their answer on a 3-point scale (1 = "*Remember*"; 2 = "*Know*"; 3 = ''*I had to guess*'') for those words that were previously judged as ''old'' (Roediger & McDermott, 1995).

Sleep recordings

Polysomnographic recordings were performed during the Sleep condition by recording six electroencephalographic (EEG channels: F3, F4, C3, C4, O1, O2, referenced against contralateral mastoids A1 and A2), two electro-oculographic (LOC-A2, ROC-A1), and a bipolar submental electromyogram channels, according to standard guidelines (Iber, 2007). Data were acquired by means of a BluNet multichannel recording system (Ne.Ro SRL, Florence, Italy) at a sample rate of 200 Hz. Sleep recordings were band-passed (0.3 – 35 Hz) and then visually scored according to standard criteria (Iber et al., 2007) by an expert technician. To verify scoring reliability, 10 randomly selected sleep recordings were also scored by another technician. Inter-rater agreement was 92%.

4.2.4. Data analysis

Outcome measures of the DRM paradigm considered were:

- At the free recall test, the number of *false recalls* (i.e. total number of falsely recalled critical *lure* words), the number of *veridical recalls* (i.e. the total number of words correctly recalled from the original *studied* word lists), and the number of *recalled intrusions* (i.e. the total number of recalled words not corresponding to *studied* words or to the critical *lure* words).
- At the recognition test, the number of *false recognitions* (i.e. ‘‘old” responses given to critical *lure* words), the number of *hits* (i.e. ‘‘old” responses given to *studied* words), and the number of *false alarms* (i.e. ‘‘old” responses given to unrelated distractors); confidence rating attributed to *false recognition, hits,* and *false alarm*s.

Classical architecture sleep variables considered in the study were: time in bed (TIB; i.e. total amount of time, in minutes, from lights off to final awakening), total sleep time (TST; i.e. total amount of time, in minutes, from the first appearance of N1 to final awakening), actual sleep time (AST; i.e. total time spent in sleep states, expressed in minutes), sleep-onset latency in minutes (SOL), sleep stage proportions over TST (N1%, N2%, N3%, REM%), percentage of wake after sleep onset over TST (WASO%) and sleep efficiency (SE%; i.e. percentage of AST over TIB).

Objective sleep quality was evaluated in IN group and GS group through an additional set of variables (Conte et al., 2012):

- *sleep continuity*: total frequency of awakenings per hour of AST; frequency of brief (< 4 epochs) and long (≥ 4 epochs) awakenings per hour of AST;

- *sleep stability*: frequency of arousals per hour of AST (arousals are defined as all transitions to shallower NREM sleep stages and from REM sleep to N1); frequency of state transitions (defined as all transitions from one state to another) per hour of TST; frequency of "Functional Uncertainty Periods" ("FU periods"; defined as periods in which a minimum of three state transitions follow one another with no longer than 1.5 min intervals; Salzarulo et al., 1997) per hour of TST; the percentage of total time spent in FU over TST (TFU%);
- *sleep organization*: number of complete sleep cycles, defined as sequences of NREM and REM sleep (each lasting at least 10 min) not interrupted by periods of wake longer than 2 min (Conte et al., 2012); total time spent in cycles over TST (TCT%).

Microarousals were also detected, by means of an automatic analysis performed through Polysmith software package (Nihon Kohden Polysmith version 9.0) and based on the following definition from the Sleep Disorders Atlas Task Force (Bonnet et al., 1992): an abrupt change in EEG frequency, including theta, alpha and/or frequencies greater than 16 Hz (but not spindles), lasting from 3 to 14s. From this analysis we derived the microarousal index (n/AST hr).

The number of sleep spindles (*total spindles*: 11 - 16 Hz; *slow spindles*: 9 – 12.5 Hz; *fast spindles*: 12.5 – 16 Hz) was automatically detected from the central derivations (C3 and C4) and the density of total, slow and fast spindles was calculated as the mean number of spindles (total, slow and fast) detected over the time spent in N2 and N3.

After checking for variables normality with the Shapiro-Wilk test, a repeated measure ANOVA was performed with "*condition*" (i.e. Sleep and Wake) as within factor, "*group*" (i.e. IN group and GS group) as between factor, and DRM variables as dependent variables. Least significant difference (LSD) post hoc comparisons were performed when appropriate.

Between-groups differences in objective sleep measures were analyzed through a *t Student's* test for independent samples. The same statistic has been adopted to evaluate discrepancies between the IN group and the GS group in subjective sleepiness levels, concentration and motivation in the Sleep and Wake condition, as well as in demographic variables. Chi-square test was carried for all binomial variables.

A Pearson's correlation analysis was performed to investigate the association between sleep measures and DRM performance in the whole sample.

Analyses were performed using SPSS (version 27) and the significance level was set at $p \leq .05$.

4.3. Results

Two subjects had to be excluded from analyses, the first one due to technical problems occurring during data collection, whereas the second one did not comply with the request of keeping habitual sleep-wake schedules during the data collection phase. Thus, the final sample included 17 subjects suffering from insomnia (IN group; 4 M, 13 F; mean age = 26.6 ± 6.71) and 15 good sleepers (GS group; 5 M, 10 F; mean age = 27.3 ± 6.18).

Demographic characteristics, circadian preference, habitual daytime sleepiness and sleep quality in the two groups

Table 1 displays characteristics of the final IN group and GS group. The two groups did not differ in age, gender distribution, circadian preference (measured through the reduced version of the Morningness–Eveningness Questionnaire; Italian version from Natale et al., 2006) and daytime sleepiness (measured through the Epworth Sleepiness Scale; Italian version from Vignatelli et al., 2003). Instead, significant between-group differences emerged in PSQI and ISI global scores (**Table 1**).

Table 1. Age, gender distribution, circadian preference, total scores in ESS, PSQI and ISI in IN group and GS groups.

	IN group	GS group	Statistical test
Age	26.6 ± 6.71	27.3 ± 6.18	$t = 0.29$, $p = .769$
Gender	4 M, 13 F	5 M, 10 F	$\chi^2 = 0.38$, $p = .538$
MEQr score	14.4 ± 3.58	13.6 ± 2.56	$t = 1.12$, $p = .271$
ESS	6.71 ± 3.60	6.47 ± 3.58	$t = -0.18$, $p = .852$
PSQI	7.75 ± 2.44	4.00 ± 1.04	$t = -5.34$, $p < .001$
ISI	10.88 ± 3.32	3.20 ± 2.21	$t = -7.52$, $p < .001$

Notes. MEQr: Morningness-Eveningness Questionnaire (reduced version); IN group: insomnia group; GS group: good sleep group; ESS: Epworth Sleepiness Scale; M: Males; F: Females; PSQI = Pittsburgh Sleep Quality Index; ISI = Insomnia Severity Index; *t* Student's test is reported for between-groups comparisons for all variables except gender. Results of the chi-squared test are reported for differences in gender distribution.

Objective sleep measures in the two groups (Sleep condition)

Classical sleep architecture variables

Table 2 displays between-groups comparisons in classical sleep architecture variables. The IN group showed longer *SOL*, higher *WASO%*, lower *SE%* and a trend to longer *TIB* compared to the GS group. No other statistically significant differences emerged between groups.

Table 2. Sleep architecture variables in the IN group and GS group.

	IN group	GS group	*t*	*p*	*Cohen's d*
TIB *(min)*	491.76 ± 49.21	454.23 ± 64.42	-1.83	.076	-0.65
TST *(min)*	494.47 ± 149.28	444.00 ± 63.78	- 1.21	.234	-0.43
AST *(min)*	424.92 ± 55.87	429.60 ± 61.45	0.23	.823	0.80
SOL *(min)*	17.44 ± 10.31	7.63 ± 4.94	- 3.34	**.002**	-1.19
N1%	9.89 ± 3.79	9.16 ± 2.79	- 0.59	.560	-0.21
N2%	44.91 ± 5.81	45.25 ± 8.36	0.13	.896	0.05
N3%	25.57 ± 11.41	26.51 ± 9.31	- 0.19	.844	-0.07
REM%	18.03 ± 5.08	19.04 ± 4.90	0.57	.203	0.21
WASO%	7.91 ± 8.23	3.18 ± 1.13	- 2.20	**.036**	-0.78
SE%	86.50 ± 7.28	94.57 ± 1.56	4.19	**< .001**	1.49

Notes. Significant between-groups differences are in bold. TIB = Time in Bed; TST = Total Sleep Time; AST = Actual Sleep Time; SOL = Sleep Onset Latency; WASO = Wake After Sleep Onset; SE = Sleep Efficiency

Sleep continuity, stability and organization measures

Table 3 displays between-groups differences in *sleep continuity*, *stability* and *organization* measures. IN participants showed higher *frequency of long awakenings* and a higher *micro-arousal index* compared to controls, as well as fewer *complete sleep cycles* and reduced *time spent in sleep cycles*. No other significant differences emerged between groups.

Table 3. *Sleep continuity*, *stability* and *organization* measures in the IN group and GS group.

	IN group	**GS group**	***t***	***p***	***Cohen's d***
Awakenings *(fq)*	3.02 ± 1.12	2.37 ± 0.75	- 1.88	.070	-0.67
Short awakenings *(fq)*	2.55 ± 1.09	2.23 ± 0.76	- 0.94	.353	-0.33
Long awakenings *(fq)*	0.48 ± 0.49	0.13 ± 0.11	- 2.70	**.011**	-0.98
Arousal *(fq)*	6.69 ± 2.44	6.28 ± 1.73	0.90	.376	-0.32
State transitions *(fq)*	22.69 ± 6.06	20.31 ± 3.23	- 1.36	.183	-0.48
FU *(fq)*	1.91 ± 0.72	1.69 ± 0.57	- 0.95	.347	-0.34
TFU%	16.08 ± 7.12	13.67 ± 4.61	- 1.12	.270	-0.39
Microarousal index	11.44 ± 4.21	8.97 ± 2.20	- 2.04	**.050**	-0.73
Sleep cycles *(n)*	0.29 ± 0.58	0.93 ± 0.96	2.30	**.029**	0.82
TCT%	4.23 ± 8.42	17.69 ± 20.12	2.54	**.017**	0.89

Notes. Significant between-groups differences are in bold. FU = Functional Uncertainty periods; TFU = Total time spent in FU; TCT = total time spent in sleep cycles

Sleep spindles density

No between-groups differences emerged for the density of *total spindles* (IN group = 2.52 ± 2.37 vs GS group = 1.60 ± 1.09; *t* (29) = -1.34, *p* =.192; *Cohen's d* = - 0.48), *slow spindles* (IN group = 0.91 ± 0.88 vs GS group = 0.69 ± 0.51; *t* (29) = - 0.82, *p* = .416; *Cohen's d* = - 0.29) or *fast spindles* (IN group = 2.21 ± 2.14 vs GS group = 1.37 ± 0.90; *t* (29) = - 1.35, *p* =.185; *Cohen's d* = - 0.49).

False memories production

Table 4 displays descriptive statistics of free recall and recognition performance of both groups in Sleep and Wake.

Table 4. Descriptive statistics of DRM performance (free recall and recognition tasks) in the two groups in both conditions. Means and standard deviations are reported.

	IN group		**GS group**	
	Sleep	**Wake**	**Sleep**	**Wake**
False recalls	1.53 ± 1.23	2.53 ± 1.84	2.87 ± 1.81	1.67 ± 1.40
Veridical recalls	22.12 ± 12.27	18.41 ± 7.88	25.07 ± 9.78	20.20 ± 9.46
Recalled intrusions	2.35 ± 1.80	2.41 ± 1.77	3.00 ± 2.03	2.60 ± 2.64
False recognitions	5.62 ± 2.21	5.12 ± 2.44	6.33 ± 1.54	5.66 ± 1.49
Hits	17.00 ± 4.27	15.75 ± 3.94	18.40 ± 3.16	16.66 ± 4.25
False alarms	3.31 ± 2.84	2.75 ± 2.95	3.07 ± 3.13	2.67 ± 2.44

Free recall task

As for *false recalls*, repeated measures ANOVA revealed a significant *group* × *condition* interaction ($F(1,30) = 12.01$, $p = .002$, $\eta^2_p = .29$), whereas the main effects of *group* ($F(1,30) = 0.26$, $p = .614$, $\eta^2_p = 0.009$) and *condition* ($F(1,30) = 0.09$, $p = .755$, $\eta^2_p = 0.003$) were not significant. Post hoc comparisons showed that the IN group produced more *false recalls* in Wake compared to Sleep ($t(30) = -2.30$, $p = .029$). The opposite result emerged for GS participants, who produced more *false recalls* in Sleep than in Wake ($t(30) = 2.59$, $p = .015$; **Figure 1**). In addition, the GS group falsely recalled more critical *lure* words than IN subjects in Sleep ($t(30) = 2.47$, $p = .019$; **Figure 1**).

As for *veridical recall*, the main effect of *condition* was significant ($F(1,30) = 6.39$, $p = .017$, $\eta^2_p = 0.18$), with participants correctly recalling more *studied* words in Sleep than in Wake ($t(30) = 2.52$, $p = .019$). Neither the main effect of *group* ($F(1,30) = 0.58$, $p = .452$, $\eta^2_p = 0.019$) nor the *group* × *condition* interaction ($F(1,30) = 0.12$, $p = .735$, $\eta^2_p = 0.004$) were significant.

As for *recalled intrusions*, repeated measures ANOVA revealed no significant effect of *group* (F(1,30) = 0.52, p = .475, η^2_p = 0.017) or *condition* (F(1,30) = 0.14, p = .709, η^2_p = 0.005), nor any *group* × *condition* interaction (F(1,30) = 0.84, p = .617, η^2_p = 0.008).

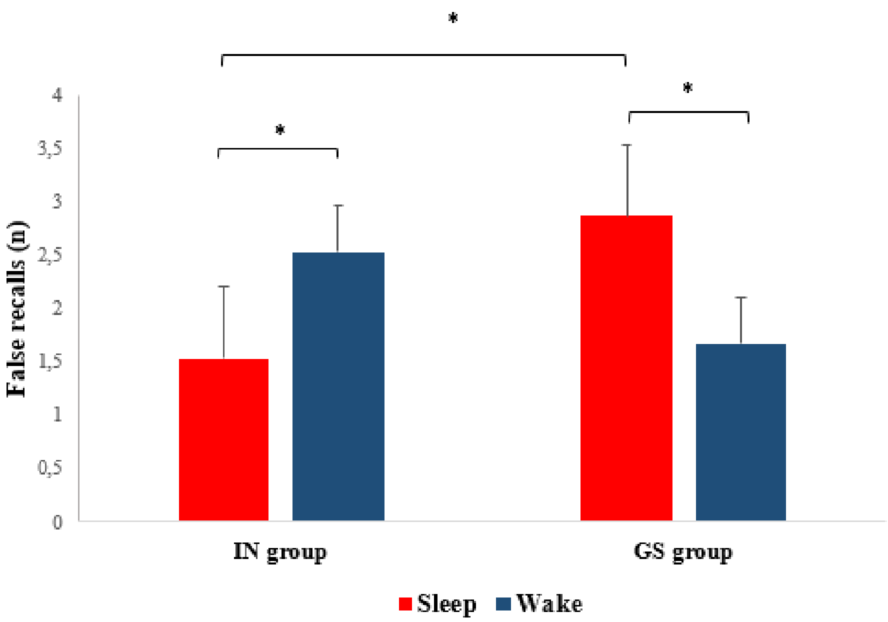

Figure 1. Number of *false recalls* in Sleep and Wake, in the IN group and GS group. *p < .05. Error bars represent standard errors.

Recognition task

As for *false recognitions*, the repeated measure ANOVA revealed no significant effect of *group* (F(1,29) = 1.02 p = .320, η^2_p = 0.034), *condition* (F(1,29) = 0.27, p = .111, η^2_p = 0.085) and *group* × *condition* interaction (F(1,29) = 0.05, p =.816, η^2_p = 0.002).

As for the number of *hits*, the main effect of *condition* was significant (F(1,29) = 4.27, p = .036, η^2_p = 0.14). Specifically, participants correctly recognized more *studied* words in Sleep than in Wake (t(29) = 2.20, p =.036). Instead, we observed no main effect of *group* (F(1,29) = 0.87, p = .359, η^2_p = 0.029) and no *group* × *condition* interaction (F(1,29) = 0.13, p = .724, η^2_p = 0.004).

The number of *false alarms* showed no significant main effect of *group* (F(1,29) = 0.03, p =.854, η^2_p = 0.001) or *condition* (F(1,30) = 0.85, p = .363, η^2_p = 0.029). The interaction *group* × *condition* was also not significant (F(1,30) = 0.02, p = .877, η^2_p = 0.001).

Finally, the interaction *group* × *condition* was not significant for confidence rating attributed to *false recognitions* ($F(1,28) = 0.15$, $p = .269$, $\eta^2_p = 0.067$), *hits* ($F(1,28) = 0.21$, $p = .647$, $\eta^2_p = 0.007$) and *false alarms* ($F(1,28) = 0.05$, $p = .942$, $\eta^2_p = 0.001$). No main effects of *group* and *condition* were observed for these variables (all *p-values* >.05).

Sleepiness, concentration and motivation

In Sleep, no between-groups differences emerged in subjective sleepiness, concentration and motivation levels referred by participants either at the *DRM learning phase* or at the *DRM test phase* (all *p*-values >.05). In Wake, the IN group reported at the learning phase higher sleepiness ($t(30) = 2.73$, $p = .010$) and higher motivation ($t(30) = -2.29$, $p = .023$) compared to the GS group (see **Table 5** for descriptive statistics).

Table 5. Subjective ratings of sleepiness, concentration and motivation (mean ± standard deviation) collected before the *learning phase* and *test phase* of the DRM paradigm in both groups in the two conditions.

	DRM Learning Phase			
	Sleep		**Wake**	
	IN group	**GS group**	**IN group**	**GS group**
KSS	4.47 ± 1.97	5.20 ± 1.97	4.18 ± 1.47	2.80 ± 1.37
Concentration	3.24 ± .83	3.20 ± .86	2.88 ± .86	2.60 ± .73
Motivation	2.24 ± .83	2.27 ± 1.03	2.94 ± 1.14	2.07 ± .88
	DRM Test Phase			
	Sleep		**Wake**	
	IN group	**GS group**	**IN group**	**GS group**
KSS	4.65 ± 1.83	4.33 ± 1.87	3.35 ± 1.45	3.07 ± 1.59
Concentration	3.12 ± 1.05	3.00 ± .85	2.88 ± .93	2.64 ± .93
Motivation	2.59 ± .87	2.27 ± .79	2.76 ± 1.25	2.14 ± 1.09

Notes. KSS = Karolinska Sleepiness Scale.

Relationship between sleep measures and DRM performance

Pearson's correlation analysis revealed that the number of *false recalls* positively correlated with SE% (*r* = 0.43, *p* = .014) and negatively with SOL (*r* = -0.39, *p* = .029) and N2% (*r* = -3.49, *p* = .050). Furthermore, the number of *veridical recalls* negatively correlated with the frequency of total (*r* = -0.41, *p* = .021) and short awakenings (*r* = -0.35, *p* = .049).

As for the recognition task, the number of *false recognitions* showed a positive correlation with AST% (*r* = 0.39, *p* = .028) and a negative one with *long awakenings frequency* (*r* = - 0.37, *p* =.043). In addition, the same variable showed a trend to a significant positive correlation with *TCT%* (*r* = 0.35, *p* = .056) and a trend to a negative one with the *microarousal index* (*r* = -0.31, *p* =.085). The number of *hits* correlated with the *number of cycles* (*r* = 0.45, *p* = .010), and showed a trend to a significant negative correlation with *WASO%* (*r* = - 0.32, *p* = .08) and *total awakenings frequency* (*r* = - 0.32, *p* = .07). Finally, the analysis revealed a positive correlation of the number of *false alarms* with N1% (*r* = 0.38, *p* = .036) and a trend to a positive correlation with frequency of FU periods (*r* = 0.35, *p* = .057).

4.4. Discussion

In this study we investigated the influence of sleep quality on false memories production in individuals suffering from insomnia and good sleepers. Here, we assumed that false memories preferentially arise from a well-functioning process of memory consolidation and reorganization occurring during an efficient, continuous, stable and well-organized sleep episode. Therefore, we expected to observe a detrimental effect of chronic poor sleep, characterizing individuals with insomnia, on false memories production.

A preliminary remark concerns our finding of significant differences in objective sleep parameters between the two groups. Indeed, as expected, IN relative to GS participants displayed longer sleep onset latency as well as less efficient sleep and higher wake after sleep onset proportion. This observation is in line with available literature on insomnia, documenting impairments in sleep efficiency in this population compared to good sleepers (Baglioni et al., 2013). In addition, here we investigated an additional set of objective sleep variables that have repeatedly been proposed as more fine-grained indices of sleep quality (Conte & Ficca, 2013; Conte et al., 2021) and have shown to sustain efficient sleep-dependent consolidation processes (e.g., Conte et al., 2012; Haimov & Shatil, 2013; Sergeeva et al., 2017), i.e. *sleep continuity*, *stability* and *organization* measures. In our study, the IN group showed higher frequency of long awakenings, higher micro-arousal index, a reduced number of sleep cycles and reduced time spent in cycles compared to the GS group. This finding represents the first objective observation that individuals with insomnia display impairments in sleep *continuity*, *stability* and *organization*, which had never been explored in this population, with possible detrimental effects on memory consolidation processes. This result adds to literature about sleep characteristics in insomnia and suggests the importance to consider these additional sleep quality measures when assessing sleep in this population.

The main aim of this research was to address whether individuals with insomnia, in light of their chronic poor sleep, show an impaired sleep effect for false memories. In other words, we expected that, in IN participants, the retention interval spent in sleep relative to wake would not promote false memories formation, at variance with what previously observed in several studies conducted on healthy participants with the same DRM paradigm (e.g., Payne et al., 2009; Diekelmann et al., 2010; Pardilla Delgado & Payne, 2017). In line with this hypothesis, the IN group did not show a higher number of *false recalls* in Sleep compared to Wake, whereas, conversely, a *sleep effect* for *false recalls* emerged in the GS group, confirming previous literature (Payne et al., 2009; Diekelmann et al., 2010; Pardilla Delgado & Payne, 2017).

These data are consistent with theoretical accounts of sleep-related memory processes. In fact, it is believed that a good night of sleep facilitates consolidation of newly acquired memories and their integration into preexisting long-term networks (Landmann et al., 2014; Rash & Born, 2013): this process is thought to induce false memories production by facilitating the extraction of the *gist* trace from the recently encoded information to a greater extent than an equivalent retention period spent awake. Therefore, our first set of results suggests that the chronic poor sleep of individuals suffering with insomnia does not ensure an efficient sleep-related memory reorganization process, consequently reducing false memories production. From a psychobiological point of view, this result could also be related to the impaired secretion of trophic factors involved in memory processes and neuronal plasticity (such as BDNF and IGF-1) that has been observed in sleep disturbances (Giese et al. 2013; Andero et al., 2014). However, future studies are needed to address how specifically these factors affect false memories production in insomnia.

Another interesting remark of our research concerns the “inverse sleep effect” observed in IN subjects, i.e., the significantly higher amount of *false recalls* produced in Wake compared to Sleep. This finding may be explained by taking into account a different mechanism involved in false memories production. In fact, some authors argue that false memories can also arise from cognitive deficits occurring during the memory retrieval phase (Diekelmann et al., 2008; Diekelmann et al., 2010). For example, Diekelmann and colleagues (2008) observed that that sleep-deprived individuals produce more false memories at morning re-test compared to subjects in an undisturbed sleep condition and explained their result by describing false memories as the consequence of an impaired retrieval process due to acute sleep loss (Durmer & Dinges, 2005; Frenda & Fenn, 2016). As for our IN group, it could be the case that the Wake condition, compared to the Sleep one, more clearly highlighted possible diurnal deficits occurring at memory retrieval. Indeed, previous studies reported that false memories can arise as a consequence of impairments in prefrontal functioning (for a review, see Schacter & Slotnick, 2004) that are commonly reported in individuals with insomnia. Specifically, this population appears to show diurnal impairments in the cognitive functions that may ensure efficient memory retrieval and avoid false memories production, i.e. retention and manipulation of information in working memory, inhibitory control, and cognitive flexibility (Fortier-Brochu et al., 2012; Ballesio et al., 2019). To this regard, a recent study from our research group (Malloggi et al., 2021) observed that people suffering from insomnia, in a condition of immediate free recall testing, produced more *false recalls* than good sleepers probably due to a deficit in source monitoring ability. Nevertheless, it could be argued that the performance of the IN group in Wake might also be the consequence of an increased sleepiness occurring during the *DRM learning phase*. However, the same group referred higher motivation in performing the DRM paradigm compared to the GS group, a variable that positively affects cognitive performance (Madan, 2017; Yee & Bravers, 2018) and that could presumably have balanced the increased sleepiness levels. Anyway, although with a significant difference between the two groups, the reported levels of subjective sleepiness are potentially high in both groups ("*more alert than drowsy*" for IN group and "*moderately alert*" in GS group), therefore it is presumable that this variable did not influence the subsequent testing.

Interestingly, in Wake, the IN group produced the same number of *false recalls* that the GS group produced in Sleep. This result can be explained by two different mechanisms possibly involved in false memories production in relation to sleep quality: on one hand, a good sleep may promote false memories production by ensuring an efficient memory consolidation process; on the other hand, false memories can be a consequence of a chronically disturbed sleep, as in insomnia, that has a detrimental impact on the ability to correctly recover previously acquired information. To this regard, a previous work of Diekelmann and colleagues (2010) already focused on these two mechanisms in healthy samples. The authors found that, relative to a diurnal wake condition, false memories increased both in a condition of sleep deprivation, by impacting memory retrieval, and in an undisturbed sleep condition, due to an efficient memory consolidation process. In this perspective, our study adds to previous results by Diekelmann and colleagues (2010) the variable "sleep quality" and suggests the relevance to take into account this variable in the study of the relationship between sleep and false memories production.

At variance with the free recall task, we did not observe between-conditions differences in the number of *false recognitions* either in the IN or GS group. As in a previous study by Pardilla Delgado & Payne (2017), our participants performed the recognition task immediately after the free recall task, so that we cannot rule out an influence of the first task on the latter. However, this result is consistent with extant literature highlighting that the promoting effect of sleep on false memories is generally more evident when adopting a free recall task compared to a recognition task, as underlined in Newbury & Monaghan's meta-analysis (2017).

Concerning veridical memories, we found that sleep, regardless of its quality, promotes the consolidation of contextual, item-specific details of the experience to a greater extent than wake both in the free recall task and in the recognition task. Overall, these findings are in line with literature showing a generic beneficial effect of sleep, relative to wake, on consolidation of declarative memory traces (see, for a review, Diekelmann & Born, 2010). Therefore, it seems that the sleep-dependent consolidation of DRM *studied* words was preserved in the IN group and that poor sleep quality selectively affected only the more refined reorganization process of these declarative memory traces at the free recall test. Moreover, our results on the GS group are in line with those of Payne and colleagues (2009), who reported an increase of *veridical recalls* after undisturbed sleep compared to a wake condition, as a consequence of an efficient sleep dependent memory consolidation process.

However, it could be the case that also the characteristics of the DRM paradigm (i.e. the strong semantic relationship between the *studied* words and *lure* word in each list) might explain the lack of differences emerging between IN and GS groups in *veridical recalls*. In fact, it has been observed that the semantic association between stimuli, such as those observed among the words of the DRM lists, can reinforce memory traces and favor their subsequent retrieval (Aka et al., 2020; Silberman et al., 2005). Therefore, individuals with insomnia symptoms could have achieved at the DRM task the same performance as good sleepers, independently of their sleep quality, thanks to this feature of the task. For the same reason both groups might have been able to avoid to recall words that were semantically unrelated to *studied* words and critical *lure* words, i.e. *recalled intrusions*.

In our study, we also performed correlation analysis in order to verify the relationship between sleep quality measures and DRM performance. In line with our hypothesis, we found that sleep measures preferentially linked to a good sleep quality (Cerasuolo et al., 2020; Conte et al., 2021) were associated with both false memories and veridical memories in the whole sample.

Specifically, we observed that the number of *false recalls* positively correlated with sleep efficiency and negatively with the duration of sleep onset latency and of time spent in light sleep (N2%). Moreover, the number of *false recognitions* positively correlated with actual sleep time and time spent in sleep cycles ($p = .056$), and negatively with frequency of awakenings. Similarly, we found that *veridical recalls* and *hits* were associated to sleep continuity and organization variables (i.e. awakening frequency and, for *hits*, WASO% and number of sleep cycles).

However, our results need to be interpreted in light of a possible limitation concerning data collection. Particularly, the PSG recordings were performed at subject's home. Although we ascertained the adequacy of sleep room through a detailed interview and in the day of the adaptation night, the experimenters were not present during the recording and had not the opportunity to monitor it during the night. However, this methodological choice was adopted to avoid the occurrence of the "reverse first night effect", which may be encountered in individuals with insomnia who reported to sleep better in laboratory compared than in their sleep-room (Toussaint et al., 1995).

In conclusion, this is the first study addressing the sleep effect for false memories in individuals with insomnia. Our results confirmed the role of sleep in false memories production and suggest that false memories preferentially arise as the consequence of a memory consolidation process occurring during an efficient, continuous, stable and well-organized sleep episode. Therefore, the study sheds light on the importance of considering sleep quality when investigating the influence of sleep on this phenomenon.

References

Aikens, J. E., & Rouse, M. E. (2005). Help-seeking for insomnia among adult patients in primary care. The Journal of the American Board of Family Practice, 18(4), 257–261. https://doi.org/10.3122/jabfm.18.4.257

Aka, A., Phan, T. D., & Kahana, M. J. (2020). Predicting recall of words and lists. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication. https://doi.org/10.1037/xlm0000964

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. The International journal of neuroscience, 52(1-2), 29–37. https://doi.org/10.3109/00207459008994241

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders, 5th ed. American Psychiatric Association.

Andero, R., Choi, D.C., Ressler, K.J. (2014). BDNF-TrkB receptor regulation of distributed adult neural plasticity, memory formation, and psychiatric disorders. Progress in Molecular Biology and Translational Science, 122,169–92.

Baglioni, C., Regen, W., Teghen, A., Spiegelhalder, K., Feige, B., Nissen, C., Riemann, D. (2013). Sleep changes in the disorder of insomnia: A meta-analysis of polysomnographic studies. Sleep Medicine Reviews, 18, 195–213. https://doi.org/10.1016/j.smrv.2013.04.001

Ballesio, A., Aquino, M., Kyle, S. D., Ferlazzo, F., & Lombardo, C. (2019). Executive Functions in Insomnia Disorder: A Systematic Review and Exploratory Meta-Analysis. Frontiers in psychology, 10, 101. https://doi.org/10.3389/fpsyg.2019.00101

Bonnet, M. H., Carley, D., Carskadon, M., & Easton, P. (1992). EEG arousals: Scoring rules and examples. Sleep, 15(2), 173–184. http://www.ncbi.nlm.nih.gov/pubmed/11032543.

Castronovo, V., Galbiati, A., Marelli, S., Brombin, C., Cugnata, F., Giarolli, L., Anelli, M. M., Rinaldi, F., & Ferini-Strambi, L. (2016). Validation study of the Italian version of the Insomnia Severity Index (ISI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 37(9), 1517–1524. https://doi.org/10.1007/s10072-016-2620-z

Cellini N. (2017). Memory consolidation in sleep disorders. Sleep medicine reviews, 35, 101–112. https://doi.org/10.1016/j.smrv.2016.09.003

Cerasuolo, M., Conte, F., Cellini, N., Fusco, G., Giganti, F., Malloggi, S. & Ficca, G. (2020). The effect of complex cognitive training on subsequent night sleep. Journal of Sleep Research, 29, e12929

Cipolli, C., Mazzetti, M., & Plazzi. G. (2013). Sleep-dependent memory consolidation in patients with sleep disorders. *Sleep Medicine Reviews*, 17 (2), 91-103.

Conte, F., Carobbi, G., Errico, B. M., & Ficca, G. (2012). The effects of pre-sleep learning on sleep continuity, stability, and organization in elderly individuals. Frontiers in Neurology, 1–9, https ://doi.org/10.3389/fneur.2012.00109

Conte, F., & Ficca, G. (2013). Caveats on psychological models of sleep and memory: A compass in an overgrown scenario. Sleep Medicine Reviews, 17(2), 105–121. https://doi.org/10.1016/j.smrv.2012.04.001

Conte, F., Cerasuolo, M., Fusco, G., Giganti, F., Inserra, I., Malloggi, S., Di Iorio, I., & Ficca, G. (2021). Sleep continuity, stability and organization in good and bad sleepers. Journal of Health Psychology, 26(12), 2131–2142. https://doi.org/10.1177/1359105320903098

Curcio, G., Tempesta, D., Scarlata, S., Marzano, C., Moroni, F., Rossini, P.M., … De Gennaro, L. (2013). Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). Neurological Sciences, 34, 511–519. https://doi.org/10.1007/s1007 2-012- 1085- y

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58, 17-22.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. Nature Reviews Neuroscience, 11(2), 114–126.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. Behavioural Brain Research, 208(2), 425–429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. PLoS ONE, 3(10), e35

Durmer, J.S., & Dinges, D.F. (2005). Neurocognitive consequences of sleep deprivation. Seminars in Neurology, 25, 117–129.

Fortier-Brochu, E., Beaulieu-Bonneau, S., Ivers, H., & Morin, C.M. (2012). Insomnia and daytime cognitive performance: A meta-analysis. Sleep Medicine Reviews, 16, 83–94. https://doi.org/10.1016/j.smrv.2011.03.008

Frenda, S.J., & Fenn, K.M. (2016). Sleep less, think worse: The effect of sleep deprivation on working memory. Journal of Applied Research in Memory and Cognition, 5, 463–469. https://doi.org/10.1016/j.jarmac.2016.10.001

Giese, M., Unternaehrer, E., Brand, S., Calabrese, P., Holsboer-Trachsler, E., Eckert, A. (2013). The interplay of stress and sleep impacts BDNF level. PLoS One, 8, e76050.

Haimov, I., & Shatil, E. (2013). Cognitive training improves sleep quality and cognitive function among older adults with insomnia. Plos ONE, 8(4), https ://doi.org/10.1371/journ al.pone.0061390

Iacullo, V.M., & Marucci, F.S. (2016). Normative data for Italian Deese/Roediger-McDermott lists. Behavior Research Methods, 48, 381–389. https://doi.org/10.3758/s1342 8-015-0582-3

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S.F. (2007). The AASM Manual for the Scoring of Sleep and Associated Events. American Academy of Sleep Medicine, Westchester.

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. Sleep medicine reviews, 18(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Madan, C. R. (2017). Motivated Cognition: Effects of Reward, Emotion, and Other Motivational Factors Across a Variety of Cognitive Domains. Collabra: Psychology. 3:24. doi: https://doi.org/10.1525/collabra.111

Malloggi, S., Conte, F., De Rosa, O., Righi, S., Gronchi, G., Ficca, G., & Giganti, F. (2021). False memories formation is increased in individuals with insomnia. Journal of Sleep Research, 00, e13527. https://doi.org/10.1111/jsr.13527

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. (2012). Interaction of sleep and emotional content on the production of false memories. PLoS ONE, 7(11), e49353

Morin, C. M., LeBlanc, M., Daley, M., Gregoire, J. P., & Mérette, C. (2006). Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep medicine, 7(2), 123–130. https://doi.org/10.1016/j.sleep.2005.08.008

Morin, C.M. (1993). Insomnia, psychological assessment and management. Guilford Press.

Natale, V., Esposito, M., Martoni, M., & Fabbri, M. (2006). Validity of the reduced version of the Morningness-Eveningness Questionnaire. Sleep Biological Rhythms, 4, 72–74. https://doi.org/10.1111/j.1479-8425.2006.00192.x

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. Psychonomic bulletin & review, 26(2), 387–400. https://doi.org/10.3758/s13423-018-1528-4

Ohayon, M. M., & Reynolds, C. F., 3rd (2009). Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). Sleep Medicine, 10(9), 952–960. https://doi.org/10.1016/j.sleep.2009.07.008

Pardilla Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. Neurobiology of Learning and Memory, 137, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., ... Stickgold, R. (2009). The role of sleep in false memory formation. Neurobiology of Learning and Memory, 92(3), 327–334.

Rasch, B., & Born, J. (2013). About sleep's role in memory. Physiological reviews, 93(2), 681–766. https://doi.org/10.1152/physrev.00032.2012.

Roediger H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of experimental psychology: Learning, Memory, and Cognition, 21(4), 803

Salzarulo, P., Formicola, G., Lombardo, P., Gori, S., Rossi, L., Murri, L., & Cipolli, C. (1997). Functional uncertainty, aging and memory processes during sleep. Acta neurologica Belgica, 97(2), 118–122.

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. Neuron, 44(1), 149–160. https://doi.org/10.1016/j.neuron.2004.08.017

Sergeeva, V., Viczko, J., Ray, L. B., Owen, A. M., & Fogel, S. M. (2017). Sleep-dependent motor sequence memory consolidation in individuals with periodic limb movements. Sleep Medicine, 40, 23–32. https://doi.org/10.1016/j.sleep.2017.09.005

Sica, C. & Ghisi, M. (2007). The Italian versions of the Beck Anxiety Inventory and the Beck Depression Inventory-II: Psychometric properties and discriminant power. In M. A. Lange (Ed.), Leading-edge psychological tests and testing research (p. 27–50). Nova Science Publishers.

Silberman, Y., Miikkulainen, R., & Bentin, S. (2005). Associating unseen events: Semantically mediated formation of episodic associations. Psychological Science, 16 161– 166.

Toussaint, M., Luthringer, R., Schaltenbrand, N., Carelli, G., Lainey, E., Jacqmin, A., Muzet, A., & Macher, J. P. (1995). First-night effect in normal subjects and psychiatric inpatients. Sleep, 18(6), 463–469. https://doi.org/10.1093/sleep/18.6.463.

Vignatelli, L., Plazzi, G., Barbato, A., Ferini-Strambi, L., Manni, R., Pompei, F., D'Alessandro, R., & GINSEN (Gruppo Italiano Narcolessia Studio Epidemiologico Nazionale (2003). Italian version of the Epworth sleepiness scale: external validity. Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 23(6), 295–300. https://doi.org/10.1007/s100720300004

Violani, C., Devoto, A., Lucidi, F., Lombardo, C., & Russo, P.M. (2004). Validity of a short insomnia questionnaire: the SDQ. Brain Research Bulletin, 63, 5, 415-421.

Yee D.M., & Braver T.S. (2018). Interactions of motivation and cognitive control, Current Opinion in Behavioral Sciences. 19: 83-90. https://doi.org/10.1016/j.cobeha.2017.11.009.

Appendix

Sleep condition

LIST 1	LIST 2	LIST 3	LIST 4
Finestra	**Fumo**	**Ago**	**Soffice**
Porta	Sigaretta	Filo	Morbido
Specchio	Boccata	Puntura	Duro
Vetro	Vampata	Pagliaio	Cuscino
Avvolgibile	Spire	Punta	Letto
Infisso	Inquinamento	Cucito	Nuvola
Davanzale	Cenere	Bilancia	Lana
Casa	Sigaro	Cucire	Panna
Aperto	Ciminiera	Dolore	Piume
Tenda	Fuoco	Abito	Sfoglia
Intelaiatura	Tabacco	Cammello	Materasso
Vista	Puzzo	Siringa	Molle
Venticello	Pipa	Ditale	Pulcino
Persiana	Polmoni	Iniezione	Piumino
Paravento	Fiamma	Pungere	Schiumoso
Tapparella	Cappa	Sarta	Tenero

LIST 5	LIST 6	LIST 7	LIST 8
Montagna	**Macchina**	**Musica**	**Rabbia**
Collina	Camion	Nota	Ira
Valle	Bus	Suono	Collera
Salita	Treno	Piano	Furore
Cima	Auto	Canzone	Imbestialire
Sommità	Veicolo	Radio	Sdegno
Cumulo	Guida	Complesso	Avversione
Picco	Jeep	Melodia	Odio
Pianura	Ford	Corno	Arrabbiato
Ghiacciaio	Gara	Concerto	Fastidio
Capra	Chiavi	Strumento	Malcontento
Sci	Garage	Sinfonia	Prepotente
Scalatore	Autostrada	Jazz	Calma
Catena	Berlina	Orchestra	Colpa
Ripido	Furgone	Arte	Malumore
Bosco	Taxi	Ritmo	Paura

Wake condition

LIST 1	LIST 2	LIST 3	LIST 4
Dolce	**Sedia**	**Re**	**Freddo**
Caramella	Tavolo	Potere	Caldo
Amaro	Legno	Regina	Neve
Pasticcino	Dondolo	Corona	Tiepido
Cioccolato	Gambe	Cavallo	Inverno
Zucchero	Rotelle	Monarca	Ghiaccio
Miele	Sedere	Cuori	Umidità
Goloso	Paglia	Sudditi	Surgelato
Crostata	Poltrona	Monarchia	Fresco
Acre	Riposo	Sole	Calore
Sapore	Scomoda	Trono	Tempo
Torta	Sdraio	Nudo	Brina
Crema	Sgabello	Elogio	Aria
Gusto	Studio	Impero	Brivido
Frutto	Comoda	Matto	Artico
Salato	Cucina	Oro	Gelo

LIST 5	LIST 6	LIST 7	LIST 8
Lento	**Piede**	**Leone**	**Ladro**
Veloce	Scarpa	Tigre	Furfante
Lumaca	Mano	Circo	Truffatore
Andamento	Dita	Giungla	Rubare
Ballo	Pedata	Domatore	Rapinatore
Treno	Sandali	Tana	Scippo
Adagio	Calcio	Cucciolo	Delinquente
Anziano	Misura	Africa	Scassinare
Calmo	Camminata	Criniera	Bandito
Ritardo	Caviglia	Gabbia	Rapina
Valzer	Braccio	Felino	Assassino
Sereno	Stivale	Ruggito	Prigione
Tartaruga	Alluce	Feroce	Sbirro
Piano	Calzino	Gazzella	Pistola
Formica	Ginocchio	Caccia	Carcere
Pigro	Collo	Branco	Crimine

Chapter 5

Methodological aspects in the relationship between false memories and sleep[4]

5.1. Introduction

Previous studies documented a relationship between sleep and false memories but obtained mixed results. For example, some researchers observed that a retention interval spent asleep, compared to wake, enhances false memories (Payne et al., 2009; Diekelmann et al., 2010; McKeon et al., 2013), whereas others found that it reduces false memories (Fenn et al., 2009; Lo et al., 2014) or even reported no effect on them (Diekelmann et al., 2008). These controversial findings have been explained in the light of methodological incongruences among available studies investigating the effect of sleep on false memories production through the Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995).

The DRM paradigm is the task most commonly adopted to study the effect of sleep on false memories and it is also one of the more frequently employed in laboratory settings (Pardilla DelGado & Payne, 2017; Gallo, 2010). In this task, subjects have to learn several lists (i.e. *DRM learning phase*) containing words (i.e., *studied* words) that are semantically related to a common associate word (i.e., the critical *lure*). After a delay that can be spent in night-sleep, night-wake or diurnal wake, the memory for these words lists is tested with a free recall and/or a recognition task. In this *test phase*, participants frequently remember the critical *lure*, despite it was never presented by the experimenter. For example, they may be presented with a list of words, such as "*bed*, *rest*, *moon*" and so on, and when tested, they probably will remember to have seen/ heard the un-presented critical *lure* word "*sleep*".

It has been argued that some DRM characteristics could influence the false memories production in relation to sleep, for example, the task adopted in the *test phase* (i.e. free recall or recognition task), the number of lists and of words within each lists, the modality presentation or emotionality of the lists (Newbury & Monaghan, 2019). Among them, the task selection represents a relevant

[4] De Rosa O., Malloggi S., Cellini N., Conte F., Giganti F., Ficca G. The role of immediate recall performance in delayed false memories production after sleep and wake. Accepted as Poster presentation at World Sleep 2022 in Rome, Italy (March 11-16, 2022).

methodological issue, since the effect of sleep on false memories appears to vary when using a free recall or a recognition task during the DRM test phase (Newbury & Monaghan, 2019).

In a recent meta-analysis, Newbury and Monaghan (2019) highlighted that the studies using a free recall task globally reported an increase of the number of critical *lure* words recalled after sleep compared to wake condition (see, for example, Payne et al., 2009; Diekelmann et al., 2010; McKeon et al., 2012), whereas this result does not emerge with a recognition task.

Up to now, only one research (Pardilla Delgado & Payne, 2017) investigated the sleep effect for false memories adopting both a free recall and a recognition task in the same study. In this research, the authors found that sleep promotes *false recalls*, whereas no differences between the sleep and the wake conditions emerged at the recognition task. However, as the same authors suggested, the discrepancies in results between the two tasks might be imputable to the order of their presentation, since the free recall was always performed before the recognition task. Indeed, it could be the case that free recall has influenced the performance to the subsequent recognition task, contributing to reduce differences between conditions. Therefore, further research is warranted directly comparing sleep-related false memory performance between a free recall and a recognition task.

Another critical aspect when addressing the relationship between sleep and false memories concerns the evaluation of the baseline measures of memory performance and its influence on subsequent, delayed, testing (after the retention period, which usually spans over several hours in a sleep-memory study). Most of the available studies did not assess baseline DRM performance (i.e., through an immediate recall task) but directly tested it after the retention interval spent in sleep or wakefulness.

However, as in classical sleep-memory studies, the evaluation of baseline performance is critical in order to exclude that performance observed at delayed testing depends on the effectiveness of the early learning of the DRM word lists rather than specifically on consolidation processes occurred over the retention interval (see Conte & Ficca, 2013). In addition, false memories performance at immediate testing may represent a useful index of individual propensity to produce false memories.

Payne and colleagues (2009) attempted to address this issue through a between-subjects design. In a DRM paradigm, an immediate free recall test was administered to a control group, while the Sleep and Wake groups performed the free recall test directly after the retention interval spent asleep or awake, respectively, as in most studies of the field. The authors observed that, compared to the control participants, while *veridical recall* deteriorated in both experimental groups, *false recalls* were reduced only in the Wake group, whereas they were preserved in the Sleep group. Further studies using within-subjects design are warranted to corroborate this interesting finding.

Furthermore, investigating the effect of immediate on delayed testing, e.g., by comparing delayed performance on stimuli that have already undergone immediate testing vs stimuli that have not, may be interesting *per se*. In fact, this procedure could produce a strengthening of the memory traces relatives to the word lists subjected to an immediate testing. To this regard, in declarative memory tasks, it has been observed that immediate testing triggers memory retrieval and enhances the accessibility of memory traces at a subsequent task (Karpicke & Smith, 2012). More specifically, it could be the case that sleep differently consolidates memory traces diversely strengthened, namely in a more deeply way (i.e. immediate testing) or not (i.e. not immediate testing), with different effects also on false memories production.

Thus, the main aim of the study was to address the above-mentioned methodological issues by evaluating, in two separate experiments, the effect of the task and of the immediate testing on false memory production in relation to sleep, introducing also a baseline measure.

In *Experiment 1*, we investigated the effect of a retention period spent asleep vs. awake on delayed recall of a modified DRM paradigm in which, as *baseline* measure, an immediate recall is requested at the DRM *learning phase* for half of the lists, while it is not for the other half. In *Experiment 2*, we investigate the *sleep effect* on DRM performance with the same procedure of *Experiment 1* but adopting a recognition task, and introducing at the *DRM learning phase*, as *baseline* measure, an immediate recognition for half of the lists.

5.2. Materials and method - Experiment 1

5.2.1. Participants

Potential participants were screened through a set of ad hoc questions, administered online through the *Google Form* platform, aimed to collect their demographic data and to assess their general medical conditions, health habits, as well as the presence of psychiatric disorders and sleep problems.

Inclusion criteria were: age 18 – 35; absence of any relevant somatic or psychiatric disorder; no history of drug or alcohol abuse; absence of sleep disorders and of any sleep apnea or respiratory symptom; having a regular sleep-wake pattern (e.g., individuals with irregular study or working habits such as shift-working were excluded); no use of psychoactive medication or alcohol at bedtime.

Finally, 20 individuals (10 M, 10 F; mean age = 22.35 ± 2.43) were recruited for the study in accordance to inclusion criteria. All participants provided informed consent.

5.2.2. Procedure

In a first phase, participants were asked to fill out a web-based set of questionnaires through the *Google Form* platform including: the Pittsburgh Sleep Quality Index (PSQI; Italian version from Curcio et al., 2013), the Insomnia Severity Index (ISI; Italian version from Castronovo et al., 2016), the Beck Depression Inventory II (BDI-II; Italian version from Sica & Ghisi, 2007), and the Beck Anxiety Inventory (BAI; Italian version from Sica & Ghisi, 2007).

In a second phase, we adopted a within-subjects design to evaluate participant's false memories production using a slightly modified version of the Deese – Roediger - McDermott paradigm (Deese, 1959; Roediger & McDermott, 1995) in a Sleep and in a Wake condition. The two conditions were balanced across participants and were carried out at a distance of a week from each other.

In Sleep, subjects' nocturnal sleep was home-monitored using the *Dreem Headband* (DREEM), a portable polysomnographic device. The days before the beginning of the study, participants were trained on how correctly wear the DREEM device and the experimenter also ascertained about its proper montage during the study.

On the day scheduled for Sleep, subjects performed the first phase of the DRM paradigm (i.e. *DRM learning phase*) at 9 PM (± 1 h) and the second part (i.e. *DRM test phase*) 12 hours later, namely at 9 AM (± 1 h). Specifically, just before 9 PM (± 1 h), the experimenter video-called the subject and sent him/her a link on Google Form Platform to perform the *DRM learning* phase. Then, participants were invited to wear the DREEM headband, to start the sleep recording and go to bed. In the morning, experimenter video-called participant again and sent him/her a link on *Google Form* platform to perform the *DRM test* phase.

In Wake, participants performed the *DRM learning phase* at 9 AM (± 1 h) and the *DRM test phase* at 9 PM (± 1 h) of the same day. In these two testing sessions, the experimenter video-called the subject and send him/her a link on *Google Form* Platform. During the retention interval, subjects were allowed to engage in non-cognitively demanding activities, which they later reported in a questionnaire, and were requested to avoid naps.

All sessions were performed at subject's home. Subjects were asked to avoid the assumption of alcohol or other substances possibly influencing sleep during the study and to maintain their normal sleep schedule for the two days prior to the experimental session. For this reason, subjects were also requested to filled a sleep diary during the study.

There was no money or credit compensation for participating in the study. The study was submitted to the local Ethical Committee, which approved the research and certified that the involvement of human participants was performed according to acceptable standards.

5.2.3. Instruments

False memory task

A slightly modified version of the Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995) was adopted to investigate false memories production in the sample. In our study, participants learned 16 DRM word lists, 8 in Sleep and 8 in Wake (see **Appendix**). Each list consisted of 15 words ("*studied*" words) all semantically related to a critical word, defined "*lure*", that was not presented (e.g. "*bed*", "*rest*", "*nap*" with "*sleep*" as *lure* word). As in Roediger and McDermott (1995) and Iacullo and Marucci (2016), the words in each list were presented in order of associative strength with the un-presented *lure* word (from strongest to weakest). Lists were selected from the set already translated in Italian language by Iacullo and Marcucci (2016) and were presented in a balanced order between subjects.

As for task administration, during the *DRM learning phase*, the experimenter read the lists aloud with an interval of 30 s between them and instructed participants to memorize them as accurately as possible. During this phase, participants performed an immediate testing with a free recall (*IT*) on 4 of the lists (lists 2, 4, 6, 8; *Immediate Free Recall lists, "IFR lists"*), in order to obtain a baseline measure of memory performance.

The *DRM test phase* was performed after an interval spent in sleep (Sleep) or diurnal wake (Wake). It included a free recall task, in which participants were asked to write, in a blank page of *Google Form* platform, any words they remembered from all the previously heard lists. They had 10 min to complete this task and were notified when they had 2 min remaining.

Sleep recordings

The DREEM Headband (DH) is a wireless device which records, stores, and automatically analyses physiological data in real time without need of connections (e.g. Bluetooth or Wi-Fi). It records physiological signals via three types of sensors: (1) brain cortical activity through 5 EEG electrodes yielding seven derivations (FpZ-O1, FpZ-O2, FpZ-F7, F8-F7, F7-O1, F8-O2, FpZ-F8; 250 Hz with a 0.4–35 Hz bandpass filter); (2) movements, position, and breathing frequency through a 3D accelerometer located over the head; (5) heart rate via a red-infrared pulse oximeter located in the frontal band (**Figure 1**).

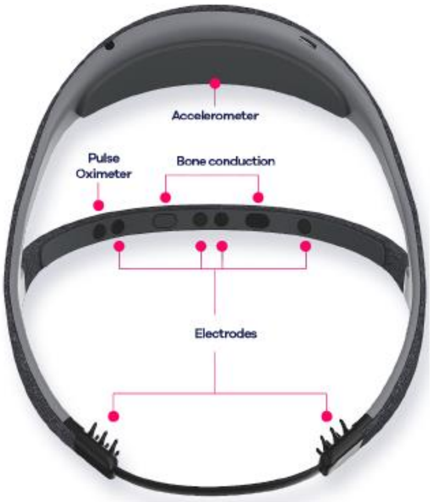

Figure 1. The DREEM headband.

The DREEM is an alternative to the classical polysomnographic recording and has been considered an adequate device for high-quality large-scale longitudinal sleep studies in the home or laboratory settings (Arnal et al., 2020). In particular, it reliably detects EEG signals, breathing frequency and heart rate during sleep; moreover, it provides an accurate automatic sleep staging classification (Arnal et al., 2020) that follows the American Academy of Sleep Medicine guidelines (AASM; Iber et al., 2007).

5.2.4. Data analysis

Outcome measures of the DRM paradigm recollected both at the *DRM learning phase* (on 4 lists) and at *DRM test phase* (on all 8 lists) were:

- the number of *false recalls* (i.e. total number of falsely recalled critical *lure* words);
- the number of *veridical recalls* (i.e. the total number of *studied* words correctly recalled from the original word lists);
- the number of *recalled intrusions* (i.e. the total number of recalled words not corresponding to the *studied* words and to the critical *lure* words).

Moreover, in the *DRM test phase*, the number of *false recalls* and *veridical recalls* has also been calculated separately for the lists subjected to immediate recall (*IFR lists*) and not (*No-Immediate Free Recall lists; N-IFR lists*) during the DRM *learning* phase.

As for objective sleep measures recollected through the DREEM headband, we considered:

- time in bed (*TIB*; i.e. total amount of time, in minutes, from lights off to final awakening);
- total sleep time (*TS*T; i.e. total amount of time, in minutes, from the first appearance of N1 to final awakening);
- sleep-onset latency in minutes (*SOL*);
- percentage of wake after sleep onset over TST (*WASO%*);
- sleep efficiency (i.e. percentage of TST over TIB; *SE%*);
- sleep stage proportions over TST (*N1%, N2%, N3%, REM%*).

Due to non-normal distribution of the data, the baseline measures collected during the DRM *learning phase* (i.e. *false recalls*, *veridical recalls*, *recalled intrusions*) were compared between conditions with a Wilcoxon signed-rank test.

To evaluate performance at the DRM *test phase*, linear mixed models were performed using DRM measures (*false recalls* and *veridical recalls*) as dependent variables, "*condition*" (Sleep vs Wake) and "*list-type*" (*IFR lists* and *N-IFR lists*) as fixed factors, participants ID as random effect, and the factor "IT*"* (i.e. *immediate testing*) as covariate. As for *recalled intrusions*, the factor "*list-type*" was not included in the analysis, because intrusions are not attributable to words belonging to *IFR lists* or *N-IFR lists*. Therefore, a linear mixed models have been performed with "*recalled intrusions"* as dependent variable, "*condition*" as fixed factor, participants ID as random effect, and "*IT"* as covariate. *Holm* correction was adopted for post-hoc tests.

Moreover, *Spearman* correlation analysis was performed to assess the relationship between DRM *test phase* performance and the sleep measures collected in Sleep.

Analyses were performed using JAMOVI (1.6.23.0) and the significance level was set at $p \leq .05$.

5.3. Results – Experiment 1

We excluded two subjects from analyses: the first because he slept < 4 hours and the second one because he did not wear the DREEM headband during the experimental night. Thus, the final sample included 18 subjects (8 M, 10 F; mean age = 22.51 ± 1.88). **Table 1** displays characteristics of the final sample.

Table 1. Circadian preference, BAI, BDI-II, PSQI and ISI total scores in the sample.

MEQr	BAI	BDI-II	PSQI	ISI
13.56 ± 4.11	10.50 ± 8.53	10.06 ± 8.59	4.78 ± 1.96	4.94 ± 3.84

Notes. MEQr = Morningness Eveningness Questionnaire (reduced version); BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; PSQI = Pittsburgh Sleep Quality Index; ISI = Insomnia Severity Index. Mean and standard deviations are reported

Sleep recordings

Table 2 displays sleep measures collected through the DREEM device in Sleep.

Table 2. Sleep measures in the Sleep condition**.**

Sleep measures	*Mean ± standard deviation*
TIB	448.63 ± 68.71
TST	410.58 ± 64.48
SOL	22.09 ± 14.48
WASO	5.90 ± 6.71
SE%	93.33 ± 3.81
N1 %	0.46 ± 0.57
N2 %	47.80 ± 9.19
N3 %	28.78 ± 7.95
REM %	24.86 ± 5.13

Notes. TIB = Time in Bed; TST = Total Sleep Time; SOL = Sleep Onset Latency; WASO = Wake After Sleep Onset; SE = Sleep Efficiency.

Baseline measures of DRM performance at the DRM learning phase

No significant differences between conditions emerged in the number of *false recalls*, *veridical recalls* and *recalled intrusions* recollected during the *DRM learning phase* (**Table 3**).

Table 3. Baseline measures of *false recalls*, *veridical recalls* and *recalled intrusions* in Sleep and Wake.

	Sleep	Wake	Z	p	ES
False recalls	1.28 ± 1.18	1.27 ± 1.18	52.00	>.999	-0.010
Veridical Recalls	29.94 ± 28.50	30.38 ± 30.00	64.00	.358	-0.251
Recalled intrusions	1.83 ± 2.60	0.83 ± 1.10	74.00	.184	-0.410

Notes. *Z* = Wilcoxon signed rank test; *p* = *p*-values. Mean ± standard deviation are reported.

Performance at the DRM test phase

As for *false recalls*, mixed models revealed a significant effect of *list-type* ($F_{1,47.6} = 16.54$, $p < .001$) and *IT* ($F_{1,59.9} = 4.83$, $p = .032$), a result showing that the *IFR lists* more frequently induced subjects to falsely recall critical *lure* words than *N-IFR lists* (t (47.3) = 4.07, $p <.001$) and that the more *false recalls* were produced at immediate recall, the more were recalled at delayed testing.

The interaction *condition* × *list-type* ($F_{1,47.6} = 9.78$, $p = .003$; **Figure 2**) was significant. Specifically, participants produced more *false recalls* at the *IFR lists* in Wake than in Sleep (t (47.3) = - 2.67, $p = .043$), whereas *false recalls* at the *N-IFR lists* were more numerous in Sleep than in Wake (t(47.3) = 1.77, $p = .16$). Moreover, in Wake the number of *false recalls* was higher for *IFR lists* than *N-IFR lists* (t (47.3) = 3.32, $p = .009$).

The *list-type* × *IT* interaction ($F_{1,47.6} = 5.25$, $p = .026$) was also significant, indicating that the more *false recalls* were produced at immediate testing (*IFR lists*), the more the same were retrieved at delayed recall (t (47.6) = - 2.29, $p = .026$).

The effect of *condition* ($F_{1,47.6} = 0.39$, $p = 0.535$), as well as the interaction *condition* × *list-type* × *IT* ($F_{1,47.6} = 0.368$, $p = 0.547$) were not significant. The interaction *condition* × *IT* approached significance ($F_{1,55.5} = 2.52$, $p = .118$), showing that immediate testing increased false memories production at delayed recall especially in Sleep.

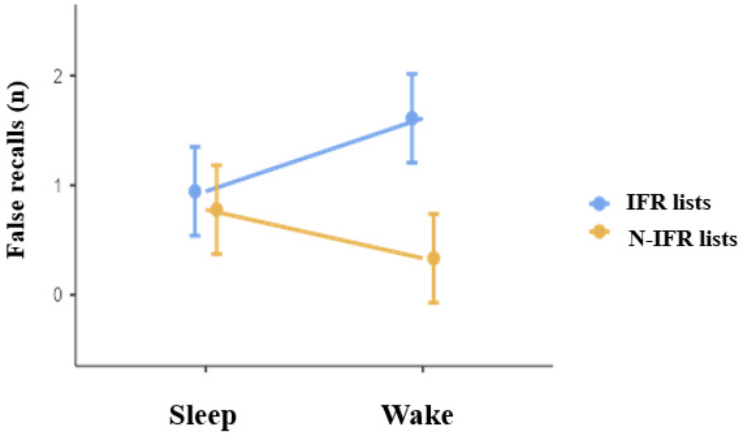

Figure 2. The *condition* × *list-type* interaction. Standard errors are reported.

As for *veridical recalls*, we observed a significant effect of *condition* ($F_{1,15.96}$ = 8.79, *p* =.009), with participants correctly recalling more *studied* words in Sleep compared to Wake (*t* (15.3) = 2.96, *p* =.009; **Figure 3**). Moreover, a significant effect of *list-type* emerged ($F_{1,48}$ = 59.21, *p* =.009), with *IFR lists* more frequently inducing *veridical recalls* compared to *N-IFR lists* (*t* (47.3) = 7.69, *p* <.001). The effect of *IT* was also significant ($F_{1,31.88}$ = 9.05, *p* = .006), indicating that the more *studied* words were recalled at the *DRM learning phase*, the more were correctly recalled at delayed testing (*t* (42.3) = 2.89, *p* =.006). Finally, we observed a significant *list-type* × *IT* interaction ($F_{1,48}$= 4.19, *p* = .048), suggesting an influence of the performance at the *DRM learning* phase especially on *IFR lists* (*t* (46) = - 0.28, *p* =.048). Finally, the interactions *condition* × *list-type* ($F_{1,48}$ = 1.07, *p* = .305) and *condition* × *list-type* × *IT* ($F_{1,48}$ = 1.05, *p* = .309) were not significant.

As for *recalled intrusions*, we found a significant interaction *condition* × *IT* ($F_{1,32}$ = 8.45, *p* = .007; **Figure 4**): the more *recalled intrusions* were produced at immediate recall, the more were produced at delayed testing, especially in the Sleep condition (*t* (32) = -2.07, *p* =.007). The effect of *condition* ($F_{1,32}$ = 1.43, *p* = .240) and of *IT* ($F_{1,32}$ = 0.264, *p* = .611) were not significant.

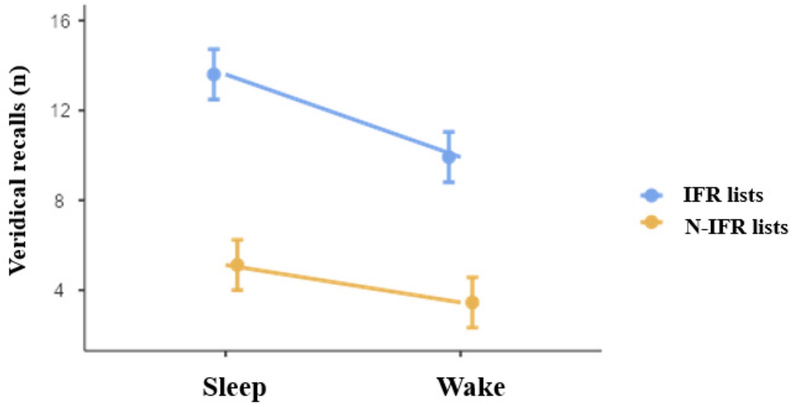

Figure 3. *Veridical recalls* in the Sleep and in the Wake conditions, for *IFR* and *N-IFR lists* separately. Standard errors are reported.

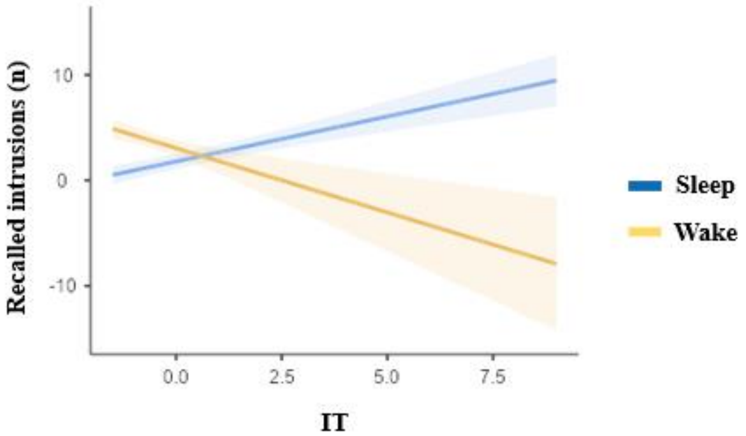

Figure 4. *Condition* × *IT* interaction for the variable *recalled intrusions*. Standard errors are reported. *IT* = immediate testing factor.

Relationship between DRM performance and sleep measures

No significant correlations emerged between DRM variables (delayed recall) and sleep measures.

5.4. Materials and method - Experiment 2

5.4.1. Participants

As in *Experiment 1*, potential participants were screened through a set of ad hoc questions, administered online through the *Google Form* platform, aimed to collect their demographic data and to assess their general medical conditions, health habits, as well as the presence of psychiatric disorders and sleep problems. Inclusion criteria were the same of *Experiment 1*.

Twenty subjects were recruited for the study (11 M, 9F; mean age = 22.45 ± 6.82). The sample was comparable to that of *Experiment 1* in age (t (36) = 0.041, p = .969), gender (χ^2 = 2.06, p = .358), MEQr scores (t (36) = 0.26, p = .793), PSQI scores (t(36) = - 0.779, p = .441), ISI scores (t (36) = - 0.01, p = .996), BAI scores (t (36) = 0.22, p = .823) and BDI-II scores (t (36) = 0.69, p = .489). All participants provided informed consent.

5.4.2. Procedure

The procedure was the same as in *Experiment 1*. The only difference concerned the adoption of a recognition rather than free recall task within the DRM paradigm (see the next paragraph).

There was no money or credit compensation for participating in the study. The study was submitted to the local Ethical Committee, which approved the research and certified that the involvement of human participants was performed according to acceptable standards.

5.4.3. Instruments

False memory task

A slightly modified version of the Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995) was adopted to investigate false memories production in the sample. In detail, participants learned 16 DRM word lists (the same of *Experiment 1;* see **Appendix**), 8 in Sleep condition and 8 in Wake condition respectively. Each list consisted of 15 words (named "*studied*" words) semantically related to a critical word, defined "*lure*", that was not presented by the experimenter (e.g. "*bed*", "*rest*", "*nap*" with "*sleep*" as *lure* word). As in Roediger and McDermott (1995) and Iacullo and Marucci (2016), the words in each list were presented in order of associative strength with the un-presented *lure* word (from strongest to weakest). Lists were selected from the set already translated in Italian language by Iacullo and Marucci (2016) and were presented in a balanced order between subjects.

As for task administration, the experimenter read the lists aloud with an interval of 30 s between them and instructed participants to memorize them as accurately as possible (i.e. *DRM learning phase*). Every two lists (namely, after lists 2,4,6 and 8), participants were requested to perform an immediate testing (*Immediate Testing* factor; "*IT"*) with a recognition task, introduced to recollect a baseline measure of memory performance.

In particular, every two lists, subjects saw on the screen a list of 15 words and were asked to indicate whether they were "*old*" or "*new*" relative to the words previously read by the experimenter. In each session, the 15 words corresponded to 8 *studied* words (serial positions in the list: 1, 3, 5, 7, 9, 11, 13, 15) and the critical *lure* word taken from the last list presented by the experimenter (i.e. 2, 4, 6, 8), together with 6 unrelated words extracted from DRM lists not enrolled in the present study. The order of the 15 words was counterbalanced between subjects.

The *DRM test phase* was performed after an interval spent in sleep (Sleep) or in diurnal wake (Wake) and consisted in a recognition task. In this task, the experimenter provided to subjects a list of 120 words, containing the *lure* words (n = 8), the original *studied* words (n = 60; taken (a) from the lists subjected to immediate recognition ("*Immediate Recognition lists" or "IR lists"*), in serial positions of 2, 4, 6, 8, 10, 12, 14, (b) from lists not submitted to immediate testing *("No-Immediate Recognition lists", or "N-IR lists"*, in serial position 1, 3, 5, 7, 9, 11, 13, 15), and unrelated distractors (n = 52), all stimuli presented in a randomized way. Participants were asked to identify words from the previous lists ("old" words), as well as new words.

Sleep recordings

As in *Experiment 1*, sleep data were recorded through DREEM device. For a detailed description of considered sleep variables, see the paragraph "sleep recordings" of *Experiment 1*.

5.4.4. Data analysis

Outcome measures of the DRM paradigm recollected both at the *DRM learning phase* (on 4 lists) and at *DRM test phase* (on all 8 lists) were:

- *False Memory Rate* (*FMR*), representing the proportion of "old" responses on the total number of critical *lures*;
- *True rate* (*TR*), representing the proportion of "old" responses on the total number of *studied* words (i.e. *hits*);
- *False Alarm Rate* (*FAR*), representing the proportion of "old" responses on the total number of unrelated distractors (i.e. *false alarm*).

Moreover, in the *DRM test phase*, the variables *FMR* and *FAR* have also been calculated separately for the lists subjected to immediate testing ("*IR lists"*) and not ("*N-IR lists"*) during the DRM *learning* phase.

Objective sleep measures were the same of *Experiment 1*.

A *paired samples t-test* has been adopted to compare the baseline measure (*IT*) recollected during the DRM *learning phase* (i.e. *FMR, TR, FAR*) between the Sleep and the Wake conditions.

To evaluate the performance at the *test phase*, linear mixed models have been performed introducing the DRM variables (*FMR* and *TR*) as dependent variables, "*condition*" (Sleep and Wake condition) and "*list-type*" (*IR lists* and *N-IR lists*) as fixed factors, participants ID as random effect, and the factor "*IT"* as covariate. As in *Experiment 1,* since intrusions are not attributable to words belonging neither to *IR* and to *N-IR lists*, a linear mixed models have been performed with "*FAR"* as dependent variable, "*condition*" (i.e. Sleep and Wake condition) as fixed factors, participants ID as random effect, and "*IT"* as covariate. *Holm* correction was adopted for post-hoc tests.

Moreover, *Pearson* correlation analysis was performed to assess the relationship between the performance to the DRM paradigm in the *test phase* and the sleep measures recollected during the Sleep condition.

Analyses were performed using JAMOVI (1.6.23.0) and the significance level was set at $p \leq .05$.

5.5. Results – Experiment 2

Table 4 displays characteristics of the final sample.

Table 4. Circadian preference, total scores in BAI, BDI-II, PSQI and ISI in the sample.

rMEQ	BAI	BDI-II	PSQI	ISI
13.20 ± 4.18	9.90 ± 7.89	8.25 ± 7.34	5.10 ± 2.43	4.95 ± 3.82

Notes. MEQr = Morningness Eveningness Questionnaire (reduced version); BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; PSQI = Pittsburgh Sleep Quality Index; ISI = Insomnia Severity Index. Mean and standard deviations are reported

Sleep recordings

Table 5 displays sleep measures in the Sleep condition collected through the DREEM headband.

Table 5. Sleep measures in the Sleep condition.

Sleep measures	*Mean ± standard deviation*
TBT	453.76 ± 46.56
TST	422.57 ± 55.65
SOL	13.42 ± 11.07
WASO	9.75 ± 17.62
SE%	94.35 ± 7.61
N1 %	0.40 ± 0.57
N2 %	45.32 ± 9.61
N3 %	27.82 ± 7.33
REM %	27.46 ± 6.40

Notes. TIB = Time in Bed; TST = Total Sleep Time; SOL = Sleep Onset Latency; WASO = Wake After Sleep Onset; SE = Sleep Efficiency

Baseline measures of DRM performance

Table 6 displays baseline measures of *FMR*, *TR* and *FAR* in Sleep and Wake condition collected during the *DRM learning phase*. There were no significant differences between conditions for any of these measures.

Table 6. Baseline measures of *FMR*, *TR* and *FAR* in the Sleep and in the Wake condition.

	Sleep	**Wake**	***t***	***p***	***Cohen's d***
FMR	0.73 ± 0.29	0.65 ± 0.36	1.29	.273	0.252
TR	0.88 ± 0.07	0.84 ± 0.07	1.38	.183	0.309
FAR	0.06 ± 0.05	0.07 ± 0.08	- 0.11	.910	- 1.521

Notes. t = *t* Student's test; *p* = *p*-value; FMR = False Memory Rate; TR = True Rate; FAR = False Memory rate. Mean ± standard deviations are reported.

DRM test phase - recognition

As for *FMR*, we observed a significant effect of *IT* ($F_{1,51}$ = 5.02, p = .029), indicating that subjects showing high *FMR* at the *DRM learning phase* more frequently recognized the *lure* words as old, independent of study condition (*t* (51) = 2.24, p = .029). No effect of *condition* ($F_{1,52.6}$ = 0.01, p = .940) and *list-type* ($F_{1,51}$ = 0.27, p = .599) emerged. In addition, the interactions *condition* × *list-type* ($F_{1,50.6}$ = 2.67, p = .108), *condition* × *IT* ($F_{1, 65.7}$ = 1.04, p = .311), *list-type* × *IT* ($F_{1,50.6}$ = 1.57, p = .215) and *condition* × *list-type* × *IT* ($F_{1,50.6}$ = 2.18, p = .143) where not significant. **Figure 5** displays *FMR* in the two conditions.

As for *TR*, we observed a significant effect of *condition* ($F_{1,54.4}$ = 6.38, *p* = .014), with participants correctly recognizing more *studied* words in Sleep compared to Wake (*t*(54.8) = 2.52, *p* = .015). No effect of *list-type* ($F_{1,52.7}$ = 1.18, *p* = .281) or *IT* ($F_{1,71}$ = 0.72, *p* = .399) emerged, as well as the interactions *condition* × *IT* ($F_{1,71.9}$ = 0.31, *p* = .579) or *list-type* × *IT* ($F_{1,52.7}$ = 1.46, *p* = .232). The *condition* × *list-type* ($F_{1, 52.7}$ = 3.19, *p* = .079) and *condition* × *list-type* × *IT* ($F_{1, 52.7}$ = 2.74, *p* =.104) were not significant either (**Figure 6**).

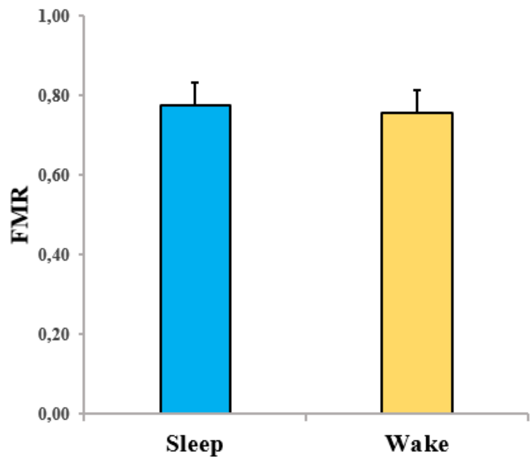

Figure 5. *FMR* in Sleep and Wake condition. Standard errors are reported.

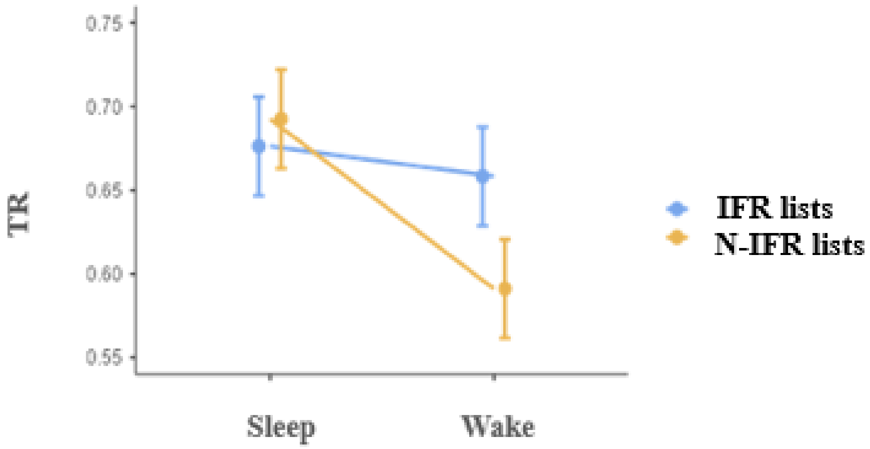

Figure 6. *TR* in the Sleep and Wake conditions, separately for the *IR* and *N-IR* lists. Standard errors are reported.

As for *FAR*, we observed a significant effect of *IT* ($F_{1,33.45} = 9.06$, p =.005), suggesting that the more subjects reported *FAR* at the *learning phase,* the more they erroneously recognized the unrelated distractors as old at delayed testing ($t(35.4) = 3.01$, p = .005). No effect of *condition* ($F_{1,22.09} = 0.38$, p = .543) and no *condition* × *IT* interaction ($F_{1,23.22} = 0.02$, p = .882) emerged. **Figure 7** displays *FAR* in the two conditions.

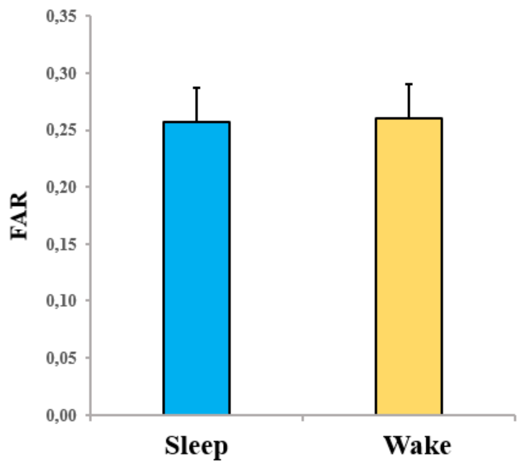

Figure 7. *FAR* in Sleep and Wake. Standard errors are reported.

Relationship between DRM variables and sleep measures

A positive correlation between *FAR* and *N3%* emerged (r = 0.44, p =.050). No other statistically significant correlation emerged.

5.6. Discussion

In this study, we addressed some methodological issues concerning the study of false memories in relation to sleep, by focusing on the influence of the task (i.e. free recall or recognition) and of the immediate testing during the *learning phase* on the DRM outcomes, and taking also into account the baseline performance.

To investigate the influence of the task, we tested the performance to the DRM paradigm, separately, in two homogeneous groups (*Experiment 1* and *Experiment 2*) who learned the same DRM word lists (n = 16; 8 in Sleep and 8 in Wake) but that were tested with two different tasks, namely a free recall (*Experiment 1*) and a recognition task (*Experiment 2*). By doing so, we were able to explore whether the *sleep effect* for false memories varied in function of the task adopted.

From the results, a *sleep effect* for false memories did not emerge in the two experiments when considering the entire set of word lists (8 in Sleep condition vs 8 in Wake condition), since we did not observe differences between the Sleep and the Wake conditions, respectively, in the number of *false recalls* (*Experiment 1)* and in the *FMR* (*Experiment 2*). This result is contrary to our initial expectations. In fact, based on the literature describing false memories as a consequence of the memory *reshaping* occurring during sleep (see, for example, Payne et al., 2009; Diekelmann et al., 2010; Pardilla Delgado & Paye, 2017), we expected to find a *sleep effect* for false memories at the free recall task at least, a task that appears more sensible to the *gist* trace extraction (Newbury & Monaghan, 2019).

However, it could be the case that the absence of differences between the two conditions in both the experiments is due to our methodological choice of testing some of the DRM lists immediately after their presentation. Specifically, when considering only the lists subjected to delayed testing (i.e., with no immediate testing required), the pattern of results emerging from both our experiments is consistent with extant researches on sleep and false memories conducted with the same procedure (i.e., with testing directly performed after the retention interval spent asleep or awake). In fact, on one hand, in *Experiment 1* we observed an increase of *false recalls* after sleep relative to wake (although quite significant) for *N-IFR* lists, as in Payne and colleagues (2009), who adopt a free-recall test procedure; on the other hand, in *Experiment 2* (recognition study) we found no between-conditions differences in *FMR* for lists not subjected to immediate testing, as in Pardilla-DelGado & Payne, (2017) and Diekelmann and colleagues (2008), who included a recognition testing. In this regard, it has been observed that studies adopting a free recall task generally report an increase of the number of critical *lure* words recalled after sleep compared to wake (see, for example, Payne et al., 2009; Diekelmann et al., 2010; McKeon et al., 2012), whereas this result does not emerge with recognition tasks, as highlighted by Newbury and Monaghan's meta-analysis (2017).

Concerning the lists subjected to immediate testing (*IFR lists* and *IR lists*), our results are different in *Experiment 1* and *2*. In *Experiment 1*, participants produced less *false recalls* for the *IFR lists* in the Sleep condition with respect to the Wake one, whereas in *Experiment 2* no between conditions differences emerged in false recognitions rate.

The pattern of findings from *Experiment 1* suggests that the presence or absence of immediate testing not only strengthens memory traces in a different way but also affects the subsequent processing of memories depending on the behavioural state in which it takes place.

Our results can be explained in the light of the *Fuzzy Trace Theory* (FFT; Brainerd et al., 1995), which provides a theoretical framework to explain false memories production at the DRM paradigm. According to this theory, subjects simultaneously encode two independent traces for each word, respectively the "*verbatim trace*" (i.e. the trace related to the contextual features of a word and especially linked to veridical memory, corresponding in the DRM paradigm to the

“*studied* words”) and the “*gist*” or “*fuzzy*” trace (i.e. the trace representing the meaning of an item, preferentially linked to false memories production). In our study, the immediate recall of the 4 words lists could have led subjects to particularly focus on item-specific features of the stimuli, encoding a stronger “*verbatim*” trace respect to the “*gist*” one. Consequently, sleep might have preferentially consolidated the “*verbatim*” trace at the expense of “*gist*”, not facilitating the recall of *lure* words at subsequent testing.
On the other hand, in the Wake condition, the “*gist*” trace originally encoded during the *DRM learning* phase appears resistant to decay, a result in line with the previous literature showing that false memories are more persistent than veridical ones in time and across long delays (Brainerd et al., 1995; Reyna & Lloyd, 1997).

On the contrary, as for *N-IFR lists* in *Experiment 1*, sleep could have strengthened the weak memory traces relatives to lists not subjected to immediate recall, consequently favoring the abstraction of the *gist* and the recall of *lure* words at morning testing. The result seems to support the idea that weak traces are the ones targeted by sleep i.e., only participants that weakly encoded the material received a benefit from sleep (Drosopoulos et al., 2007; Kuriyama et al., 2004). Therefore, it is plausible that our participants relied more heavily on gist/semantic processing of *N-IFR lists* words due to a sleep-dependent benefit in *gist* extraction. On the contrary, in the Wake condition, the weaker memory trace for *N-IFR* lists, initially induced by the lack of immediate testing during the *DRM learning phase*, could have hampered the retrieval of *lure* words after a long delay.

An alternative interpretation concerns the activation of source monitoring processes occurring at the times of testing. It can be hypothesized that immediate testing triggers during the *DRM learning phase* a source monitoring process which, in case of word lists directly recalled after the delay, would only be activated after the retention interval. In this case, it could be that word lists already subjected to source monitoring at immediate testing undergo some kind of labelling which prevents further “*gist* extraction” processes during sleep but not during wake. In other words, sleep would preferentially benefit the “*verbatim* trace” when the accuracy of memories has already been “checked”, regardless of outcome accuracy of the source monitoring process.

Our results also highlighted that the baseline performance is an additional factor possibly influencing the production of false memories in relation to sleep. In fact, we observed that more false memories participants produced at the *DRM learning phase*, the more they retrieved at the subsequent test phase in both Experiments. In addition, we found that this effect is particularly evident in the Sleep condition of *Experiment 1*, a result suggesting that sleep could particularly

benefit the consolidation of *gist* traces already formed during the *DRM learning phase* when adopting a free recall testing procedure.
In parallel, the baseline performance could also be interpreted as a measure of propensity to produce false memories, influenced by various individual factors, such as imagery and perceptual ability, executive functioning efficiency, or personality characteristics (Chen et al., 2010; Zhou et al., 2010). In this perspective, our result also suggests that future studies should take into account also this individual factor when evaluating the effect of sleep on false memories production.

As for the results obtained on false recognitions in *Experiment 2* (i.e., no preferential effect of sleep over wake regardless of the presence of immediate recall), the finding is probably linked to the nature of the task, which is less susceptible to produce memory errors compared to the free recall procedure (Cabeza et al., 2001; Tulving & Madigan, 1970). Specifically, in the recognition task subjects are explicitly asked to discriminate between "old" and "new" items among a pool of stimuli provided by the experimenter. Conversely, the free recall task does not provide at retrieval any external memory cues, so that subjects have to generate their own cues to reinstate the original memory traces. This process of "self-cueing" is at risk of inducing memory errors. Instead, the presentation of words in a recognition procedure would presumably reinstate the context of encoding and favour the retrieval of the details of the *studied* words consolidated during sleep. This process would simplify the discrimination between *studied* and *lure* words, preventing false recognitions.

Regarding veridical memories, our results clearly highlight that sleep, relative to wake, protects the *studied* words from decay both in the free recall and the recognition task, regardless of the presence of immediate testing. This finding is in line with previous studies adopting the DRM paradigm (Payne et al., 2009; McKeon et al., 2013; Pardilla Delgado & Payne, 2017), as well as with the general, robust literature on sleep and declarative memory (for a review, see Conte & Ficca, 2013).

A further remark concerns our finding of a very consistent effect of immediate testing on delayed performance, regardless of the behavioural state in which consolidation took place. The presence of immediate testing influenced all memory performance variables in both conditions and at both free recall and recognition, except for *true recognitions* in *Experiment 2*. Indeed, in most cases, the more words of a certain type (*lures, studied* or *intrusions*) were produced at immediate testing, the more words of the same type were produced at delayed testing. This is compatible with common theories on the strength of the memory trace (Conte & Ficca, 2013), since immediate testing has presumably reinforced the labile traces formed during the presentation of the lists.

Finally, we also evaluated the relationship between sleep measures and the performance of the DRM paradigm. We found a positive correlation between N3% and *FAR*, a result possibly suggesting the occurrence of a general reorganization of previously acquired information during deeper sleep (Landmann et al., 2013). At variance with previous studies (Payne et al., 2009; Pardilla Delgado & Payne, 2017), we did not observe other significant correlations between DRM performance and sleep measures. This result is probably due to the small number of DRM words lists introduced in the present study and to the slightly different procedure we adopted, namely the presence of immediate testing performed on some of the lists. In addition, sleep recordings were not performed through standard PSG, as in the above-mentioned studies, but with a different portable device (although validated and reliable in its outcomes).

The results of this study need to be interpreted in the light of some limits. First, the immediate testing on half of the list could have reduced the emergence of a clearer *sleep effect* for *N-IFR lists* in *Experiment 1*, due to the small number of lists tested. However, this procedure permitted us to recollect, in parallel, a baseline measure of the DRM performance. To better address the effect of the baseline performance on false memories production after sleep, future studies should introduce an immediate free recall testing after the presentation of each DRM list, also possibly incrementing the number of words lists. Moreover, the sleep recordings have been performed at participants' homes through a portable device that they wear the night of the experimental condition. This procedure could be subjected to some bias regarding for example the impossibility for the experimenter to control the sleep recording and the proper positioning of the device during the night.

In conclusion, in this study, we tried to clarify the influence of sleep on false memories production by overcoming some methodological issues of previous researches in the literature. Our results underline the importance of cautiously evaluating the introduction of immediate testing in any investigation of false memories formation over long delays. Specifically, our findings point to a different sleep- and wake-related memory re-processing, over the retention period, depending on the presence or absence of an immediate free recall test, whereas this phenomenon appears not to be evidenced when testing is conducted through recognition rather than free recall paradigms.

References

Arnal, P. J., Thorey, V., Debellemaniere, E., Ballard, M. E., Bou Hernandez, A., Guillot, A., Jourde, H., Harris, M., Guillard, M., Van Beers, P., Chennaoui, M., & Sauvet, F. (2020). The Dreem Headband compared to polysomnography for electroencephalographic signal acquisition and sleep staging. Sleep, 43(11), zsaa097. https://doi.org/10.1093/sleep/zsaa097

Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. Journal of Memory and Language, 34(2), 157–185. https://doi.org/10.1006/jmla.1995.1008

Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R., & Schacter, D.L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. PNAS, 98, 4805–10.

Castronovo, V., Galbiati, A., Marelli, S., Brombin, C., Cugnata, F., Giarolli, L., Anelli, M. M., Rinaldi, F., & Ferini-Strambi, L. (2016). Validation study of the Italian version of the Insomnia Severity Index (ISI). Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 37(9), 1517–1524. https://doi.org/10.1007/s10072-016-2620-z

Chen, C.; Zhu, B.; Loftus, EF; Lin, C.; He, Q.; Chen, C., et al. (2010). Individual differences in false memory from misinformation: personality characteristics and their interactions with cognitive abilities. Personality and Individual Differences, 48(887-894). Retrieved from https://escholarship.org/uc/item/6jz8k98w

Conte, F., & Ficca, G. (2013). Caveats on psychological models of sleep and memory: A compass in an overgrown scenario. Sleep Medicine Reviews, 17(2), 105–121. https://doi.org/10.1016/j.smrv.2012.04.001

Curcio, G., Tempesta, D., Scarlata, S., Marzano, C., Moroni, F., Rossini, P.M., … De Gennaro, L. (2013). Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). Neurological Sciences, 34, 511–519. https://doi.org/10.1007/s1007 2-012- 1085- y

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58, 17-22.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. Nature Reviews Neuroscience, 11(2), 114–126.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. Behavioural Brain Research, 208(2), 425–429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. PLoS ONE, 3(10), e35

Drosopoulos S, Schulze C, Fischer S, Born J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. J Exp Psychol.136(2):169–83.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. Learning & Memory, 16(9), 509–513.

Gallo, D. A. False memories and fantastic beliefs: 15 years of the DRM illusion. Memory and Cognition 38 (7), 833–848 (2010).

Iacullo, V.M., & Marucci, F.S. (2016). Normative data for Italian Deese/Roediger-McDermott lists. Behavior Research Methods, 48, 381–389. https://doi.org/10.3758/s1342 8-015-0582-3

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S.F. (2007). The AASM Manual for the Scoring of Sleep and Associated Events. American Academy of Sleep Medicine, Westchester.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. Journal of Memory and Language, 67, 17-29.

Kuriyama K, Stickgold R, Walker MP. (2004). Sleep-dependent learning and motor-skill complexity. Learning and Memory, 11(6):705–13.

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. Sleep medicine reviews, 18(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Lo, J. C., Sim, S. K., & Chee, M. W. (2014). Sleep reduces false memory in healthy older adults. Sleep, 37(4), U247–U665

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. (2012). Interaction of sleep and emotional content on the production of false memories. PLoS ONE, 7(11), e49353

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. Psychonomic bulletin & review, 26(2), 387–400. https://doi.org/10.3758/s13423-018-1528-4

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. Neurobiology of Learning and Memory, 137, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., ... Stickgold, R. (2009). The role of sleep in false memory formation. Neurobiology of Learning and Memory, 92(3), 327–334.

Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. Learning and Individual Differences, 9(2), 95–123. https://doi.org/10.1016/S1041-6080(97)90002-9

Roediger H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of experimental psychology: Learning, Memory, and Cognition, 21(4), 803

Sica, C. & Ghisi, M. (2007). The Italian versions of the Beck Anxiety Inventory and the Beck Depression Inventory-II: Psychometric properties and discriminant power. In M. A. Lange (Ed.), Leading-edge psychological tests and testing research (p. 27–50). Nova Science Publishers

Tulving, E., & Madigan, S.A. (1970). Memory and verbal learning. Annual Review of Psychology, 21, 437–84.

Zhu, B., Chen, C., Loftus, E. F., Lin, C., He, Q., Chen, C., Li, H., Xue, G., Lu, Z., & Dong, Q. (2010). Individual differences in false memory from misinformation: cognitive factors. Memory (Hove, England), 18(5), 543–555. https://doi.org/10.1080/09658211.2010.487051

Appendix

Sleep condition (Experiment 1 and Experiment 2)

LIST 1	LIST 2	LIST 3	LIST 4
Alto	**Bandiera**	**Dolce**	**Fiume**
Basso	Italia	Caramella	Foce
Magro	Rossa	Amaro	Affluente
Palazzo	Tricolore	Pasticcino	Scorrere
Adige	Vento	Tenero	Diga
Dritto	Nazione	Zucchero	Pescare
Fragile	Bianca	Miele	Ponte
Medio	Alza	Goloso	Torrente
Gigante	Asta	Crostata	Corrente
Muro	Colore	Acre	Flusso
Grattacielo	Italiana	Sapore	Pesce
Scaffale	Strisce	Torta	Arno
Metro	Cuore	Crema	Lungo
Grado	Mondiali	Panna	Valle
Palo	Innalzare	Frutto	Detriti
Torre	Inno	Salato	Tevere

LIST 5	LIST 6	LIST 7	LIST 8
Ladro	**Lampada**	**Musica**	**Sedia**
Furfante	Luce	Concerto	Tavolo
Truffatore	Lume	Pentagramma	Legno
Rubare	Sfregare	Radio	Dondolo
Rapinatore	Accesa	Ritmo	Gambe
Scippo	Spina	Melodia	Rotelle
Delinquente	Voce	Sassofono	Sedere
Scassinare	Volta	Strumento	Paglia
Bandito	Desiderio	Nota	Poltrona
Rapina	Ragione	Sinfonia	Riposo
Assassino	Calore	Arpa	Scomoda
Prigione	Alogena	Disco	Sdraio
Sbirro	Lampadina	Pianoforte	Sgabello
Pistola	Aladino	Suono	Studio
Carcere	Neon	Chitarra	Comoda
Crimine	Genio	Orchestra	Cucina

Wake condition (Experiment 1 and Experiment 2)

LIST 1	LIST 2	LIST 3	LIST 4
Freddo	**Giustizia**	**Re**	**Ago**
Ghiaccio	Polizia	Potere	Filo
Guanti	Salto	Regina	Puntura
Vento	Colle	Corona	Pagliaio
Coperta	Colpa	Cavallo	Punta
Metallico	Crimine	Monarca	Cucito
Montagna	Cassazione	Cuori	Bilancia
Neve	Giudice	Sudditi	Cucire
Pioggia	Privata	Monarchia	Dolore
Monte	Tribunale	Sole	Abito
Tiepido	Equità	Trono	Cammello
Coperto	Divina	Nudo	Siringa
Maglione	Ingiusta	Elogio	Ditale
Naso	Legge	Impero	Iniezione
Corrente	Martello	Matto	Pungere
Sciare	Utopia	Oro	Sarta

LIST 5	LIST 6	LIST 7	LIST 8
Uomo	**Penna**	**Fumo**	**Lento**
Donna	Inchiostro	Pipa	Veloce
Barba	Scrivere	Sigaro	Lumaca
Forte	Matita	Sigaretta	Andamento
Genere	Foglio	Tabacco	Ballo
Persona	Biro	Cenere	Treno
Adamo	Calamaio	Fumare	Adagio
Onore	Nera	Nicotina	Anziano
Bello	Piuma	Polmoni	Calmo
Capelli	Scrittura	Camino	Ritardo
Forza	Scuola	Inquinamento	Valzer
Virile	Sfera	Cancro	Sereno
Cravatta	Stilo	Vapore	Tartaruga
Bambino	Astuccio	Puzza	Piano
Grande	Grafia	Accendere	Formica
Maschile	Oca	Arrabbiato	Pigro

Conclusions

The first three studies presented in this thesis introduced the variable "sleep quality" as a useful approach to investigate the relationship between sleep and false memories, and focused on the disorder of insomnia, actually representing one of the most relevant sleep disorders in the general population. The insomnia disorder allows to study the influence of sleep quality on false memories assuming two different perspectives: on one hand, assessing whether a chronic poor sleep, by impacting on frontally-mediated executive functioning, could increase the subjects' susceptibility to commit source monitoring errors and to produce false memories; on the other hand, investigating whether a poor sleep affects the process of memory reorganization expected to occur during sleep, with consequences on false memories production.

Both results from studies presented in **Chapters 2** and **3** support our hypothesis that false memories can represent one of the consequences of poor sleep and that the efficiency of executive functioning may significantly influence the production of false memories. Furthermore, our results highlighted that the characteristics of the stimuli adopted to assess false memories production may represent an additional determining factor when studying this phenomenon. In fact, we observed that individuals with insomnia produced more false memories for the sleep-related word list relative to the neutral one, a result documenting a different allocation of attentional resources to sleep-related stimuli in this sample when performing the DRM paradigm.

Findings from the study presented in **Chapter 4**, where we conceptualized false memories as the consequence of a sleep-dependent memory reorganization process, confirmed the role of sleep in false memories production, by documenting a *sleep effect* for false memories only in good sleepers and not in individuals with insomnia. This result globally suggests that false memories preferentially arise as the consequence of a reorganization of memory traces occurring during an efficient, continuous, stable, and well-organized sleep episode.

Finally, addressing the methodological aspects of the relationship between sleep and false memories, the results from the last study underline the importance of cautiously evaluating the introduction of immediate testing in any investigation of false memories production aver long delays. Specifically, our findings point to a different sleep- and wake-related memory re-processing, over the retention period, depending on the presence or absence of an immediate free recall test, whereas this phenomenon appears not to be evidenced when testing is conducted through recognition rather than free recall paradigms. Furthermore, our results provide some evidence that individual differences might modulate memory performance, suggesting also the importance of introducing a baseline measure in studies addressing the relationship between sleep and false memories production.

In conclusion, the present thesis contributes to shed light on factors possibly influencing the effect of sleep on false memories production. In this perspective, our findings could be useful for future studies addressing the role of sleep on consolidation and qualitative reorganization, as well as on recovery of memory traces.

Acknowledgements

At the end of this long and significant academic journey, I would like to sincerely thank the people without whom the fulfilment of this thesis would have not been possible:

- The supervisor of this thesis, Prof. Fiorenza Giganti, who has patiently followed my work over the past three years, believing in my abilities and providing me, thanks to her valuable suggestions, with all the tools I need to continue enhancing my research work;

- Prof. Gianluca Ficca, who has always provided a precious contribution to the research projects of this thesis, both from a theoretical and methodological point of view. His remarks have consistently offered great insight into the topics examined, and have also represented inspiration for future researches;

- Prof. Francesca Conte. Despite the geographical distance, she has been close to me in carrying out my research, and also suggested innovative perspectives in understanding psychological phenomena. In addition, she was always ready to help me when the working load was heavy. I hope one day to be able to return, at the smallest extent at least, this great favour!

- Prof. Maria Pia Viggiano and the whole research group of the CogPsy Lab of the University of Florence. They provided me stimulating theoretical insights and were always available to help me in carrying out my research;

- Prof. Nicola Cellini, who I would like to thank for his precious support in statistical analyses and for his insightful comments on my research;

- To the colleagues and collaborators who helped me with data collection despite the very difficult pandemic period.

At last, but not least, I would like to sincerely express gratitude to my parents and family for the support they have always given me, even in the most critical times. I would also like to thank my lifelong friends and Marco, who always encouraged me and provided valuable advices.