INTRODUCTION

Impaired sleep is highly prevalent in the general population (Morin, Bootzin, LeBlanc, Daley, Gregoire, & Mérette, 2006; Ohayon, 2002), and is associated with substantial disease burden (Kyle, Morgan, & Espie, 2010; Weinberg, Noble, & Hammond, 2016) and economic costs (Daley, Morin, LeBlanc, Grégoire, & Savard, 2009; Léger & Bayon, 2010). It is a feature of both clinically significant insomnia and non-clinically significant sleep disturbances. Insomnia can be defined as: (a) sleep difficulties concerning quantity and quality; (b) despite adequate opportunity and circumstances for sleep; (c) for at least 1 month; (d) resulting in specific daytime impairment related to the nighttime sleep difficulty (Buysse, Ancoli-Israel, Edinger, Lichstein, & Morin, 2006; Edinger et al., 2004). Sleep
disturbances include cases in which symptoms of insomnia do not reach clinical significance, as well as cases in which there is not enough information present to diagnose insomnia.

College students in particular are at risk for experiencing impaired sleep, in large part due to their high levels of perceived stress; likewise, an epidemiological study found that 60% of college students surveyed reported experiencing sleep disturbances (Lund, Reider, Whiting, & Prichard, 2010). The prevalence of clinical insomnia in college students reaches nearly 10% (Schlarb, Kulessa, & Gulewitsch, 2012; Taylor et al., 2011). As emerging adults trying to earn academic degrees, young college students face numerous new challenges, such as establishing identity and future life goals, reaching momentous life decisions, living independently, handling finances, managing academic demands, and modifying existing and adopting new social roles (Arnett, 2000; Brougham, Zail, Mendoza, & Miller, 2009; Pierceall & Keim, 2007). In contrast to other young adults, college students face a less prestructured everyday life due to self-organized schedules. In this respect, evidence shows that college students often shift to an irregular sleep–wake cycle by short sleep length on weekdays and later wake-up times on weekends (Buboltz, Brown, & Soper, 2001; Machado, Varelle, & Andrade, 1998) due to, for example, long learning hours, parties and noise pollution in student residences.

Given this transition phase in life, there are several reasons to assume that the efficacy of psychological interventions to improve sleep might differ between college students and the general population. Young adults (ages 19–29 years) show especially high strain in their everyday lives due to sleep disturbances and the worst sleep hygiene (SH) behaviour (Rosenberg et al., 2011). More severe sleep disturbances have been shown to be a predictor for larger sleep improvements within the scope of psychological intervention (Espie, Inglis, & Harvey, 2001; van Houdenhove, Buyse, Gabriëls, & van den Bergh, 2011; Murwaski, Wade, Plotnikoff, Lubans, & Duncan, 2018). Furthermore, psychological interventions aimed at changing beliefs and behaviour patterns might be more effective for college students than for other populations. Because cognitive flexibility is correlated with intelligence (Colzato, van Wouwe, Lavender, & Hommel, 2006) – which in turn is fostered by education (Ritchie & Tucker-Drob, 2018) – and decreases with increasing age (Peltz, Gratton, & Fabiani, 2011), college students may have above average cognitive flexibility. This increased cognitive flexibility may lead to improved response to psychological intervention. Finally, research has found increasing evidence that early intervention predicts better treatment outcomes for a variety of psychological disorders (McGorry, Purcell, Goldstone, & Amminger, 2011). Because sleep disturbances and insomnia often begin during the college years, this is an essential time for intervention (Lund et al., 2010).

Insomnia and sleep disturbances have been shown to be a precipitating factor for incidence of mental illness (Harvey, Murray, Chandler, & Soehner, 2011; Pigeon, Bishop, & Krueger, 2017). People suffering from insomnia show a two- to threefold (odds ratio $[OR] = 1.98–2.98$) increased risk for depression, a significantly greater risk for anxiety disorders ($[OR] = 1.63–2.64$), and a significantly greater risk for substance abuse disorders, including alcohol abuse (Baglioni et al., 2011; Pigeon et al., 2017; Sivertsen et al., 2014). Insomnia has also been reported to be a significant risk factor for medical conditions like arthritis ($[OR] = 1.87$), whiplash ($[OR] = 1.71$), arthrosis ($[OR] = 1.68$), headache ($[OR] = 1.50$), asthma ($[OR] = 1.47$) and myocardial infarction ($[OR] = 1.46$; Sivertsen et al., 2014). Studies investigating the mental and physical health consequences of insomnia and sleep disturbances in college students specifically provide results that indicate similar deleterious effects (Hershner & Chervin, 2014; Lund et al., 2010; Taylor & Bramoweth, 2010; Taylor et al., 2011; Wong et al., 2013). Moreover, sleep disturbances and insomnia in college students are associated with personal distress (Meerlo, Sgoifo, & Suchecki, 2008; Rosenberg et al., 2011), lower quality of life, lower academic achievement, and impaired declarative and procedural memory performance (Cucic, Ferrara, & de Gennaro, 2006; Gomes, Tavares, & de Azevedo, 2011; Kelly, Kelly, & Clanton, 2001; Wong et al., 2013). Furthermore, sleep disturbances increase the risk of self- and third-party endangerment. For example, 16% of an undergraduate sample reported having fallen asleep while driving at least once in their lifetime (Hershner & Chervin, 2014; Taylor & Bramoweth, 2010).

Cognitive behaviour therapy for insomnia (CBT-I), including stimulus control, sleep restriction, progressive muscle relaxation, SH and cognitive therapy, meets the American Psychological Association’s criteria for empirically validated psychological interventions for insomnia (Morgenthaler et al., 2006; Morin, Bootzin, et al., 2006). CBT-I is recommended for the general population as first-line treatment for insomnia (Qaseem, Kansagara, Forciea, Cooke, & Denberg, 2016; Riemann et al., 2017), as it has been shown to be effective in improving sleep quality (SQ) and quantity in adult samples (van Straten et al., 2018).

Despite this, little is known about the efficacy of psychological interventions among college students, as there are no meta-analyses available referencing this specific population. Thus, investigating treatment response in college students is of particular importance due to their increased vulnerability to sleep disturbances and the negative health outcomes they experience as a result. Given that emerging adulthood – including college years – represents a time of major life transition and the highest risk of developing mental disorders, reaching emerging adults through interventions is of paramount importance (Auerbach et al., 2018; Kessler, Berglund, et al., 2018).
College students experiencing a variety of stressors show an increased onset of mental health problems (Pedrelli, Nyer, Yeung, Zuluf, & Wilens, 2015; Thurber & Walton, 2012), with over 20% of all college students suffering from a mental disorder (Auerbach et al., 2018). This high prevalence depicts the elevated vulnerability for mental disorders at a time of major life transitions, that also influences further life of college students considerably, as this critical time period is substantial for basic life events as, for example, attaining educational levels/degrees (Auerbach et al., 2018; Bruffaerts et al., 2018). The number of affected college students far exceeds the resources of most treatment options at universities, resulting in substantial treatment gaps of mental health issues among college students (Auerbach et al., 2018; Beiter et al., 2015).

Further, college students are a key group in society in terms of human capital (Abel & Deitz, 2012), driving future societal economic growth and innovation. Higher education institutions provide an opportunity to become a key setting for the prevention and early treatment of sleep disturbances and insomnia, with a high reachability of over 50% of emerging adults being enrolled in higher education (Aud et al., 2011; Reavley & Jorm, 2010). Beyond that, because sleep disturbances are a common complaint that lack the stigma associated with other mental health issues, psychological interventions aimed at improving sleep may provide an acceptable low-threshold opportunity for a first step in a care pathway that have the potential to positively affect other co-morbid mental and physical health issues (Freeman et al., 2017; Friedrich & Schlarb, 2017; Wu, Appleman, Salazar, & Ong, 2015).

To the best of our knowledge, there are no meta-analyses available so far, focusing on psychological sleep treatments in college students. To date, two descriptive reviews have synthesized findings on psychological interventions to overcome sleep disturbances in college students. One of these reviews (Dietrich, Francis-Jimenez, Knibbs, Umali, & Truglio-Londrigan, 2016) focuses on education programmes to improve sleep. SH education in varying formats was identified as the main intervention form in the articles included in the review. The interventions on SH education differed in their format with regard to the integration of additional content, delivery strategies, length and duration of the programme. Studies were included that meet the primary outcome measures of SH knowledge, SH behaviour or SQ. However, because such programmes are a poorly explored subsection of psychological sleep interventions (at least among college students), this review includes only three randomized controlled trials (RCTs) and one quasi-experimental study. Significant effects of sleep education on SH knowledge and behaviour, respectively, were found in one of two studies, and effects on SQ were found in one of four studies. Thus, this review only provides preliminary evidence, and does not allow to draw conclusions on the overall effectiveness of sleep education interventions in college students. Furthermore, the review focuses on sleep education only, without targeting and analysing all types of psychological sleep interventions. Friedrich and Schlarb (2017) examined 27 evaluation studies on sleep treatments that included psychological components. The treatments were categorized in SH interventions, CBT-I, relaxation and other treatments (e.g. gestalt therapy or imagery rehearsal therapy). They found small to moderate effect sizes for SH interventions ($d = 0.32–0.61$), large effects for CBT-I ($d = 1.06–1.77$), and moderate effects for other psychological interventions ($d = 0.45–0.61$). Fifteen out of the 27 studies were RCTs, seven studies used a controlled design without randomization, and five trials had no control condition. Because it was the first review on the topic of sleep disturbances in college students, rather broad eligibility criteria were chosen that focused more on sensitivity than specificity (Friedrich & Schlarb, 2017).

Generalizability and interpretability of these two reviews are limited due to the small number of included studies, and the inclusion of studies that were not RCTs. The authors could, therefore, not calculate pooled effect sizes using meta-analytic techniques. Because several RCTs on this issue have recently been published, we performed a meta-analysis to evaluate the efficacy of psychological interventions aimed at improving sleep among college students compared with control conditions. Thus, the aim of the current meta-analysis is to deliver quantitative information on the effectiveness of sleep interventions in college students including RCTs only, focusing on psychological treatment with reported outcomes that are related to sleep.

## 2 | MATERIALS AND METHODS

This study was carried out as part of the WHO World Mental Health International College Student initiative (WMH-ICS), which aims to: (a) obtain accurate cross-national information on the prevalence, incidence and correlates of mental, substance and behavioural problems among college students worldwide; (b) describe patterns of service use, barriers to treatment and unmet need for treatment; (c) investigate the associations of these disorders with role function in academic and other life domains; (d) evaluate the effects of a wide range of preventive and clinical interventions on student mental health, functioning and academic performance; and (e) develop precision medicine support tools to help select the right interventions for the right students (Cuijpers et al., 2019). The WHM-ICS's meta-analysis initiative, of which the present meta-analysis is a subproject, has been registered with PROSPERO (CRD42017068758). The procedure and results of this systematic review are outlined in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberatio et al., 2009).

### 2.1 | Study selection

Publications were identified by searching three major electronic databases, starting from database inception on 28 September 2017 up to 20 March 2018: the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and PsycINFO. As part of the WMH-ICS, the search string was compiled on a superordinate level to provide a foundation for investigating a variety of psychological interventions in college students. Thus, the search string did not contain terms that
would restrict the search to the disorders or delivery modes targeted in this analysis; it accepted a high number of references for screening in order to minimize the risk of missing relevant studies. There were no restrictions on publication date or status. The search was based on a string combining terms text words, index, and free terms indicative of RCTs that evaluated psychological interventions in tertiary education settings (see Appendix A). From there, references in identified studies and previous systematic reviews of overlapping topics were checked for earlier publications. The WHO International Clinical Trials Registry Platform was also searched for unpublished trials as a way to identify grey literature. Authors of study protocols without published results were contacted to determine the eligibility of unpublished data for the meta-analysis.

We screened the titles and abstracts of all articles for overall fit for this analysis. Full texts of selected articles were then retrieved and independently assessed for eligibility. Both steps were performed independently by two researchers (KSa, TB). Discussion between researchers was initiated in case of assessor disagreement; two senior researchers (DEE, HB) were consulted when disagreement could not be resolved.

2.2 | Inclusion criteria

We included: (a) RCTs in which (b) individuals enrolled at a tertiary education facility (university, college or comparable postsecondary higher education facility) at the time of randomization, (c) received a sleep-focused psychological intervention, (d) that was compared with a passive control condition, defined as a control condition in which no active manipulation was induced as part of the study (wait-list, treatment as usual). For the purposes of this analysis, "sleep-focused" means: (e) effects on symptoms of sleep disturbances (global measures of sleep disturbances, sleep-onset latency [SOL], fatigue and daytime functionality, pre-sleep behaviour and experiences) were assessed as a (f) target outcome (by declaring a sleep outcome as the primary outcome or by stating the intervention was primarily aimed at this outcome) using (g) standardized symptom measures (objective sleep measures, standardized sleep or fatigue questionnaires, sleep diaries, items recording sleep quantity, quality or hygiene). Only studies (h) published in English or German were considered for inclusion.

Consistent with Cuijpers, van Straten, Warmerdam, and Andersson (2008), psychological interventions were defined as: interventions in which verbal communication between a therapist and a client was the core element; or in which a systematic psychological method was written down in book format or on a website (bibliotherapy), while the client worked through it more or less independently. Psychological interventions consisting of educational elements, cognitive methods, behavioural methods or other CBT-I-related techniques delivered face-to-face (F2F), written, by phone, tape or in computerized form were included. We elected to include all manner of psychological interventions currently practiced in order to provide a representative and overarching depiction of practiced psychological interventions. A subgroup analysis that focused on CBT-I exclusively was also conducted.

Thus, non-psychological interventions like medication, manipulation of light and sound, physical exercise or surgery were excluded. Studies exploring non-specific and hybrid interventions were also excluded, as they did not focus primarily on sleep. Furthermore, studies containing any participants who were not enrolled in higher education institutions were excluded, even if the majority of the sample was. Finally, case studies, cross-sectional studies, non-randomized trials, non-controlled trials, and trials with control groups containing interventions or other active manipulation such as placebo control conditions were also excluded. Our main research question was to analyse non-specific treatment effects of sleep interventions in college students as there was, to date, no evidence whether psychological interventions to treat sleep disturbances were also effective in the vulnerable group of college students (Ohayon & Reynolds, 2009). Thus, active control conditions were excluded as we did not intend to analyse specific effects of treatment in superiority or non-inferiority trials.

2.3 | Data extraction and classification

The following data were extracted for each article, if reported or applicable: (a) bibliographical data (first author, year of publication); (b) sample characteristics (sample size, gender and age distribution, dropout rate, university course[s] of the participants, study subject, compensation for participation, country); (c) participation criteria (inclusion and exclusion criteria); (d) intervention characteristics (mode of delivery, frequency and number of contacts, duration of intervention, therapeutic content, type of control group); (e) outcome measures (time points of assessments, outcomes of interest). In addition to sleep-related outcomes, we were also interested in measures of academic performance due to their reported relationship to sleep disturbances.

If relevant information to examine the eligibility of a study was not available, authors were contacted a maximum of two times to attain or clarify information. If the authors did not respond, and the information given in the publication was insufficient to perform a meta-analysis, the article was excluded. Data extraction was conducted independently by two researchers (KSa, TB).

2.4 | Quality assessment

Study validity was evaluated by two researchers (KSa, TB). The risk of bias assessment was carried out according to the updated method guidelines for systematic reviews with RCTs of the Cochrane Collaboration (Furlan, Pennick, Bombardier, & van Tulder, 2009). We examined bias due to: (a) selection (sequence generation, allocation concealment and the similarity of participants in the different conditions); (b) lack of blinding (blinding of participants, personnel and outcome assessors); (c) incomplete outcome data (amount of dropout and the way dropout data were handled [imputation methods]); and (d) four other threats to validity (e.g. selective outcome reporting, presence of parallel interventions to the study intervention, differences in compliance with the interventions across
conditions, including intensity, duration, number and frequency of sessions and differences in timing of outcome assessments in the conditions). In addition to the recommendations of Furlan et al. (2009), we reported: (e) researcher allegiance, as it has been shown to be associated to outcome (Munder, Brütsch, Leonhart, Gerger, & Barth, 2013); and (f) whether the sample was a convenience sample chosen due to reachability. The risk of researcher allegiance was considered high if at least one of the authors developed the intervention and: (a) delivered it himself; or (b) did not deliver it himself but trained or supervised those that did (Gaffan, Tsaousis, & Kemp-Wheeler, 1995; Leykin & DeRubeis, 2009). The studies’ risk of bias was considered low if: (a) at least six of Furlan et al.’s (2009) 12 suggested risk of bias criteria were rated low; and (b) the risk of researcher allegiance was rated low.

### 2.5 Outcome measures

The included studies reported a number of different outcome measures and used different measurement methods (primarily questionnaires or daily sleep diaries, but also actigraphy and electroencephalography [EEG]). Sleep-related outcomes included all measures assessing: (a) direct sleep disturbances (e.g. global measures of sleep disturbances, SQ, sleep duration, SOL, number of awakenings or sleep patterns); (b) pre-sleep behaviour and experiences (e.g. pre-sleep arousal, technology use before sleep); (c) attitudes towards sleep (e.g. dysfunctional beliefs, attribution patterns concerning sleep); (d) fatigue and daytime functioning attributed to SQ and sleep quantity; and (e) behaviour patterns fostering or decreasing SQ and sleep quantity (e.g. SH). The different sleep-related

---

**TABLE 1** List of outcomes reported in at least one of the included studies

| Outcome | Outcome type | Abbreviation | Frequency | Specific construct |
|---------|--------------|--------------|-----------|--------------------|
| Sleep diary and objective data | Sleep quality | Sleep diary SQ | 2 | GMoSD |
| Sleep-onset latency | Sleep diary SOL | 6 | SOL |
| | Actigraphy SOL | 1 | SOL |
| Difficulties falling asleep | EEG SOL | 1 | SOL |
| Number of awakenings | Sleep diary DFA | 1 | SOL |
| | Sleep diary NWAK | 2 | SOL |
| | Actigraphy | 2 | SOL |
| Sleep efficiency | Sleep diary SE | 2 | SOL |
| Technology use before and at bedtime | Actigraphy | 2 | SOL |
| Terminal wakefulness | Sleep diary Techuse | 1 | SOL |
| | Sleep diary TWAK | 2 | SOL |
| Total sleep time | Actigraphy | 2 | SOL |
| | Sleep diary TST | 1 | SOL |
| Wake after sleep onset | Actigraphy | 2 | SOL |
| | Sleep diary WASO | 3 | SOL |
| | Actigraphy | 1 | SOL |

Questionnaires

- Insomnia Severity Index (Bastien, Vallières & Morin, 2001)
- Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman & Kupfer, 1989)
- Sleep Condition Indicator with 8 items (Espie et al., 2014)
- Epworth Sleepiness Scale (Johns, 1991)
- Fatigue Severity Scale (Krupp, LaRocca, Mizrahi-Eisenberg & Steinberg, 1989)
- Insomnia Impact Scale (Hoelscher, Ware & Bond, 1993)
- Multidimensional Fatigue Inventory (Smets, Garssen, Bonke & De Haes, 1995)
- Work and Social Adjustment Scale (Mundt, Marks, Shear & Greist, )
- Dysfunctional Beliefs and Attitudes about Sleep (Short Form) (Morin, Stone, Trinkle, Mercer, & Remsberg, 1993; Short-form: Morin, Vallières & Ivers, 2007)
- Sleep Hygiene Index (Mastin, Bryson, & Corwyn, 2006)
- Disturbing Dreams and Nightmare Severity Index (Krakow et al., 2002)

Abbreviations: DFF, daytime-functionality and fatigue; Frequency, number of comparisons reporting the outcome; GMoSD, global measure of sleep disturbances; PSBE, pre-sleep behaviour and experiences; SOL, sleep-onset latency; Specific constructs, constructs, the outcomes were allocated to, for calculations of effect sizes of specific constructs.
outcomes reported in the studies and their frequency can be seen in Table 1.

The majority of studies reported more than one outcome assessing sleep, and a primary outcome was not defined. Because of this, we used combined effect sizes for each study. Thus, effect sizes contain effects of the psychological intervention on all obtained sleep-related facets. Effect sizes were calculated by standardizing all sleep-related outcomes within a study and combining them to one effect size. When calculating a combined effect size within a single study, the intercorrelations between included outcomes need to be considered. We assumed a conservative intercorrelation of $r = 1$. To account for the potential distortion of effect size created by this assumption, we added a sensitivity analysis with an intercorrelation estimation of $r = 0$ (Becker, 2000).

### 2.6 Meta-analyses

#### 2.6.1 Power calculation

An a priori power analysis (Borenstein, Hedges, Higgins, & Rothstein, 2010) indicated that at least 15 studies with at least 20 participants per condition, respectively, six studies with at least 50 participants per condition, would be needed to detect an effect of $g = 0.3$ (with moderate heterogeneity) with a statistical power of 0.8 and a significance level of $\alpha = .05$.

#### 2.6.2 Effect size calculations

An effect size for each sleep-related outcome in each study was calculated. These effect sizes compared post-treatment values of intervention groups and control groups. As many included studies had small sample sizes, bias was corrected for by calculating Hedges' $g$ as an effect size metric (Hedges & Olkin, 2014). The effect sizes were calculated using reported means and standard deviations or t-values (Lipsey & Wilson, 2010). Effect sizes below $g = 0.32$ were rated as small, between $g = 0.33$ and 0.55 as moderate, and above 0.55 as large (Lipsey & Wilson, 1993). To facilitate clinical interpretation of standardized mean difference (Hedges' $g$), we calculated numbers-needed-to-treat (NNT) to generate one additional clinically significant change using the formulae of Kraemer and Kupfer (2006). We applied the random effects model for pooled effect sizes (Borenstein et al., 2010). Effect sizes, measures of heterogeneity (visual forest-plot inspection, Q and $I^2$) and publication bias (visual funnel-plot analysis, Egger's test, Duval and Tweedie's Trim and Fill method) were calculated with the programme CMA (version 3.3.070; Borenstein, Hedges, Higgins, & Rothstein, 2014). For the calculation of effect sizes, we only conducted meta-analyses when there were five or more studies available for pooling. The confidence interval of $I^2$ was calculated according to Borenstein et al. (2010). $I^2$ heterogeneity of 25% can be regarded as low, 50% as moderate, and 75% as substantial heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003).

#### 2.6.3 Subgroup analyses

Subgroup analyses were based on the mixed effects model; a fixed effects model was used to test differences across subgroups, while a random effects model was used within subgroups (Borenstein et al., 2010). Due to the small number of studies included, only three covariates were tested in subgroup analyses: (a) type of psychological intervention; (b) mode of delivery; and (c) risk of bias. Meta-regressions were not conducted for this same reason. (a) Type of psychological intervention could be categorized in behavioural, cognitive, educational, and cognitive behavioural treatment. (b) Mode of delivery was divided into individual F2F therapy, group therapy, and self-help treatment based on the included study; and (c) risk of bias was categorized in high and low risk of bias for the subgroup analysis.

#### 2.6.4 Specific sleep constructs

In addition to the calculation of the overall effect size, we analysed the effect of the interventions on specific sleep constructs. The construct "global measures of sleep disturbances" includes questionnaires assessing the conditions of sleep (Pittsburgh Sleep Quality Index [PSQI]), Insomnia Severity Index [ISI], Sleep Condition Indicator with eight items [SCI-8], and the sleep diary item "sleep quality"). To reduce the risk of over-representing single aspects of sleep disturbances, sleep diary items focusing on a single aspect (e.g. "number of awakenings") were excluded. The construct "fatigue and daytime functionality" includes all questionnaires assessing fatigue and daytime functionality in connection to SQ, sleep continuity or sleep duration. The construct "pre-sleep behaviour and experiences" includes all questionnaires assessing SH or attitudes towards sleep. The construct "SOL" includes all measures assessing the duration of time spent trying to fall asleep to actually falling asleep; this includes measures derived from the sleep diary, actigraphy and EEG (Table 1).

#### 2.6.5 Follow-up analysis

The effect sizes for available follow-up measurements were calculated by comparing the treatment conditions with the control conditions. Follow-up measurements include all measurements declared by the authors of the studies as such and timed after the end of intervention and post-test measurement.

### 3 RESULTS

#### 3.1 Selection of studies

We screened a total of 12,206 studies for titles and abstracts (search string n = 11,936; secondary search strategies n = 270). After this screening, 23 studies of the main search string and 12 studies of the additional secondary search strategies remained to be assessed...
on the full text level. The study selection process is described in a PRISMA flowchart (Figure 1). Finally, 10 studies were included in the meta-analysis. The inter-rater reliability concerning study selection was very good ($\kappa = 0.83$).

### 3.2 | Missing data

With regard to missing data, 14 authors were asked via email to provide the full text. Five of the missing studies were conference abstracts and thus not able to be included. Four studies could be identified as unpublished master theses with no access given. Three authors did not provide full text articles needed, and thus respective studies had to be excluded. Two of the studies received were excluded because they: (a) examined a treatment focusing on headache with sleep disturbances as a co-morbidity; and (b) investigated the effectiveness of booster sessions after the initial treatment only.

### 3.3 | Study characteristics

Detailed study characteristics can be found in Table 2.

The 10 studies included 2,408 college students in total (treatment conditions: $n = 1,002$; control conditions $n = 1,406$). In the seven studies reporting gender, the majority of participants were female (70.1%). Five studies included only psychology students, while the other five included students from a variety of disciplines. In five of the studies, participants did not receive any incentive for study participation. Conversely, of the remaining five studies, two offered course credit, two offered monetary compensation, and the final study offered both.

![PRISMA flowchart of included studies](image)
Seven studies were conducted in the USA, two in the UK and one in China. Two studies did not have any inclusion criteria.

Four studies compared CBT-I with control groups, and one of these studies added music therapy to CBT-I. Freeman et al. (2018) used the unguided web-based CBT-I intervention Sleepio, which is based upon a validated manual (Espie et al., 2001, 2007, 2008) presented by an animated therapist. An initial assessment drove the algorithms used to personalize the programme. Participants completed interactive therapy sessions once a week for 6 weeks; each session lasted 20 min on average. Additionally, participants completed daily sleep diaries. Throughout the intervention, participants had access to a moderated online community and an online library for information about sleep. The psychological intervention included sleep restriction, stimulus control (getting out of bed after 15–20 min not being able to fall asleep) and relaxation as behavioural techniques. The cognitive techniques included paradoxical intention, belief restructuring, mindfulness, imagery, and putting the day to rest. The educational component covered information about SH and the processes of sleep.

Gao et al. (2014) delivered CBT-I components on three consecutive days in 45-min group sessions, including: (a) sleep education; (b) stimulus control; and (c) progressive muscle relaxation (PMR). The first session included sleep education on: (a) proper sleep habits, sleeping process, and sleeping structure; (b) techniques to improve sleep; and (c) instructions to listen to calming music (provided in a music library by the researchers) for 30 min within 1 hr before night-time sleep. The stimulus control portion of this intervention included six steps participants were required to complete daily (e.g. getting out of bed after 20 min not being able to fall asleep). Finally, for PMR, participants were trained to divide their muscles into 16 parts and muscles into 16 groups that were contracted and relaxed successively. Participants were instructed to do PMR daily for 30 min, and were encouraged to combine music therapy with PMR. The process was supervised by researchers visiting students twice a week for 1 month.

Morris et al. (2016) used an unguided web-based CBT-I programme that contained seven modules and monitoring tools that depicted graphical feedback on sleep. Participants received periodic reminders to complete a given module; these reminders included motivational input. In order, the modules contained: (a) psychoeducation; (b) guided imagery; (c) relaxation techniques; (d) SH, stimulus control therapy and sleep restriction; (e) relaxation techniques; (f) PMR; and (g) information and tips.

The CBT-I programme delivered by Taylor et al. (2014) was based on two treatment manuals (Morin & Espie, 2004; Perlis, Benson-Jungquist, Posner, & Smith, 2005). The programme included: (a) stimulus control therapy (e.g. getting out of bed after 20 min not being able to fall or return to sleep); (b) sleep restriction; (c) SH information addressing behaviours that may influence SQ and quantity; (d) a relaxation exercise combining deep breathing, PMR and autogenic training to be practiced twice a day; and (e) cognitive restructuring (e.g. eliciting the participants’ dysfunctional beliefs and attitudes about sleep, and developing more adaptive ones).

Five studies used standalone behavioural interventions: relaxation (k = 3), stimulus control (k = 1) and exposition (k = 1). The three relaxation interventions contained either group sessions (Borkovec & Weerts, 1976; Steinmark & Borkovec, 1974) or individual sessions (Menas, Lichstein, Epperson, & Johnson, 2000) that trained participants in PMR and instructed them to complete PMR twice daily, including once prior to sleep. The stimulus control intervention (Zwart & Lisman, 1979) required participants to follow stimulus control rules, such as getting up after 10 min of not being able to fall or return to sleep. In the exposition intervention (Carrera & Elenewski, 1980), participants were asked to visualize lying in bed and having difficulties falling asleep. Then they were encouraged to attend to bodily sensations and interpret them as a strange disease. Finally, they were instructed to imagine their own death.

One study used a sleep education programme (Barber & Cucalon, 2017) in which participants were informed about general SH information and technology boundary management via PowerPoint presentation.

Five of the psychological interventions were self-help programmes, of which four were delivered electronically and two were preceded by F2F training sessions. The other psychological interventions were delivered F2F as group therapy (k = 3) and individual therapy (k = 2). The group size ranged from two to seven. Considering all 10 studies, the number of contacts ranged from one to 11, and averaged 4.7 contacts (including e-mail contacts). The post-test measurements were conducted on average 4.5 weeks after the baseline measurements (range 1-10 weeks). This duration also reflects the intervention duration.

To measure sleep outcomes, sleep diaries (k = 6), objective measures such as actigraphy and overnight EEG (k = 3), and questionnaires (k = 6) were used. No study evaluated effects on academic functioning. Inter-rater agreement concerning study characteristics was very good (κ = 0.87).

### 3.4 Quality assessments

The complete assessment for each study is presented in Table 3. Overall risk of bias was considerable. Of the 10 studies, five showed a high risk of bias. On average 7.5 of 14 assessments per study were rated unclear or high. Six studies did not report information on sequence generation and allocation concealment. Two studies did not report on demographic and baseline differences between conditions. One study had a high risk of bias in allocation concealment. All studies revealed an unclear or high risk of bias in at least two of the three blinding situations. Five of the studies revealed low risk of bias in the domain of incomplete outcome data. Concerning other threats to validity, most criteria showed on average low risk of bias. Only selective reporting showed on average a higher risk of bias, with nine studies being rated unclear and only one study being rated low. Three studies had an unclear or high risk of compliance differences between the treatment conditions, and three showed a high risk of researcher allegiance. Inter-rater reliability regarding risk of bias assessments was good (κ = 0.76).
### TABLE 2  Selected characteristics of included RCTs (n = 10)

| Author (Year) | Participants | Criteria | Outcome measures |
|---------------|--------------|----------|-----------------|
|               | N | Age M | Subject | Inclusion symptoms | Exclusion | MoD | Frequency | Treatment condition (n) | Control condition (n) | Baseline | Follow-up | Questionnaires | Objective measures |
| Barber and Cucalon (2017) | 78 | 20.00 | all | none | none | SH-E | once | Sleep education (43) | TAU (35) | 0 | SHI | ICS | Actg(TST, NWAK, SE) |
| Borkovec and Weerts (1976) | 22 | / | psy | SOL > 30 min | CUD | GT (2) | weekly | Relaxation (11) | TAU (11) | 0 | SOL | diary (SOL) |
| Carrera and Elenewski (1980) | 39 | / | psy | SOL > 45 min | CWOPS | SH-E | once | Death-exposure (16) | WLC (23) | 4 weeks | diary (SOL) |
| Freeman et al. (2017) | 1,875 | 24.70 | All | SCI < 16 | none | SH-E | weekly | CBT-I (733) with sleep education, sleep restriction, stimulus control, relaxation, cognitive restructuring, imagery, mindfulness, paradox, thought stopping, putting day to rest | TAU (1,142) | 0 | SCI-8 | ISI |
| Gao et al. (2014) | 84 | 20.49 | all | PSQI > 7 | PMPEI | SH-T | daily | CBT-I + M (42) with sleep education, stimulus control, relaxation, music therapy | TAU (42) | 0 | PSQI | |

(Continues)
| Author (Year)                  | Participants                                                                 | Criteria                  | Outcome measures                                                                 |
|-------------------------------|-----------------------------------------------------------------------------|---------------------------|----------------------------------------------------------------------------------|
| Means et al. (2000)           | 57 | 20.59 | psy  | SOL > 30 min | CUD-S | IT | Relaxation (28) | WLC (29) | 0 & 2 weeks | IIS | Sleep diary (SOL, SQ WASO, SE) |
|                               | 68.5% | 4.84 | cc   | AASO > 30 min | PMPEI | 3 – 7 days | 3 x 15 min | 8 weeks | / | DBAS |
|                               | / | all | US | IDF > 2 months | ASS | / | / | / | / | ESS |
| Morris et al. (2016)          | 95 | 20.48 | All | None | none | SH-E | weekly | 7 x 20 min | 6 weeks | PSQI |
|                               | 67.4% | 2.09 | mon | none | none | weekly | 4 weeks | / | / |
|                               | 15 | all | GB | None | none | weekly | 4 weeks | / | / |
| Steinmark et al. (1974)       | 24 | / | psy | SOL > 30 min | CUD | GT (5 – 7) | Relaxation (12) | WLC (12) | 0 | Sleep diary (SOL, WASO, TWAK, NWAK, DFS) |
|                               | / | / | cc | > 6 months | CWOPS-S | weekly | 4 weeks | / | / |
|                               | 4 | 1 | US | / | / | / | 5 months | / | / |
| Taylor et al. (2014)          | 29 | 19.71 | all | SOL > 30 min | CUD | IT | CBT-I (16) with sleep education, sleep restriction, stimulus control, relaxation, guided imagery | WLC (13) | 0 – 1 week | PSQI |
|                               | 58.8% | 2.10 | mon | AASO > 30 min | PMPEI | weekly | 3 months | 6 weeks | ISI |
|                               | 5 | all | US | IDF > 3 months | 6 | / | / | / | ESS |
| (Continues)                   |                                             |                           |                                                 |                                                 |                                                 |                                                 |                                                 |                                                 |                                                 |
3.5 | Overall effect of psychological intervention

The overall effect size of combined outcomes of the 10 studies was moderate to large ($g = 0.61$; 95% confidence interval [CI]: 0.41–0.81) with moderate heterogeneity ($I^2 = 41; 95\% \text{ CI: } 0–72$). Negative effect sizes favour intervention success in the intervention group compared with the control group. The effect size for the combined outcomes equals a NNT of 3.

Visual analysis of the corresponding forest-plot (Figure 2) revealed one potential outlier, not overlapping with the pooled confidence interval. Excluding this study generated a moderate effect size of $g = 0.54$ (95% CI: 0.46–0.63) corresponding to a NNT of 3.36, with no heterogeneity ($I^2 = 0; 95\% \text{ CI: } 0–52$). The inter-rater reliability regarding the parameters to calculate effect sizes was excellent ($\kappa = 0.98$).

3.6 | Follow-up

Only three studies reported follow-up results. The studies gathered follow-up measures 1 month, 3 months or 12 months, respectively, after the end of psychological intervention. The combined follow-up effect size was $g = 0.56$ (95% CI: 0.45–0.66) with low heterogeneity of $I^2 = 0$ (95% CI: 0–95).

3.7 | Subgroup analyses

Subgroup analyses yielded no significant differences between subgroups, indicating that mode of delivery, risk of bias and type of treatment did not affect the effect size (see Appendix B).

3.8 | Sensitivity analyses

Calculating the overall effect size assuming independence of the outcomes (an intercorrelation of $r = 0$) yielded a slightly smaller effect size ($g = 0.57$; 95% CI: 0.38–0.75) and a higher heterogeneity ($I^2 = 64; 95\% \text{ CI: } 30–82$).

3.9 | Specific sleep constructs

For a summary of the effect sizes of specific sleep constructs, see Table 4.

3.9.1 | Global measures of sleep disturbances

Six studies reported at least one global outcome measure of sleep disturbances. Psychological interventions were able to improve the global measure of sleep disturbances compared with passive control conditions ($g = 0.79; 95\% \text{ CI: } 0.52–1.06$), but also showed moderate
| Study                        | Selection bias | Blinding | Incomplete outcome data | Other threats to validity |
|------------------------------|----------------|----------|-------------------------|---------------------------|
|                             | Generation a   | Concealment b | Similar groups c         | Participants d | Personnel e | Outcome assessors f | Dropout g | Intention to treat analysis h | Selective Reporting i | Compliance j | Timing k | Cointerventions l | Researcher allegiance m | Convenience sample n | Risk of Bias o |
| Barber and Cucalon (2017)    | Low a          | Low      | High                    | High                     | Low          | Low       | Low          | Low       | Unclear                  | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Borkovec and Weerts (1976)   | Unclear        | Unclear  | Unclear                | High                     | High         | Low       | Unclear     | Low       | Unclear                  | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Carrera and Elenewski (1980) | Unclear        | Unclear  | Low                    | Unclear                   | Low          | Low       | Unclear     | Low       | Unclear                  | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Freeman et al. (2017)        | Low a          | Low      | Low                    | High                     | High         | Low       | Low          | High      | Low                      | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Gao et al. (2014)            | Unclear        | Unclear  | Unclear                | Low                      | High         | High      | Low          | Unclear   | Unclear                  | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Means et al. (2000)          | Unclear        | Unclear  | Low                    | Unclear                   | Unclear      | Unclear   | Unklear     | Unclear   | Unclear                  | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Morris et al. (2016)         | Low a          | Low      | Low                    | High                     | High         | Low       | Low          | High      | Low                      | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Steinmark and Borkovec (1974)| Unclear        | Unclear  | Low                    | High                     | High         | Low       | Low          | High      | Low                      | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Taylor et al. (2014)         | Low a          | Low      | Low                    | High                     | High         | Unclear   | Low          | High      | Low                      | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |
| Zwart and Lisman (1979)      | Unclear        | Unclear  | Unclear                | Unclear                  | Unclear      | Unclear   | High         | High      | High                     | Low          | Low       | Low       | Low          | High          | Low       | Low       | Low       |

*aRisk of selection bias due to inadequate generation of a randomized sequence.

bRisk of selection bias due to inadequate concealment of allocations prior to assignment.

cRisk of selection bias due to baseline differences between the treatment conditions in outcome measures and demographic measures.

dRisk of performance bias due to knowledge of the allocated interventions by participants during the study.

eRisk of performance bias due to knowledge of the allocated interventions by personnel during the study.

fRisk of detection bias due to knowledge of the allocated interventions by outcome assessors (objective outcomes/subjective outcomes).

gRisk of attrition bias due to incomplete outcome data.

hRisk of attrition bias due to handling of incomplete outcome data.

iRisk of reporting bias due to selective outcome reporting.

jRisk of performance bias due to differing compliance between the conditions.

kRisk of detection bias due to differences in time of outcome measurement across groups.

lRisk of performance bias due to differences in cointerventions across groups.

mRisk of detection bias due to researchers’ attitudes to the treatments.

nSample chosen for reasons of reachability.

oOverall estimated risk of bias.

pLow, low risk of bias; high, high risk of bias; unclear, not enough information given to estimate the risk of bias.

qSerious flaw: less than six low risk of bias assessments (researcher’s allegiance and convenience sample excluded).

rSerious flaw: less than six low risk assessments and high risk of researcher’s allegiance.

sSerious flaw: high researcher allegiance.
to high heterogeneity ($I^2 = 65$; 95% CI: 16–85; see Appendix C). We carried out the same subgroup analyses as in the sensitivity analyses for the overall effect size.

Including only studies using the PSQI (the most frequently reported questionnaire) for a global measurement of sleep disturbances led to even stronger effects ($g = 1.09$; 95% CI: 0.42–1.77) with high heterogeneity ($I^2 = 79$; 95% CI: 32–93).

### 3.9.2 | Sleep-onset latency

Sleep-onset latency was assessed by six studies through sleep diaries, actigraphy and EEG. The effect size of SOL was large ($g = 0.65$; 95% CI: 0.36–0.94) with very little heterogeneity ($I^2 = 0$; 95% CI: 0–63; see Appendix D).

### 3.10 | Publication bias

Both Egger’s tests and visual inspection of funnel plots did not show any significant publication bias for all but one analysis. Only the visual inspection of the funnel plot for the effect size distribution of SOL suggested some publication bias. According to Duval and Tweedie’s trim and fill procedure, three studies were imputed, resulting in a smaller but still significant effect size of $g = 0.46$.

### 4 | DISCUSSION

This meta-analysis suggests that psychological intervention may reduce sleep disturbances in college students, as determined by global measures of sleep disturbances and SOL. Overall risk of bias was substantial, but no indication of larger effects through distortion could be found relating to the risk of bias of a study. Indication for publication bias was only found for the construct SOL, in which adjusting for potential unpublished studies resulted in change. Thus, the present analysis implies psychological intervention to be an appropriate way to help the large number of college students with sleep disturbances.

Our results are in line with previous research among college students. In their review, Friedrich and Schlarb (2017) found comparable medium to high effect sizes regarding SQ ($d = 0.61–1.56$), a high range of effects regarding SOL ($d = 0.38–1.16$), and higher effects regarding daytime functioning for CBT-I ($d = 1.10$). In a meta-analysis, including general and clinical populations, van Straten et al. (2018) found smaller effects for SOL ($g = 0.4$) and PSQI ($g = 0.65$), and slightly smaller effects for SOL ($g = 0.57$), although CIs did overlap. Due to different operationalization of SQ (in our analysis called global measures of sleep disturbances to emphasize the different operationalization) and broad CIs in the present meta-analysis, those differences in effect size need to be interpreted cautiously. Still, they are in line with results found by van Straten et al. (2018) when examining age as a potential moderator. These authors found significant age differences in two (SOL and sleep efficiency [SE]) of the three main outcome variables (young adults: $g = 0.7$ for SOL, $g = 0.99$ for SE, $g = 1.02$ for ISI; all adults: $g = 0.42$ for SOL, $g = 0.55$ for SE, $g = 0.73$ for ISI). Assuming this difference in effect sizes is due to an actual increased benefit for college students compared with older adults (rather than representing methodological or other biases), it might be explained by two factors. First, young adults experience especially high strain from sleep disturbances, and show the worst SH compared with older adults (Rosenberg et al., 2011). Thus, college students (who are, in general, young adults) may have on average more room for improvement regarding their sleep. This assumption is similar to results found in research on depression, showing that higher pre-intervention symptom severity is associated with higher efficacy of psychological intervention (Driessen, Cuijpers, Hollon, & Dekker, 2010). In accordance with these findings, the study by van Straten et al. (2018) found higher effects of psychological treatment on sleep-related outcomes in primary care or specific care settings, in which participants are presumed to have stronger symptoms and psychological distress compared with the general population (care settings: $g = –0.64$ to –0.93; general population: $g = –0.47$ to –0.82; van Straten et al., 2018).

Second, as outlined above, college students are likely to show above average cognitive flexibility (Colzato et al., 2006), resulting in a higher capacity to change their sleep behaviour and thus a potentially heightened benefit of psychological intervention. Future research is needed to confirm such assumptions.

The present meta-analysis did not find efficacy differences between different types of psychological interventions. This is inconsistent with the suggestion of Friedrich and Schlarb (2017) that CBT-I is the most effective psychological intervention, and SH interventions the least effective, a specification supported by research in the general population (Morin, Culbert, & Schwartz, 1994; Morin et al., 1999; Morin, LeBlanc, et al., 2006; Qaseem et al., 2016). It is noteworthy, though, that the latest and most comprehensive meta-analysis on psychological interventions for sleep disturbances and insomnia in the general population also failed to find this specification. The authors suggested that this may be due to variability within each type of psychological intervention, or a potential correlation of psychological intervention type, study quality and variation in effect size with time of publication; this would result in smaller effects for the types of psychological interventions investigated more recently with higher study quality (van Straten et al., 2018). This explanation might also apply to the results of the present meta-analysis. It is also noteworthy that the small number of trials in our meta-analysis might have reduced power to detect true differences in efficacy between different types of psychological intervention. However, effects of specific psychological intervention techniques could also be overestimated, and the reduction of sleep disturbances might be evoked by generic mechanisms of change common to all types of psychological intervention like therapeutic alliance, empathy, goal consensus and collaboration.

Although a previous study found F2F psychological intervention to be more effective in improving sleep than self-help psychological intervention in the general population (Lancee, van Straten, Morina, Kaldo, & Kamphuis, 2016), the present analysis does not support this
finding. This lack of significant difference in effect size is supported by Blom et al. (2015). Van Straten et al. (2018) found mixed results with F2F psychological intervention being more effective in improving SOL and SE, but not concerning the ISI. If the efficacy of self-help psychological intervention is supported by future studies, it may become the psychological intervention of choice for college students wanting to improve their sleep. Assuming equal efficacy between different modes of delivery, self-help interventions would combine efficacy with superior cost-effectiveness and accessibility compared with F2F psychological intervention.

The current meta-analysis has some limitations. First, the number of included studies was small. However, it is not only the number of studies that makes a meta-analysis reliable, but also the number of participants per study (Borenstein et al., 2010). The 10 studies that were included in this meta-analysis had an average sample size of more than 100 participants per condition, which is sufficient to reliably detect effect sizes of 0.3 and above, even given a small number of trials with high heterogeneity (Borenstein et al., 2010). Yet the small number of studies does enhance the risk of missing true effects due to insufficient power, especially in subgroup analyses. Second, half of the studies showed a high risk of bias, which can increase the chances of inflated effect sizes in comparison to the true effects in the population. However, subgroup analyses (grouped by risk of bias assessment) did not show significant differences in effect size. Finally, this meta-analysis had limited generalizability for the analysis of long-term effects, as only three studies reported usable follow-up data at different time periods.

Despite these limitations, we conclude that psychological interventions aimed at improving sleep among college students are effective, perhaps even more so than in other populations. Thus, extending additional validated psychological interventions to college students suffering from sleep disturbances is likely to be a worthwhile investment for both the individual and society. By improving sleep, psychological interventions reduce the risk of clinical sleep disorders, mental disorders and medical conditions that are triggered or amplified by sleep disturbances and insomnia (along with reducing the associated economic costs). Importantly, these reductions in risk would come during a critical period for the development of mental disorders (Breslau, Roth, Rosenthal, & Andreski, 1996; Daley et al., 2009; Kessler, Berglund, et al., 2005; Kessler, Chiu, et al., 2005; Sivertsen et al., 2014). Above that, such interventions may improve the cognitive performance of college students, who represent a large portion of a society’s human capital (Gomes et al., 2011). More RCTs are needed to evaluate the effects and moderators of psychological interventions to improve sleep. It is particularly important to investigate the effects of psychological intervention design (e.g., type of psychological intervention, mode of delivery and number of contacts).

Future research should focus on creating more precise and consistent definitions of constructs (e.g., SQ) in order to differentiate constructs and their operationalization. Doing so may reduce the high number of sleep-related outcomes used in different sleep studies, and improve the comparability and quality of outcomes. It would also contribute to the quality of meta-analyses evaluating psychological sleep interventions. Furthermore, academic performance should be assessed and reported as an outcome in studies examining sleep disturbances in college students. This would help to investigate whether sleep interventions can decrease cognitive impairment

| Study name          | Outcome            | Statistics for each study | Hedges’s g and 95% CI |
|---------------------|--------------------|----------------------------|-----------------------|
| Barber et al. (2017)| Combined           | -0.17                      | [0.28, 0.61]          |
| Borkovec et al. (1976)| Combined         | -0.59                      | [0.24, 1.43]          |
| Carrera et al. (1980)| SOL (diary)       | -0.96                      | [0.30, 1.62]          |
| Freeman et al. (2017)| Combined         | -0.55                      | [0.46, 0.65]          |
| Gao et al. (2014)   | PSQI               | -1.29                      | [0.82, 1.75]          |
| Means et al. (2000) | Combined           | -0.47                      | [0.05, 0.99]          |
| Morris et al. (2016)| PSQI              | -0.51                      | [0.10, 0.91]          |
| Steinmark et al. (1974)| Combined       | -0.20                      | [0.58, 0.99]          |
| Taylor et al. (2014)| Combined           | -0.72                      | [0.03, 1.48]          |
| Zwart et al. (1979) | SOL (diary)        | -0.91                      | [0.10, 1.91]          |
| **Combined sleep related outcomes overall** |                     | **-0.61**                  | **[0.41, -0.80]**     |

**Note.** Combined: Mean of all reported sleep related outcomes within a study; PSQI: Pittsburgh Sleep Quality Index; SOL (diary): Sleep onset latency measured via sleep diary

**FIGURE 2** Overall effects of psychological treatment on combined sleep-related outcomes compared with passive control groups
induced by sleep disturbances and insomnia, or maybe even improving academic performance independent of former losses.

5 | CONCLUSION

Psychological interventions for improving sleep in college students may improve global measures of sleep disturbances and daytime functionality, and reduce SOL. Providing validated psychological interventions to college students who suffer from sleep disturbances may therefore be a worthwhile investment for both society and the individual. By improving sleep, psychological interventions reduce the risk of clinical sleep disorders, mental disorders and medical conditions triggered or amplified by sleep disturbances and insomnia. This in turn reduces the economic cost associated with these disorders and conditions. Above that, such interventions may improve the cognitive performance of college students, who represent a large portion of a given society’s human capital. The task of providing psychological treatment to college students includes not only establishing sufficient provision of psychological interventions, it also includes raising awareness of intervention opportunities and providing psychological interventions that are suited to the needs of college students in particular.

ACKNOWLEDGEMENTS

DDE reports to have received consultancy fees or served in the scientific advisory board of Minddistrict, Sanofi, Lantern, Schön Kliniken, and German health insurance companies (BARMER, Techniker Krankenkasse). DDE and MB are stakeholders of the Institute for Health Trainings online (GET.ON), which aims to implement scientific findings related to digital health interventions into routine care. HB reports to have received consultancy fees and fees for lectures or workshops from chambers of psychotherapists and training institutes for psychotherapists. In the past 3 years, RCK received support for his epidemiological studies from Sanofi Aventis, was a consultant for Johnson & Johnson Wellness and Prevention, Sage Pharmaceuticals, Shire, Takeda, and served on an advisory board for the Johnson & Johnson Services Inc and Lake Nona Life Project. RCK is a co-owner of DataStat, Inc, a market research firm that carries out health care research. KSp has received fees for workshops from training institutes for psychotherapists.

This study was carried out as part of the WHO World Mental Health International College Student project (WMH-ICS): https://www.hcp.med.harvard.edu/wmh/college_student_survey.php

CONFLICT OF INTEREST

No conflicts of interest declared.

AUTHOR CONTRIBUTIONS

DDE, HB and KSa conceived and designed the meta-analysis. KSa and TB analysed the data. KSa wrote the paper with important contributions from A-CZ and DDE. All authors participated in the review and revision of the manuscript, and have approved the final manuscript to be published.

ORCID

Karina Saruhanjan https://orcid.org/0000-0001-7960-084X

REFERENCES

Abel, J. R., & Deitz, R. (2012). Do colleges and universities increase their region’s human capital? *Journal of Economic Geography*, 12, 667–691. https://doi.org/10.1093/jeg/lbr020

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55, 469–480. https://doi.org/10.1037//0003-066X.55.5.469

Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., & Tahan, K. (2011). *The condition of education 2011 (NCES 2011–033)*. Washington, DC: U.S. Government Printing Office.

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., ...Kessler, R. C. (2018). *WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders.* *Journal of Abnormal Psychology*, 127, 623–638. https://doi.org/10.1037/abn0000362

Bagioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., ...Riemann, D. (2011). Insomnia as a predictor of depression: A meta-analytic evaluation of longitudinal epidemiological studies. *Journal of Affective Disorders*, 135, 10–19. https://doi.org/10.1016/j.jad.2011.01.011

Barber, L. K., & Cucalon, M. S. (2017). Modifying the Sleep Treatment Education Program for Students to include technology use (STEPS-TECH): Intervention effects on objective and subjective sleep outcomes. *Stress and Health: Journal of the International Society for the Investigation of Stress*, 33, 684–690. https://doi.org/10.1002/smi.2746

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Medicine*, 2(4), 297–307.

---

**TABLE 4** Effects of psychological treatment on specific sleep constructs

| Construct specification | Outcomes | Studies (n) | Effect size | Heterogeneity |
|-------------------------|----------|------------|-------------|---------------|
| 1. Global measures of sleep disturbances | ISI, PSQI, SCI-8, SQ(diary) | 6 | 0.76 | 95% CI: 0.52 to 1.06 | p: 0.000 | NNT: 2.36 | I²: 65 | 95% CI: 16–85 |
| 2. Sleep-onset latency | All SOL measures | 6 | 0.65 | 95% CI: 0.36 to 1.03 | p: 0.000 | NNT: 2.82 | I²: 0 | 95% CI: 0–63 |

Note: All SOL measures: sleep-onset latency measured via sleep diary, actigraphy and EEG; CI: confidence interval; ISI: Insomnia Severity Index; NNT: numbers needed to treat; PSQI: Pittsburgh Sleep Quality Index; SCI-8: Sleep Condition Indicator with 8 items; SQ (diary): sleep quality assessed via sleep diary.
Ohayon, M. M. (2002). Epidemiology of insomnia: What we know and what we still need to learn. Sleep Medicine Reviews, 6, 97–111. https://doi.org/10.1053/smrv.2002.0186

Ohayon, M. M., & Reynolds, C. F. III (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.

Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College students: Mental health problems and treatment considerations. Academic Psychiatry, 39(5), 503–511.

Peltz, C. B., Gratton, G., & Fabiani, M. (2011). Age-related changes in electrophysiological and neuropsychological indices of working memory, attention control, and cognitive flexibility. Frontiers in Psychology, 2, 190. https://doi.org/10.3389/fpsyg.2011.00190

Perlis, M. L., Benson-Jungquist, C., Posner, D. A., & Smith, M. T. (2005). Cognitive behavioral treatment of insomnia: A session-by-session guide. New York, NY: Springer Science+Business Media, Inc.

Pierceall, E. A., & Keim, M. C. (2007). Stress and coping strategies among 18 of 22 University of Virginia inpatients. American Journal of Health Behavior, 31, 703–712. https://doi.org/10.1080/1068926060866579

Pierceall, E. A., & Keim, M. C. (2007). Stress and coping strategies among community college students. Community College Journal of Research and Practice, 31, 703–712. https://doi.org/10.1080/1068926060866579

Pigeon, W. R., Bishop, T. M., & Krueger, K. M. (2017). Insomnia as a precipitating factor in new onset mental illness: A systematic review of recent findings. Current Psychiatry Reports, 19, 44. https://doi.org/10.1007/s11920-017-0802-x

Qaseem, A., Kansagara, D., Forciea, M. A., Cooke, M., & Denberg, T. D. (2016). Management of chronic insomnia disorder in adults: A clinical practice guideline from the American College of Physicians. Annals of Internal Medicine, 165, 125–133. https://doi.org/10.7326/M15-2175

Reavley, N., & Jorm, A. F. (2010). Prevention and early intervention to improve mental health in higher education students: A review. Early Intervention in Psychiatry, 4, 132–142. https://doi.org/10.1111/j.1751-7893.2010.00167.x

Riemann, D., Baglioni, C., Bassetti, C., Borovcnik, B., Dolenc Groselj, L., Ellis, J. G., ... Spiegelhalder, K. (2017). European guideline for the diagnosis and treatment of insomnia. Journal of Sleep Research, 26, 675–700. https://doi.org/10.1111/jsr.12594

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A Meta-Analysis. Psychological Science, 29, 1358–1369. https://doi.org/10.1177/0956797618742453

Rosenberg, R., Gradišar, M., Hale, L., Harvey, A. G., Wolfson, A. R., & Czeisler, C. A. (2011). Sleep in America Poll 2011. Retrieved from http://sleepfoundation.org/sites/default/files/sleepinameripoll_SIAP_2011_Summary_of_Findings.pdf

Scharlb, A. A., Kulessa, D., & Gulewitsch, M. D. (2012). Sleep characteristics, sleep problems, and associations of self-efficacy among German university students. Nature and Science of Sleep, 4, 1–7. https://doi.org/10.2147/NSS.S27971

Siverts, B., Lallukka, T., Salo, P., Pallesen, S., Hysing, M., Krostad, S., & Simon, Ø. (2014). Insomnia as a risk factor for ill health: Results from the large population-based prospective HUNT Study in Norway. Journal of Sleep Research, 23, 124–132. https://doi.org/10.1111/jsr.12102

Smets, E.M.A., Garssen, B., Bonke, B., De Haes, J.C.J.M. (1995). The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. Journal of Psychosomatic Research, 39(3), 315–325.

Steinmark, S. W., & Borkovec, T. D. (1974). Active and placebo treatment effects on moderate insomnia under counterdemand and positive demand instructions. Journal of Abnormal Psychology, 83, 157–163. https://doi.org/10.1037/h0036489

Taylor, D. J., & Bramoweth, A. D. (2010). Patterns and consequences of inadequate sleep in college students: Substance use and motor vehicle accidents. Journal of Adolescent Health, 46, 610–612. https://doi.org/10.1016/j.jadohealth.2009.12.010

Taylor, D. J., Gardner, C. E., Bramoweth, A. D., Williams, J. M., Roane, B. M., Grieser, E. A., & Tatum, J. I. (2011). Insomnia and mental health in college students. Behavioral Sleep Medicine, 9, 107–116. https://doi.org/10.1080/15402002.2011.557992

Taylor, D. J., Zimmerman, M. R., Gardner, C. E., Williams, J. M., Grieser, E. A., Tatum, J. I., ... Ruggiero, C. (2014). A pilot randomized controlled trial of the effects of cognitive-behavioral therapy for insomnia on sleep and daytime functioning in college students. Behavior Therapy, 45, 376–389. https://doi.org/10.1016/j.beth.2013.12.010

Thurber, C. A., & Walton, E. A. (2012). Homicidess and adjustment in university students. Journal of American College Health, 60(5): 415–419.

van Houdenhove, L., Buyse, B., Gabriëls, L., & van den Bergh, O. (2011). Treating primary insomnia: Clinical effectiveness and predictors of outcomes on sleep, daytime function and health-related quality of life. Journal of Clinical Psychology in Medical Settings, 18, 312–321. https://doi.org/10.1007/s10880-011-9250-7

van Straten, A., van der Zweerde, T., Kleinboer, A., Cuipers, P., Morin, C. M., & Lancee, J. (2018). Cognitive and behavioral therapies in the treatment of insomnia: A meta-analysis. Sleep Medicine Reviews, 38, 3–16. https://doi.org/10.1016/j.smrv.2017.02.001

Weinberg, M. K., Noble, J. M., & Hammond, T. G. (2016). Sleep well feel well: An investigation into the protective value of sleep quality on subjective well-being. Australian Journal of Psychology, 68, 91–97. https://doi.org/10.1111/ajp.12098

Wong, M. L., Lau, E. Y. Y., Wan, J. H. Y., Cheung, S. F., Hui, C. H., & Mok, D. S. Y. (2013). The interplay between sleep and mood in predicting academic functioning, physical health and psychological health: A longitudinal study. Journal of Psychosomatic Research, 74, 271–277. https://doi.org/10.1016/j.jpsychores.2012.08.014

Wu, J. Q., Appleman, E. R., Salazar, R. D., & Ong, J. C. (2015). Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: A meta-analysis. JAMA Internal Medicine, 175, 1461–1472. https://doi.org/10.1001/jamainternmed.2015.3006

Zwart, C. A., & Lism, S. A. (1979). Analysis of stimulus control treatment of sleep-onset insomnia. Journal of Consulting and Clinical Psychology, 47, 113–118. https://doi.org/10.1037/0022-006X.47.1.113

How to cite this article: Saruhanjan K, Zarski A-C, Bauer T, et al. Psychological interventions to improve sleep in college students: A meta-analysis of randomized controlled trials. J Sleep Res. 2021;30:e13097. https://doi.org/10.1111/jsr.13097
| Search | Query |
|--------|-------|
| #77    | Add Search (#19 AND #33 AND #76) |
| #76    | Add Search (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54 OR #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61 OR #62 OR #63 OR #64 OR #65 OR #66 OR #67 OR #68 OR #69 OR #70 OR #71 OR #72 OR #73 OR #74 OR #75) |
| #75    | Add Search (anxiety OR fear OR dental anxiety OR panic[MeSH Terms]) |
| #74    | Add Search (brain[Text Word] AND doping[Text Word]) |
| #73    | Add Search pornography[Text Word] |
| #72    | Add Search (problematic[Text Word] AND internet[Text Word]) |
| #71    | Add Search (internet[Text Word] AND addiction[Text Word]) |
| #70    | Add Search (social[Text Word] AND media[Text Word]) |
| #69    | Add Search (cognitive AND enhance*) |
| #68    | Add Search neuroenhance[Text Word] |
| #67    | Add Search nicotine[Text Word] |
| #66    | Add Search (test[Text Word] AND anxiety[Text Word]) |
| #65    | Add Search self-efficacy[Text Word] |
| #64    | Add Search (time[Text Word] AND management[Text Word]) |
| #63    | Add Search time manag*[Text Word] |
| #62    | Add Search (social[Text Word] AND skill*[Text Word]) |
| #61    | Add Search (social[Text Word] AND competence[Text Word]) |
| #60    | Add Search stress*[Text Word] |
| #59    | Add Search resilience[Text Word] |
| #58    | Add Search (health[Text Word] AND behaviour[Text Word]) |
| #57    | Add Search (health[Text Word] AND behavior) |
| #56    | Add Search (behaviour[Text Word] AND modification[Text Word]) |
| #55    | Add Search (behaviour[Text Word] AND modification[Text Word]) |
| #54    | Add Search (behaviour[Text Word] AND change[Text Word]) |
| #53    | Add Search (behaviour[Text Word] AND change[Text Word]) |
| #52    | Add Search eating*[Text Word] |
| #51    | Add Search weight*[Text Word] |
| #50    | Add Search procrastination*[Text Word] |
| #49    | Add Search rumination*[Text Word] |
| #48    | Add Search fear*[Text Word] |
| #47    | Add Search worry*[Text Word] |
| #46    | Add Search obesity*[Text Word] |
| #45    | Add Search exercise*[Text Word] |
| #44    | Add Search (physical[Text Word] AND activity[Text Word]) |
| #43    | Add Search smoking*[Text Word] |
| #42    | Add Search marijuana*[Text Word] |
| #41    | Add Search cannabis*[Text Word] |
| #40    | Add Search substance*[Text Word] |
| #39    | Add Search drug*[Text Word] |
| #38    | Add Search alcohol*[Text Word] |
| #37    | Add Search HIV*[Text Word] |
| #36    | Add Search condom*[Text Word] |
| #35    | Add Search (sexual[Text Word] AND health[Text Word]) |
Search (mental disorders OR anxiety disorders OR bipolar and related disorders OR disruptive, impulse control, and conduct disorders OR dissociative disorders OR elimination disorders OR feeding and eating disorders OR mood disorders OR neurocognitive disorders OR neurodevelopmental disorders OR neurotic disorders OR paraphilic disorders OR personality disorders OR schizophrenia spectrum and other psychotic disorders OR sexual dysfunctions, psychological OR sleep wake disorders OR somatoform disorders OR substance-related disorders[MeSH Terms])

#33 Add Search (#32 NOT #31)
#32 Add Search (#20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #28 OR #29 OR #30)
#31 Add Search (animals [mh] NOT (animals [mh] AND humans [mh]))
#30 Add Search clinical trials as topic [mh]
#29 Add Search trial[Title]
#28 Add Search (#26 AND #27)
#27 Add Search (waitlist*[Title/Abstract] OR wait* list*[Title/Abstract] OR treatment as usual[Title/Abstract] OR TAU[Title/Abstract])
#26 Add Search (control*[Title/Abstract] OR group* 1[Title/Abstract])
#25 Add Search placebo*[Title/Abstract]
#24 Add Search randomly[Title/Abstract]
#23 Add Search randomised[Title/Abstract]
#22 Add Search randomized[Title/Abstract]
#21 Add Search controlled clinical trial[Publication Type]
#20 Add Search randomized controlled trial[Publication Type]
#19 Add Search (#14 OR #15 OR #16 OR #17 OR #18)
#18 Add Search (tertiary*[Title/Abstract] AND education*[Title/Abstract])
#17 Add Search (college*[Title/Abstract] AND student*[Title/Abstract])
#16 Add Search (undergraduate*[Title/Abstract] AND student*[Title/Abstract])
#15 Add Search (university*[Title/Abstract] AND student*[Title/Abstract])
#14 Add Search student*[Text Word]
#13 Add Search internet*[Text Word]
#12 Add Search online*[Text Word]
#11 Add Search web*[Text Word]
#10 Add Search online therap*[Text Word]
#9 Add Search therap*[Text Word]
#8 Add Search psychotherap*[Text Word]
#7 Add Search program*[Text Word]
#6 Add Search psychoeducation*[Text Word]
#5 Add Search course*[Text Word]
#4 Add Search counsel*[Text Word]
#3 Add Search treatment*[Text Word]
#2 Add Search training*[Text Word]
#1 Add Search intervention*[Text Word]
APPENDIX B

Subgroup analyses of potential moderators between psychological treatment and sleep relevant outcomes

| Covariate                | Q       | df | p   |
|--------------------------|---------|----|-----|
| Mode of delivery <sup>a</sup> | 0.16870 | 1  | .681 |
| Risk of bias <sup>b</sup>     | 0.61030 | 1  | .435 |
| Type of treatment <sup>c</sup> | 2.53953 | 2  | .281 |

<sup>a</sup>Separated in face-to-face treatment and self-help treatment.
<sup>b</sup>Separated in high and low risk of bias.
<sup>c</sup>Separated in behavioral, cognitive, educational and cognitive behavioral treatment

APPENDIX C

Effects of psychological treatment on global measure of sleep disturbances compared to passive control groups

| Study name         | Outcome                | Statistics for each study | Hedges's g and 95% CI |
|--------------------|------------------------|---------------------------|-----------------------|
|                    |                        | Hedges's g    | Lower limit | Upper limit | p-Value | Favors intervention | Favors control |
| Borkovec et al. (1976) | Combined              | -0.59         | -1.42       | 0.24        | .164    |                     |               |
| Carrera et al. (1980) | SOL (diary)           | -0.96         | -1.62       | -0.30       | .004    |                     |               |
| Means et al. (2000)  | SOL (diary)           | -0.29         | -0.81       | 0.22        | .266    |                     |               |
| Steinmark et al. (1974) | SOL (diary)         | -0.73         | -1.53       | 0.07        | .075    |                     |               |
| Taylor et al. (2014)  | Combined              | -0.87         | -1.62       | -0.12       | .023    |                     |               |
| Zwart et al. (1979)   | SOL (diary)           | -0.91         | -1.91       | 0.10        | .078    |                     |               |
| Sleep onset latency overall |                    | -0.65         | -0.94       | -0.36       | .000    |                     |               |

<sup>Note.</sup> Combined: Mean of all reported sleep onset latency outcomes (via sleep diary, questionnaire, actigraphy and EEG); SOL: Sleep onset latency
APPENDIX D
Effects of psychological treatment on sleep onset latency compared to passive control groups

| Study name          | Outcome       | Statistics for each study | Hedges's g and 95% CI |
|---------------------|---------------|----------------------------|-----------------------|
| Barber et al. (2017)| SQ (diary)    | -0.29 –0.73 0.16 .207     |                       |
| Freeman et al. (2017)| Combined    | -0.72 –0.82 –0.62 .000    |                       |
| Gao et al. (2014)   | PSQI          | -1.29 –1.75 –0.82 .000    |                       |
| Means et al. (2000) | SQ (diary)    | -0.99 –1.53 –0.44 .000    |                       |
| Morris et al. (2016)| PSQI          | -0.51 –0.91 –0.10 .015    |                       |
| Taylor et al. (2014)| Combined    | -1.43 –2.23 –0.62 .000    |                       |
| Global measures of sleep disturbances overall | -0.79 –1.06 –0.52 .000 |                       |

Note. Combined: Mean of all reported sleep quality outcomes (PSQI, Sleep Condition Indicator – 8 items, Insomnia Severity Index, SQ [diary]); PSQI: Pittsburgh Sleep Quality Index; SQ (diary): Sleep quality reported by sleep diary