Sleep loss and emotion: A systematic review and meta-analysis of over fifty years of experimental research

Cara A. Palmer1*, Joanne L. Bower2*, Kit W. Cho3, Michelle A. Clementi4, Simon Lau5, Benjamin Oosterhoff6, Candice A. Alfano5,7

1Montana State University, Department of Psychology, Bozeman, MT, United States
2University of East Anglia, School of Psychology, Norwich, United Kingdom
3University of Houston-Downtown, Houston, TX, United States
4University of Colorado Anschutz Medical Campus, Aurora, CO, United States
5University of Houston, Department of Psychology, Houston, TX, United States
6The Hackett Center for Mental Health, Meadows Mental Health Policy Institute, Houston, TX, United States
7HEALTH Research Institute, University of Houston, Houston TX

*These authors contributed equally to this work and share first authorship.

Corresponding authors:
Cara A. Palmer, Ph.D. Joanne L Bower, Ph.D.
319 Traphagen Hall School of Psychology
Department of Psychology University of East Anglia
Montana State University Norwich Research Park
Bozeman, MT 59717 Norwich, NR4 7TJ, UK
cara.palmer@montana.edu joanne.bower@uea.ac.uk

Author Note
This study design and analytic plan were preregistered at crd.york.ac.uk/prospero/display_record.php?ID=CRD42016045754. The datasets generated and analyzed during the current study, along with all analytic code are available at https://osf.io/fy349/?view_only=6c8b9edc094a39ba7e2696b8a0041e. The findings from this manuscript have not been disseminated elsewhere.
Abstract

In a largely sleep-deprived society, quantifying the effects of sleep loss on emotion is critical for promoting psychological health. This preregistered systematic review and meta-analysis quantified the effects of various forms of sleep loss on multiple aspects of emotional experiences. Eligible studies used experimental reductions of sleep via total sleep deprivation, partial sleep restriction, or sleep fragmentation in healthy populations to examine effects on positive affect, negative affect, general mood disturbances, emotional reactivity, anxiety symptoms, and/or depressive symptoms. In total, 1338 effect sizes across 154 studies were included (N = 5,717; participant age range = 7-79 years). Random effects models were conducted, and all forms of sleep loss resulted in reduced positive affect (SMD = -0.27 to -1.14), increased anxiety symptoms (SMD = 0.57 to 0.63), and blunted arousal in response to emotional stimuli (SMD = -0.20 to -0.53). Findings for negative affect, reports of emotional valence in response to emotional stimuli, and depressive symptoms were mixed and depended on the type of sleep loss. Non-linear effects for the amount of sleep loss as well as differences based on the stage of sleep restricted (i.e., rapid eye movement sleep or slow wave sleep) were also detected. This study represents the most comprehensive quantitative synthesis of experimental sleep and emotion research to date, and provides strong evidence that periods of extended wakefulness, shortened sleep duration, and/or nighttime awakenings adversely influence human emotional functioning. Findings provide an integrative foundation for future research on sleep and emotion and elucidate the precise ways that inadequate sleep may impact our daytime emotional lives.

Keywords: emotion; mood; mental health; sleep; meta-analysis
Impact/Public Significance Statement

This study synthesizes more than 50 years of experimental research to reveal the multiple ways loss of sleep undermines human emotional functioning, and increases risk for psychiatric disorders. Implications for individual and public health are considerable in a largely sleep-deprived society. Industries and sectors prone to sleep loss (e.g., first responders, pilots, truck drivers) should develop and adopt policies that prioritize sleep to mitigate against the risks to daytime function and wellbeing.
Sleep loss and emotion: A systematic review and meta-analysis of over fifty years of experimental research

Sleep and emotion share an intimate relationship. Just as sleep is a universal human need, emotional experiences are ubiquitous and modulate virtually every aspect of our day-to-day lives, including our health, motivation, decision making, social interactions, learning, and memory (Engen & Anderson, 2018; Ferrer & Mendes, 2018; Fredrickson, 2001; Gallo et al., 2005; Sznycer et al., 2022; Tyng et al., 2017). There is overwhelming evidence that poor sleep serves as a direct catalyst for the development of emotional difficulties and affective disorders across the lifespan (Goldstein & Walker, 2014; Gregory et al., 2009; Kelly & El-Sheikh, 2014; Triantafillou et al., 2019). The practical implications of these relationships are considerable and far-reaching, as epidemiological studies confirm large segments of society regularly fail to obtain adequate sleep (Hafner et al., 2017). In the United States, more than 30% of adults and as many as 90% of teenagers obtain less than the recommended amount of nightly sleep (Basch et al., 2014; Liu, 2016), and mobile technology data provide an indication of similar truncated sleep patterns across the world (Walch et al., 2016). Quantifying the effects of sleep loss on human emotional responses therefore has direct relevance not only for individual and public health, but for the policies and regulations that govern our lives (e.g., school start times for youth, maximum work periods for certain occupational roles, and military/combat operations).

A substantial body of experimental research accumulated in recent decades demonstrates adverse effects of inadequate sleep on emotion (Palmer & Alfano, 2017; Tempesta et al., 2018). The first meta-analysis of these effects was published over 25 years ago and examined the effects of sleep loss on cognitive, motor, and mood domains (Pilcher & Huffcutt, 1996). At the time, only 20 out of 143 identified effect sizes focused on mood. However, while each functional
domain examined (e.g. cognitive performance, motor performance, mood) was adversely impacted by sleep loss, the largest effects were observed for changes in mood. Since this seminal meta-analysis, studies of sleep’s impact on affective responses have rapidly proliferated, informed by increasingly sophisticated theoretical models/frameworks (Palmer & Alfano, 2017; Walker & van der Helm, 2009). Different types of sleep manipulations (e.g., deprivation of sleep for extended periods, truncated sleep periods, and fragmenting sleep throughout the night) have all been shown to induce adverse changes in the generation of emotions and, to a lesser extent, emotion regulatory abilities (cf. Palmer & Alfano, 2017). Precise mechanisms underlying these effects are complex, but prevailing models emphasize sleep’s influence over connectivity between limbic structures involved in emotional arousal and pre-frontal regions that control emotional responses, rapid eye movement (REM) sleep’s modulatory role in the processing of emotional experiences, and, more recently, the importance of slow wave sleep (SWS) for positive affect and emotion (Finan et al., 2015, Finan et al., 2016, Goldstein & Walker, 2014; Kahn et al., 2013; Tempesta et al., 2018).

Recent qualitative reviews that have attempted to collate and synthesize findings from these experimental studies generally conclude that inadequate sleep poses a broad risk to emotional functioning, though specific relationships have been difficult to quantify (Beattie et al., 2015; Palmer & Alfano, 2017; Tempesta et al., 2018). For example, Tempesta and colleagues (2018) determined that evidence for the impact of sleep loss on emotional reactivity (i.e., pleasantness and/or arousal in response to emotional stimuli or situations) is somewhat mixed, citing studies finding no changes, increases, and decreases in reactivity following sleep loss. These authors underscore the small sample sizes and considerable heterogeneity of research designs and methods that characterize this body of research, including different experimental
paradigms, manipulations of sleep, and measurement of emotional constructs. Other recent reviews of this literature have concluded that sleep deprivation reliably impacts the neurobiological processing of emotion, but effects on subjective emotional reports are less consistent (Beattie et al., 2015). As a result, the adverse impacts of sleep loss on human emotion, while perhaps intuitive, are still inadequately understood.

Complicating matters further, many experimental sleep studies do not adequately define and/or differentiate among emotional outcomes of interest (Palmer & Alfano, 2017). Core affect or mood has been the most commonly examined emotional outcome in response to sleep loss. These affective states are diffuse and long-lasting, and are typically not attributed to a single event or stimulus (Barrett & Bliss-Moreau, 2009; Beedie et al., 2005). Measures of mood/affect reflect overall subjective states of pleasure (positive affect) or displeasure (negative affect), which are proposed to serve distinct evolutionary functions (Fredrickson, 2001; Tooby & Cosmides, 2008) and are therefore likely to be differentially influenced by sleep loss. Despite these significant functional differences, many experimental sleep studies fail to adequately distinguish between positive and negative affect and instead measure impact on overall mood, irrespective of valence, or conflate affect with other psychiatric symptoms. Unlike these more diffuse affective states, emotional reactivity is a functional response to a specific event or stimulus (Lang & Bradley, 2010) that varies in its degree of pleasantness/unpleasantness (valence) as well as arousal (i.e., level of activation). Emotional reactivity differs considerably in intensity, duration, and its dynamic features across emotional situations. Differentiating among these constructs is not merely semantic, since alterations in affect and emotional reactivity are central to the diagnosis and management of psychiatric disorders, and in non-clinical samples,
day-to-day changes in emotional experiences predict psychological well-being and life satisfaction (Beal & Ghandour, 2011; Burrow et al., 2014; Cohn et al., 2009).

A recent meta-analysis (Tomaso et al., 2021) examined the effects of both total sleep deprivation and partial sleep restriction on several emotional outcomes, including positive and negative mood (which encompassed measures of affect, along with specific symptoms of anxiety and depression) as well as measures of emotional arousal in response to emotional stimuli. Large effects of sleep loss were observed for positive mood, medium effects for negative mood, and small effects for emotional arousal. Overall, these findings represent a major step forward in quantifying the effects of sleep loss on some aspects of human emotion, but additional questions remain. There is still a need to quantify the impact of other aspects of sleep loss (i.e., sleep fragmentation) and to clearly delineate emotional constructs (e.g., valence and arousal, depression/anxiety vs. affect).

The Present Research

To integrate and consolidate findings across this heterogeneous research literature, a comprehensive quantitative synthesis is needed. The current pre-registered study is the most comprehensive systematic review and meta-analysis of experimental sleep and emotion research to date. Our first aim was to examine how various forms of experimental sleep reduction (i.e., sleep deprivation, sleep restriction, and sleep fragmentation) impact different types of emotional experiences, including positive affect, negative affect, general mood disturbance (i.e., studies that include mixed valence mood states as an outcome measure), and emotional reactivity (i.e., reports of valence and arousal in response to emotional stimuli). Due to problematic emotional experiences that characterize many affective disorders (APA, 2013), we also included studies that assessed anxiety symptoms and depressive symptoms as specific outcomes. Given the
considerable methodological heterogeneity across studies, we examined how the ‘dose’ (i.e., amount) of sleep loss relates to effects on emotional outcomes including whether these associations were linear or non-linear. Further, consistent with theoretical models positing the importance of both REM sleep (van der Helm & Walker, 2009) and SWS (Finan et al., 2015, 2016) for emotion-based outcomes, we examined how effect sizes differ in relation to loss of REM sleep versus SWS. Finally, based on prior research that has found inter-individual differences in responsivity to sleep loss (Boccabella & Malouf, 2017; Conklin et al., 2018; Zitting et al., 2018), we examined both age and sex as moderators of emotional outcomes.

Notably, we did not examine sleep’s impact on emotion regulation in the current set of analyses for a few reasons. First, as described in previous work (Palmer & Alfano, 2017), a majority of available studies do not adequately disentangle emotion regulation from the generation of discrete emotions or overall emotional reactivity, and/or conflate specific constructs. Further, in their recent meta-analysis, Tomaso and colleagues (2021) attempted to examine the effects of sleep loss on emotion regulation. The authors noted several other limitations in this small body of literature that preclude firm conclusions at this time, including insufficient variability in the types of emotion regulatory strategies examined, the use of sleep restriction protocols only, and a focus on youth rather than adult populations. Thus, we believe this emerging literature is inadequate for meta-analytic procedures or conclusions at this time.

Based on extant research, we expected all forms of sleep reduction to impact each emotional outcome, with the largest effect sizes for positive affect compared to other emotional outcomes. We also expected that greater amounts of sleep loss (i.e., more hours of wakefulness in full deprivation studies, and fewer hours of sleep in restriction studies) and loss of REM sleep
(compared to SWS) would produce larger effects. Finally, we expected that the effects of sleep on emotional outcomes would be larger for younger samples and female participants.

**Method**

**Search Strategy and Study Identification**

Search criteria included human studies available in any language with mention of experimental manipulations of sleep, and any type of emotion-related outcome in the title and/or abstract. In attempt to mitigate the file drawer problem (Strube & Hartmann, 1983), and to estimate publication bias using both published and unpublished data, our search was open to both published and unpublished records (e.g., theses/dissertations, conference abstracts). Searches were conducted using PubMed/MEDLINE, PsychINFO, and Web of Science. We had no restrictions on the location or date that the study took place, and all studies available through December 2022 were included. Our full search criteria, dates, and procedures can be found in Supplemental File 1. Our searches returned 14,936 records after duplicates were removed. All abstracts and titles of the search records were reviewed to determine their eligibility for the meta-analysis by the two first authors (CP and JB). We also obtained data and articles by (a) placing a call for unpublished data on several sleep-related listservs, (b) reviewing reference lists of all included studies and several qualitative review articles on sleep and emotion, (c) hand searches in the journals *Sleep* and *Journal of Sleep Research*, and (d) review of our reference list for missing studies by several content experts. These methods revealed an additional 17 studies. In total, 365 records were selected for full review.

**Study Screening and Inclusion**

Eligibility criteria for inclusion were:

1) Participants did not have any known psychiatric or sleep disorders, or any reported medical conditions known to impact sleep and/or emotional functioning. We also
excluded any specialized samples with high risk for sleep or circadian disruptions (e.g., pilots, military personnel, medical professionals). We placed no restrictions on the age of study participants.

2) The study experimentally manipulated reductions in nighttime sleep, including sleep deprivation (SD), partial sleep restriction (SR), or sleep fragmentation (SF) paradigms over the course of one or more nights. Nap studies (i.e., studies manipulating daytime sleep), forced desynchrony, and phase shift protocols were not included. Studies or conditions using sleep extension or recovery sleep protocols were not included.

3) An adequate comparison or control group condition was used. We defined this as a baseline indicator of the dependent variable of interest, before the sleep manipulation, after an adequate or typical night of sleep (for within-subjects designs), or a control group who completed similar assessments of the dependent variable after an adequate or typical night of sleep (for between-subjects designs).

4) At least one emotion-related dependent variable was assessed after the sleep manipulation. Measures could include subjective reports of affect/mood, and/or subjective reports of emotional valence or arousal in response to an emotional stimulus. Since many depressive and anxiety symptom questionnaires include components of negative mood, (lack of) positive mood, along with other related symptoms (e.g., somatic symptoms), these were included as separate indicators of emotional experience. We did not include any studies that only reported fatigue subscales that were a part of larger mood scales (e.g., Profile of Mood States) if no other emotion or mood outcomes were included. The pre-registration for this meta-analysis also initially included physiological or behavioral indicators of emotional
experience in response to an emotional stimulus (e.g., facial expressions, skin conductance response, cortisol), but due to the small number and heterogeneity of studies focused on these outcomes, we opted to not include these in our final dataset. We did not include any neurological or neuroimaging (e.g., event related potentials, fMRI) dependent variables as the study aims were focused on subjective experiences of affective functioning following sleep loss, thus these were beyond the scope of the current study.

5) No other intervention or study procedure occurred prior to the assessment of the dependent variable that was expected to affect emotional experience or sleep (e.g., intense exercise or extreme training, simulated work shifts, emotion extinction paradigms, pharmacological agents).

All articles and unpublished datasets were reviewed for inclusion criteria by two independent coders and disagreements were resolved during regular consensus meetings (CP and JB; 98.5% agreement, $\kappa = .95$). If sufficient data for calculating effect sizes and variance were not reported, authors were contacted to request the necessary statistics and/or raw data. The final number of studies that were included in the final meta-analysis were 146 peer-reviewed studies published between 1966-2022 and 8 unpublished studies or datasets (e.g., dissertations, additional data provided by authors). A full list of reports that were reviewed but not included, along with the reasons for exclusion, can be found at the following link: https://osf.io/fy349/?view_only=6c8b9edbd094a39ba7e2696b8a0041e. Descriptive statistics for participant and study characteristics for all included studies appear in Table 1. Tables with details for each individual study, including author names, country, sample characteristics, study design, sleep manipulation, outcome variables, and findings are available in Supplemental File 2.
A flow diagram outlining the search, and final included and excluded studies is displayed in Figure 1.

**Data Extraction**

Two reviewers (CP and JB) independently extracted and coded all study codes and data using a coding sheet developed in Comprehensive Meta-Analysis 3.3.070 (Borenstein et al., n.d.). All codes were verified during regular consensus meetings. When studies reported data for the dependent variable(s) for multiple sleep conditions (e.g., after restricting sleep to 4 hours or 2 hours, or after conditions of both SR and SF) or at multiple timepoints during an ongoing sleep condition (e.g., collected emotion reports hourly from 24-48 hours of SD), all available effect sizes and conditions were coded.

**Sleep Manipulation.** SD studies included paradigms where participants were asked to stay awake with no intermittent periods of sleep for extended time periods. We coded the length of deprivation (i.e., hours of wakefulness) for any effect size occurring at 18+ hours of wake. This cut-off was chosen based on minimum sleep duration recommendations by the National Science Foundation for adults (7 hours; Hirshkowitz et al., 2015), which implies that more than 17 hours awake would be an extended period of wakefulness. Effect sizes coded ranged from 18-110 hours of SD. In several instances, a range of SD times were reported (e.g., all participants were awake from 36-38 hours), and in these circumstances we used the mid-point of the range reported. For studies that did not report exact wake times prior to the SD paradigm, the hours of SD were estimated (e.g., 24 hours for studies using 1 night of total SD and testing in the morning).

SR studies included paradigms that shortened typical sleep duration for one or more nights. Across studies, we coded the sleep opportunity provided for participants (e.g.,
participants were allowed to sleep from 4:00-8:00), which we used to determine sleep duration for each effect size. Some studies also reported actual hours of sleep obtained within that sleep opportunity (i.e., using objective measures such as actigraphy or polysomnography), and we used this more precise data for sleep duration when it was available. Sleep duration ranged from 125-390 minutes for adult studies and 117-502 minutes for pediatric studies. A number of studies also included emotional measurements after multiple days of SR, which ranged from 1-42 days (mean = 2.64, SD = 3.63, median = 2, mode = 1). For studies that reported sleep duration or sleep opportunity across multiple nights of restriction, we used data from the most recent night prior to the measure of the dependent variable.

Studies using SF included paradigms that forced awakenings during a night of sleep. Similar to SR studies, we also coded the total sleep duration or sleep opportunity (and used actual sleep duration when available). We also coded whether SR and SF studies likely reduced primarily SWS or REM sleep. Specifically, since SWS is more abundant in the first half of the sleep period and greater REM occurs during the latter half of the sleep period, we coded whether effect sizes varied based on whether participants’ sleep was restricted in the early or late portion of the typical sleep period for SR studies, or whether REM/SWS was directly targeted in the fragmentation studies (i.e., through forced awakenings during particular sleep stages determined by polysomnography).

**Emotional Outcomes**

*Positive Affect.* A number of studies included outcomes that assessed positive affect. These included studies that primarily used the positive affect subscale from the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988), Visual Analogue Scales assessing mood/affect that was positively-valenced (e.g., calm), and the vigor subscale from the Profile of
Mood States (POMS; Heuchert & McNair, 2012; McNair et al., 1981). Because items on the POMS vigor scale primarily assess “energy” (e.g., active, full of pep, vigorous), we also coded whether or not positive affect was assessed using the POMS or a different measure (e.g., PANAS).

**Negative Affect.** Studies assessing negative affect primarily included the negative affect subscale from the PANAS (Watson et al., 1988), Visual Analogue Scales of negative affect (e.g., anger), and negative mood subscales from the POMS (Heuchert & McNair, 2012; McNair et al., 1981).

**General Mood Disturbance.** There were a number of studies that included outcomes that incorporated a mix of positively- and negatively-valenced emotional experiences (i.e., it was impossible to distinguish whether effects were driven by changes in positive affect, negative affect, or both). As a result, we also included a general mood disturbance outcome that encompassed these outcomes. These included measures such as the total POMS scale (which incorporates subscales that measure positive and negative affect; Heuchert & McNair, 2012; McNair et al., 1981) and Visual Analogue Scales using bipolar anchoring (e.g., calm as a response option on one end of the scale and anxious as a response option on the other end of the scale), and reports of emotion without a clear valence (e.g., surprise).

**Emotional Reactivity: Valence and Arousal.** A number of studies included subjective emotional reports in response to various emotional stimuli or emotional experiences, including images from the International Affective Picture System (IAPS), or a laboratory stressor (e.g., completing a difficult task and receiving negative feedback). Theories of emotion separate emotional reactivity along the dimensions of valence (i.e., degree of feeling pleasant or unpleasant) and arousal (i.e., intensity of emotional response or activation level) (Russell &
Barrett, 1999). Thus, we separated these two important indicators of emotional reactivity in our analyses. Reports were frequently measured using the Self-Assessment Manikin (Lang, 1980) or Visual Analogue Scales. All effect sizes were coded so that higher values indicated reports of greater arousal or more unpleasantness.

It is important to note that some studies explicitly asked participants to report on how they felt while viewing an image, while undergoing a stressor, or immediately after, which are measures of reactivity to a stimulus. However, other studies asked participants to report whether an emotional stimulus itself is unpleasant/pleasant or arousing, which does not assess emotional reactivity, but rather emotional perception or appraisal. Thus, when studies explicitly provided instructions directed at assessing the participants perception or appraisal of an emotional stimulus, and not the participant’s own emotional reactions, they were not included.

Anxiety Symptoms. We opted to include measures of anxiety as a separate dependent variable from our other mood outcomes, since many anxiety assessments include physiological (e.g., rapid heart rate) and cognitive (e.g., worry) symptoms in addition to emotional experiences. The studies assessing anxiety symptoms primarily included reports on the State-Trait Anxiety Inventory (Spielberger et al., 1983), the anxious arousal subscale on the Mood and Anxiety Symptom Questionnaire (Watson et al., 1995), or clinician-rated scales (e.g., Hamilton Anxiety Rating Scale; Hamilton, 1959).

Depressive Symptoms. Similar to anxiety, we opted to include a separate depressive symptom dependent variable since many depressive symptom assessments include physical, cognitive, and behavioral symptoms along with emotional experiences (e.g., anhedonia). Various scales assessing depressive symptoms were used in the included studies, such as the Beck Depression Inventory (Beck et al., 1996), Hamilton Depression Rating Scale (Hamilton, 1960),
Minnesota Multiphasic Personality Inventory - Depression scale (Marek et al., 2015), the Anhedonia scale on the Psychotomimetic States Inventory (Mason et al., 2008), the Patient Health Questionnaire-9 (Kroenke et al., 2001), the Mood and Anxiety Symptom Questionnaire Anhedonia scale (Watson et al., 1995), and clinician rating scales.

**Effect Size Calculation.** The standardized mean difference (Hedges’ g) and standard errors were calculated for each effect size using established formulas (see Borenstein et al., 2021). All effect sizes were double coded independently by CP and JB and verified during regular consensus meetings. If a discrepancy was present, both authors went back to the original data or paper and independently recoded the effect size to determine the correct values. Following this procedure, effect size agreement was 100%. If codable statistics were not available, or if the direction of the effect was unable to be determined based on the data reported, the author was contacted. If authors did not respond or could not provide data, these effects were excluded. In some instances, both between and within-subjects results were reported. Many studies did not report the full necessary statistics for calculating effect sizes for within-subjects data (e.g., did not include a correlation between pre-sleep manipulation and post-sleep manipulation for the outcome variables of interest), so in these scenarios we opted to report between-subjects effects when possible. For within-subjects studies without relevant between-subjects data that were missing pre-post correlation data or adequate data on change scores, we coded the available data and estimated the pre-post correlation in our analyses by taking the average pre-post correlation for each outcome type from the data that was available (all estimated correlations ranged from .45-69). For within-subjects designs that reported multiple baseline emotion-related outcomes prior to sleep loss, we chose the baseline that corresponded most closely to the time of day of the experimental measurement in attempt to minimize
circadian effects. When a baseline value at the same approximate time of day was not available, we chose the most recent baseline measurement. Interpretation of effect sizes was based on standard Cohen’s $d$ cut offs (above 0.2 for small effects, 0.5 for medium effects and 0.8 for large effects; Cohen, 1988).

**Other Study Characteristics.** In addition to effect size, data was extracted for study design and characteristics for each effect size (published or unpublished, procedure details, between-subjects or within-subjects design), participant characteristics (sample size, demographics, sample type, participant exclusion/inclusion criteria), and study location (if no location was reported, we based location on the country of origin of any grant data or university ethics IRB reported, or location of the corresponding and/or first author). The control or comparison condition procedures and method of measuring adherence to the sleep condition (e.g., actigraphy or polysomnography) was also extracted. When data separated by sex or age group was available, effect sizes were coded separately for each group. When data were only available in the aggregate, we coded mean age of the sample, along with the sex distribution (coded as % female).

**Risk of Bias/Study Quality Coding.** All effect sizes were coded for study quality using established relevant metrics of risk using the Cochrane Risk of Bias tool (Higgins et al., 2022). Additional risk of bias criteria relevant to this literature were also assessed. As risk of bias occasionally varied within study for different effect sizes, a separate risk of bias score was calculated for each individual effect size. Each risk of bias measure was calculated as 1 (high risk or unclear/not enough information reported) or 0 (low risk). A total risk of bias score was calculated for each effect size by summing scores for all categories (possible range = 0-8, with higher scores indicating higher risk of bias) and was included in analyses. An image displaying
risk of bias across all studies and categories is available in Figure 2. All risk of bias measures were double coded by CP and JB and reviewed during consensus meetings.

**Risk of Bias Categories**

*Randomization.* Studies that randomized participants to the experimental or control group (for between-subjects studies), or used a randomized cross-over design (for within-subjects studies) were coded as low risk. Studies that used systematic assignment or did not use a randomized cross-over design for within-subjects designs were characterized as high risk. A number of studies stated that they randomized participants, but provided no further detail about the randomization procedures and were coded as unclear risk of bias. Agreement among coders was excellent (92% agreement).

*Allocation concealment.* Studies that concealed randomization or group assignment up until the point of the sleep manipulation were coded as low risk, whereas studies where the group or manipulation order was possible to know or suspect in advance of the sleep manipulation were coded as high risk. Studies that did not provide enough detail were coded as unclear risk of bias. Agreement among coders was excellent (97% agreement).

*Adequate masking of participants to study condition.* By nature, sleep loss studies are unable to fully mask participants to their study condition. As a result, we did not include this criterion in our coding scheme.

*Selective outcome reporting.* Studies where expected outcomes were missing based on measures that are typically reported together (e.g., only some subscales of the POMS were reported, or the PANAS positive affect scale was reported but not the negative affect scale) were coded as high risk of bias for selective outcome reporting. Studies where all expected outcomes...
that are typically grouped together in the literature were reported were coded as low risk. Agreement among coders was excellent (93% agreement).

**Incomplete outcome data.** Studies that had missing data that was likely related to the outcome (e.g., a participant dropped out mid-study because they could not stay awake as directed) were coded as high risk. Studies with complete data were coded as low risk. Some studies did not clearly report the number across conditions or information about incomplete data was not reported and could not be deduced from other aspects of the paper (e.g., degrees of freedom), and these were coded as unclear risk of bias. Agreement among coders was acceptable (83% agreement).

**Adequacy of baseline sleep.** We coded for the quality of participants’ baseline sleep. For between-group studies, if there were similar baseline sleep patterns across groups prior to the sleep manipulation, or if they required healthy baseline sleep from participants (e.g., 8 hours a night), this was coded as low risk. For within-subject designs, if a participant’s baseline sleep prior to the experimental manipulation was typical for healthy participants or within a reasonable sleep duration (i.e., >7 hours for adults or within age-adjusted recommendations for pediatric studies), this was coded as low risk. Studies that had pre-experimental sleep patterns that differed between groups, or had no measure of baseline sleep, were coded as high risk. If differences were not reported or if this was not assessed, it was coded as unclear risk of bias. Agreement among coders was excellent (90% agreement).

**Sleep Manipulation Verification.** Some studies verified that participants adhered to the study sleep manipulation using reliable, verifiable methods. For instance, some participants were directly observed in the lab continuously, or sleep (or lack thereof) was verified using objective sleep measurement tools such as actigraphy or polysomnography. These studies were coded as
Some studies verified adherence to the sleep manipulation using self-report or sleep diaries, which may be prone to bias, or had periodic check-ins (and not continuous observation) by researchers, which does not verify the presence or absence of sleep between check-ins. These methods were coded as high risk of bias. Other studies did not report how they confirmed their sleep manipulation and were coded as unclear risk of bias. Agreement among coders was excellent (91% agreement).

**Control of Circadian Influences.** Some studies assessed the dependent variable at a different time of day between the conditions, which introduces biases due to circadian effects on mood. If the emotion measurements occurred at different phases of the day (e.g., morning, afternoon, evening, or nighttime), and/or with greater than 3 hours difference in the measurement time points, we coded these as high risk. If studies had multiple time points that ranged more than 3 hours and were only available in the aggregate, this was also coded as high risk. If the measurements occurred during the same phase of the day (e.g., both happened during the morning, within 3 hours), these were coded as low risk. If timing of the measurements were not reported or if the description was vague (e.g., “both occurred in the morning”), it was coded as unclear risk of bias. Agreement among coders was acceptable (86% agreement).

**Control of Other Possible Confounders.** A number of studies restricted activities that may alter sleep or mood during the sleep manipulation period, including the use of (a) caffeine, (b) medications that may impact sleep or mood, (c) vigorous exercise, (d) alcohol, (e) nicotine, and (f) naps. For each of these, studies were coded as low risk if this confounder was prohibited or if it did not apply to the sample (e.g., alcohol or nicotine in pediatric studies), or it was coded as high risk if they did not report prohibiting the confounder. An overall risk of bias score for these confounders was calculated by summing the scores for all of these confounders together, with a
range from 0-6. This score was weighted in the final risk of bias score so that the highest risk was coded as 1 (i.e., all 6 confounders were present) and the lowest was coded as 0 (i.e., none of the confounders were present). Agreement for each category and the total score was acceptable (82-90% agreement).

**Analytic Procedure**

Summary analyses were conducted by performing separate meta-analyses for the three types of sleep manipulations (SD, SR, and SF), and seven emotion domains (positive affect, negative affect, general mood disturbance, emotional reactivity – valence, emotional reactivity – arousal, anxiety symptoms, depressive symptoms). Sample age and sex, study design (between or within-subjects), sample size, and risk of bias score for each effect size were included as covariates in all analyses. Each study could contribute multiple effect sizes to the same sleep condition (SD, SR, or SF) if they reported effect sizes for different emotion domains and vice versa. To account for these non-independent effect sizes (i.e., nested structure), we computed multilevel/multivariate (three-level) meta-analyses. The multivariate models estimated and accounted for heterogeneity between studies, and between samples within studies.

To estimate the overall effects, multilevel random-effects models were used because of the considerable heterogeneity between studies (e.g., different populations, settings, and cultures). Mixed-effects models were used in analyses of potential moderators, which were included as fixed effects, following the recommendation of Borenstein et al. (2021). $Q$, $I^2$, and $\tau^2$ values were used to denote heterogeneity. For the $I^2$ and $\tau^2$ values, we reported values for between-studies (level 3) and within-studies (level 2) separately. Meta-analyses were performed when at least two studies were available. Forest plots for all individual types of sleep loss and emotion outcomes were created to display the effect sizes for each individual study (see
Supplemental Files 3-5), along with summary forest plots. Probability values and 95% confidence intervals were used to assess confidence in the body of evidence, with \( p \) values < .05 and 95% confidence intervals excluding zero indicating a statistically meaningful effect. Betas (\( b \)) and standard errors are reported.

For the SD paradigms, in addition to examining overall effect sizes for each emotion domain, we also examined hours of wakefulness as a predictor of effect size for each outcome. These models included both linear and non-linear (i.e., quadratic) effects of hours of wakefulness. We also examined sleep duration as a predictor of the effects for SR, including linear and non-linear (i.e., quadratic) effects. These models were likewise estimated using the multi-level meta-analyses procedures described above. Finally, some SF and SR studies primarily fragmented/restricted REM sleep or SWS, and follow-up analyses examined how these different types of sleep loss may differentially impact outcomes (adjusting for sleep opportunity/duration). For all models examining positive affect, we also examined whether or not the dependent variable was assessed using the POMS Vigor subscale (i.e., primarily measuring high arousal positive affect such as energy), or if a different measurement of positive affect was used. Finally, we examined age (mean age of the sample) and sex distribution (percentage of the sample that was female) as predictors of the effect sizes in each model to determine whether sensitivity to the emotional effects of sleep loss is associated with sex or age. Based on recommendations from Borenstein and colleagues (2021), moderator analyses were conducted when \( k \geq 10 \).

Outlier/influence analysis was conducted following the approach recommended by Viechtbauer and Cheung (2010). An effect size was defined as an outlier when it had a large influence on the pooled estimated effect. If an outlier was detected, as determined by a DFBETA
value greater than |1|, we proceeded to analyze the data with and without the extreme effect size. This process was done iteratively to assess for multiple outliers. When relevant, results from analyses without outliers are described. Statistics for models with and without outliers are presented in Supplemental Files 3-5.

To assess for possible selective reporting in the results, we conducted two tests (Egger’s regression and three-parameter selection model [3PSM]), per the recommendation of Rodgers and Pustejovsky (2021). We used the multilevel meta-analysis variant of Egger’s regression, which maintains type I error rate that could be inflated when using dependent effect sizes (Rodgers & Pustejovsky, 2021). This test assesses funnel plot asymmetry or small study effects. To further control for type I error in Egger’s test, the standard error term was corrected to remove the effect size from its computation using the equation provided in Pustejovsky and Rodgers (2019). Evidence for funnel plot asymmetry or a small study effect is denoted by a significant p value testing that the slope of the regression equation is different from 0. We also assessed selective reporting by computing the 3PSM (McShane et al., 2016; Vevea & Hedges, 1995; Vevea & Woods, 2005) to evaluate whether non-significant findings (defined in our model as p > .05) are less likely to be published relative to significant results. The presence of a selective reporting bias is indicated by a significant likelihood ratio test. Because, at present, the 3PSM can only be applied to univariate meta-analysis models, we followed the recommendation of Rodgers and Pustejovsky (2021) and aggregated the dependent effect sizes in each analysis before computing the 3PSM test. Although both the Egger’s regression test and 3PSM can be extended to models that include moderators, and we report these results in the Supplemental Files, the validity of their application in such models have not been well-studied (Rodgers & Pustejovsky, 2021). As such, these results should be interpreted with some caution.
Transparency and Openness

This study design and analytic plan were preregistered at PROSPERO (CRD42016045754). All reporting in this systematic review and meta-analysis adhered to the APA MARS guidelines (Appelbaum et al., 2018), as well as the PRISMA 2020 guidelines (Page et al., 2021). All analytic code and datasets are available at https://osf.io/fy349/?view_only=6c8b9edbdc094a39ba7e2696b8a0041e. All effect sizes were calculated using established formulas (Borenstein et al., 2021), and the full raw data and R formulas to calculate effect sizes are also provided on OSF. Meta-analytic calculations were performed using the *metafor* package (Viechtbauer, 2010) using R statistical software version 4.3 (http://www.R-project.org/).

Results

The total number of participants included in the analyses was $N = 5,717$ across 28 countries and 154 different studies. Descriptive summaries of participant and study characteristics for all included studies are provided in Table 1. Participants were predominantly young adults ($M = 23.66$, $SD = 8.31$, range 6-79), with an approximately equal overall number of male and female participants (50.14% female). Studies were predominantly conducted in the USA or Europe, and most commonly used a within-subjects design. The full list of individual studies and their sample characteristics is included in Supplemental File 2.

Sleep Deprivation (SD)

Eighty-five studies used SD paradigms, with a total of 599 effect sizes. Summary forest plots for all SD studies for each emotional outcome are available in Figure 3. There were 45 studies ($k = 190$) examining the association between SD and positive affect. Heterogeneity was significant, $\chi^2(189) = 981.72$, $p < .001$; between-studies: $I^2 = 61\%$, $\tau^2 = 0.23$; within-studies: $I^2 = 29\%$, $\tau^2 = 0.11$. The overall effect was negative and significant (standardized mean difference
[SMD] = -0.86; 95% CI [-1.05, -0.72]), indicating that SD resulted in a decrease in positive affect. Post-hoc moderation analyses examining whether positive affect was assessed using the POMS indicated that effect size was significantly moderated by measurement type ($b = 0.46, SE = 0.09, p < .001, 95\% CI [0.29, 0.62]$). Effect sizes were larger for studies that used the POMS compared to studies using other measures (POMS SMD = -1.14, SE = 0.16, 95\% CI [-1.46, -0.82]; Non-POMS SMD = -0.76, SE = 0.07, 95\% CI [-0.91, -0.62]). There were 48 studies ($k = 278$) examining negative affect. Heterogeneity was significant, $\chi^2(279) = 1245.77, p < .001$; between-studies: $I^2 = 56\%, \tau^2 = 0.11$; within-studies: $I^2 = 27\%, \tau^2 = 0.06$, and the overall effect was significant, SMD = 0.37, 95\% CI [0.26, 0.48], indicating an increase in negative affect after SD. There were 23 studies ($k = 57$) examining general mood disturbance. Heterogeneity was significant, $\chi^2(56) = 245.27, p < .001$; between-studies: $I^2 = 84\%, \tau^2 = 0.36$; within-studies: $I^2 = 5\%, \tau^2 = 0.02$, as well as the overall effect, SMD = 0.71, 95\% CI [0.44, 0.98], indicating that SD increased mood disturbances.

There were 6 ($k = 27$) studies examining emotional valence. Heterogeneity was significant, $\chi^2(26) = 76.47, p < .001$; between-studies: $I^2 = 75\%, \tau^2 = 0.20$; within-studies: $I^2 = 6\%, \tau^2 = .02$. The overall effect was non-significant, SMD = -0.05; 95\% CI [-0.42, 0.34]. There were 2 studies ($k = 6$) examining emotional arousal. One effect size was identified as an outlier, and after removing this case, heterogeneity was non-significant, $\chi^2(4) = 0.66, p = .96$; between-studies: $I^2 = 0\%, \tau^2 < 0.01$; within-studies: $I^2 = 0\%, \tau^2 < 0.01$, and the overall effect of SD on arousal was significant, SMD = -0.53, SE = 0.17, 95\% CI [-0.86, -0.19], indicating that SD blunted arousal in response to emotional stimuli.

There were 22 studies ($k = 30$) examining anxiety symptoms. Heterogeneity, $\chi^2(29) = 106.64, p < .001$; between-studies: $I^2 = 78\%, \tau^2 = 0.17$; within-studies: $I^2 = 0\%, \tau^2 < 0.01$, and the
overall effect were significant, SMD = 0.63, 95% CI [0.43, 0.83], indicating SD increased anxiety symptoms. There were 7 studies ($k = 10$) examining depressive symptoms. Heterogeneity, $\chi^2(9) = 77.63, p < .001$; between-studies: $I^2 = 97\%, \tau^2 = 2.02$; within-studies: $I^2 = 0\%, \tau^2 < 0.01$, and the overall effect was non-significant, SMD = 0.44, 95% CI [-0.64, 1.41]. Full forest plots and individual effect sizes, model estimates, and figures and statistics for all SD analyses are available in Supplemental File 3.

**Dose-Response Effects.** Hierarchical linear models (HLMs) were estimated to examine linear and quadratic effects of the hours of wakefulness on each emotional outcome. These models were conducted for outcomes with at least ten effect sizes (positive affect, negative affect, general mood disturbance, valence, anxiety, depression). There was a linear ($b = 0.03, SE = 0.01, p = .001$) and quadratic ($b = -0.0002, SE = 0.0001, p = .002$) effect of hours of wakefulness on negative affect. With increased hours of wakefulness, the effect size between SD and negative affect gradually increased, peaking at about 60 hours awake. There was also a linear and positive effect of hours of wakefulness on general mood disturbances ($b = 0.16, SE = 0.07, p = .027$), indicating that mood disturbances increased with hours of wakefulness. After removing an outlier, there was a linear effect of hours of wakefulness on anxiety ($b = 0.68, SE = 0.26, p = .01$), and the quadratic effect was also significant ($b = -0.01, SE = 0.004, p = .01$), and indicated a gradual increase in anxiety symptoms with hours of wakefulness which peaked around 30 hours of wakefulness. There was also a significant linear ($b = 0.53, SE = 0.21, p = .01$) and quadratic ($b = 0.01, SE = 0.002, p = .01$) effect on depression, with effects peaking between 30-40 hours of wakefulness. Positive affect and valence did not show significant dose-response effects. These dose-response analyses are displayed in Figures 4a-d, and full model statistics are available in Supplemental File 3.
**Age, Sex, and Study Characteristics.** A series of HLMs examined effects of study design (between or within-subjects design), risk of bias, and sample size as well as age and sex on each emotional outcome. These models were conducted for outcomes with at least 10 effect sizes (positive affect, negative affect, general mood disturbances, valence, anxiety, depression). Lower risk of bias was associated with larger effect sizes for negative affect ($b = -0.11, SE = 0.04, p = .02$), general mood disturbances ($b = -0.28, SE = .10, p = .01$), valence ($b = -0.43, SE = 0.15, p = .004$), anxiety symptoms ($b = -0.25, SE = 0.07, p < .001$), and depression ($b = -1.06, SE = 0.18, p < .001$). In addition, study design ($b = -1.72, SE = 0.51, p < .001$) and sample size ($b = 0.06, SE = 0.01, p < .001$) moderated the effect of SD on depression. Studies with larger sample sizes and within-subjects designs had larger effect sizes. Age also moderated associations between SD and valence ($b = 0.19, SE = 0.04, p < .001$), indicating that SD had a stronger effect for samples with an older average sample age. Conversely, the opposite effect was shown for the impact of age on the associations between SD and general mood disturbances ($b = -0.19, SE = 0.05, p < .01$), with younger adults showing stronger effects. The sex distribution of each sample also moderated the association between SD and valence ($b = 0.16, SE = 0.04, p < .001$), with stronger effect sizes for samples that included a greater percentage of females. The remaining 20 effects were non-significant. Full model statistics for all moderators can be found in the Supplemental File 3.

**Sleep Restriction (SR)**

Fifty-six studies used SR, with a total of 483 effect sizes. Summary forest plots for all SR studies averaged across each outcome are available in Figure 3. There were 37 ($k = 140$) studies examining the association between SR and positive affect. Heterogeneity, $\chi^2(139) = 589.27, p < .001$; between-studies: $I^2 = 57\%, \tau^2 = 0.14$; within-studies: $I^2 = 24\%, \tau^2 = 0.06$, and the overall
effect was significant, $SMD = -0.56$, $95\% CI [-0.70, -0.42]$, indicating SR decreased positive affect. Post-hoc moderation analyses examining whether positive affect was assessed using the POMS indicated that effect size was not moderated by measurement type. There were 43 studies ($k = 223$) examining negative affect. Heterogeneity was significant, $\chi^2(222) = 595.19$, $p < .001$; between-studies: $I^2 = 46\%$, $\tau^2 = 0.47$; within-studies: $I^2 = 19\%$, $\tau^2 = 0.20$, and the overall effect was significant, $SMD = 0.20$, $95\% CI [0.10, 0.29]$. There were 12 studies ($k = 34$) examining general mood disturbances. One effect size was identified as an outlier, and after removing this case heterogeneity was significant, $\chi^2(32) = 211.04$, $p < .001$; between-studies: $I^2 = 70\%$, $\tau^2 = 0.23$; within-studies: $I^2 = 7\%$, $\tau^2 = 0.02$, and the overall effect of SR on general mood disturbances was significant, $SMD = 0.56$, $SE = 0.16$, $95\% CI [0.16, 0.87]$.

There were 12 studies ($k = 54$) examining emotional valence. Heterogeneity, $\chi^2(53) = 177.89$, $p < .003$; between-studies: $I^2 = 66\%$, $\tau^2 = 0.04$; within-studies: $I^2 = 4\%$, $\tau^2 < 0.01$, and the overall effect was significant, $SMD = 0.23$, $95\% CI [0.11, 0.36]$, indicating SR resulted in more unpleasant valence. There were 4 studies ($k = 8$) examining emotional arousal ratings. Heterogeneity was non-significant, $\chi^2(7) = 3.05$, $p = .88$; both $I^2s < 0.001\%$ and $\tau^2s < 0.001$. The overall effect was significant, $SMD = -0.20$, $95\% CI [-0.29, -0.11]$, indicating SR blunted arousal.

There were 9 studies ($k = 18$) examining anxiety symptoms. Heterogeneity was significant, $\chi^2(17) = 31.41$, $p = .02$; between-studies: $I^2 = 59\%$, $\tau^2 = 0.08$; within-studies: $I^2 = 0\%$, $\tau^2 < 0.01$. The overall effect was significant, $SMD = 0.57$, $95\% CI [0.33, 0.80]$, indicating SR increased anxiety. There were 4 studies ($k = 6$) examining depressive symptoms. One outlier was identified, and after removal of this case heterogeneity was non-significant, $\chi^2(4) = 3.84$, $p = .43$; between-studies: $I^2 = 32\%$, $\tau^2 = 0.03$; within-studies: $I^2 = 0\%$, $\tau^2 = 0.01$, and the overall effect
was significant, SMD = 0.46, SE = 0.15, p = .002, 95% CI [0.16, 0.75]. Full forest plots and individual effect sizes, model estimates, and figures and statistics for all SR analyses are available in Supplemental File 4.

**Dose-Response Effects.** HLMs estimated linear and quadratic effects of the total hours slept on each emotional outcome. Similar to the sleep deprivation models, these models were conducted for outcomes with at least ten effect sizes (positive affect, negative affect, general mood disturbance, valence, arousal, anxiety). To account for age-adjusted sleep duration recommendations, we ran these models conditionized on adult studies only. There was a significant linear \( (b = -0.01, SE = 0.01, p = .04) \) and quadratic \( (b = 0.00003, SE = 0.00002, p = .04) \) effect of sleep duration on positive affect. The effect of SR gradually increased with shorter sleep opportunities, peaking at approximately 4 hours. One outlier was identified for valence, and after removal of this outlier the linear \( (b = 0.02, SE = 0.01, p = .004) \) and quadratic \( (b = -0.00004, SE = 0.00001, p = .01) \) effect of sleep duration on effect size was significant, which also indicated that effects peaked at approximately 4 hours. Dose-response analyses for negative affect, general mood disturbance, arousal, and anxiety were non-significant. These significant dose-response analyses are displayed in Figures 5a-b, and full model statistics are available in Supplemental File 4.

**Age, Sex, and Study Characteristics.** A series of HLMs examined between-study effects of study design, risk of bias, and sample size as well as age and sex on each emotional outcome with at least 10 effect sizes (positive affect, negative affect, general mood disturbance, valence, arousal, anxiety). Between-subjects studies \( (b = 0.46, SE = 0.16, p = .01) \) and greater risk of bias were associated with larger effect sizes \( (b = 0.10, SE = 0.04, p = .01) \) for valence. Two outliers were identified for general mood disturbance, and after removal of these outliers,
age ($b = 0.03, \ SE = 0.01, p < .001$) and study design ($b = -1.10, \ SE = 0.17, p < .001$) both were significant, indicating that studies with an older sample age on average and those with within-subjects designs had larger effect sizes. Three effect sizes were identified as outliers for anxiety, and after removal of these effects age ($b = -0.31, \ SE = 0.14, p = .03$), sex ($b = -0.18, \ SE = 0.08, p = .02$), risk of bias ($b = 0.52, \ SE = 0.15, p < .001$), and sample size ($b = 0.30, \ SE = 0.11, p = .01$) were all significant. Studies with younger average samples, a greater proportion of males, higher risk of bias, and larger samples had larger effect sizes. The remaining 21 effects were non-significant. These full model statistics are available in Supplemental File 4.

**Sleep Fragmentation (SF)**

Twenty-one included studies used SF, with 256 effect sizes. Summary forest plots for all SF studies are available in Figure 3. There were 16 studies ($k = 59$) examining the association between SF and positive affect. Heterogeneity, $\chi^2(58) = 93.99, p < .001$; between-studies: $I^2 = 47\%, \tau^2 = 0.53$; within-studies: $I^2 = 0\%, \tau^2 < 0.01$, and the overall effect were significant, $\text{SMD} = -0.40, 95\% \text{CI} [-0.53, -0.26]$, indicating SF resulted in a decrease in positive affect. Post-hoc moderation analyses indicated that measurement type significantly moderated the effect of SF on positive affect ($b = 0.19, \ SE = 0.09, p = .04, 95\% \text{CI} [0.01, 0.37]$). Effect sizes for studies using the POMS ($\text{SMD} = -0.53, \ SE = 0.07, 95\% \text{CI} [-0.67, -0.38]$) were larger than studies using other measurements of positive affect ($\text{SMD} = -0.27, \ SE = 0.08, 95\% \text{CI} [-0.42, -0.12]$). There were 13 ($k = 120$) studies examining negative affect. Heterogeneity was significant, $\chi^2(119) = 171.79, p = .001$; between-studies: $I^2 = 52\%, \tau^2 = 0.04$; within-studies: $I^2 = 15\%, \tau^2 = 0.01$. The overall effect was non-significant, $\text{SMD} = 0.10, 95\% \text{CI} [-0.03, 0.23]$. There were 7 studies ($k = 59$) examining the association between SF and general mood disturbance. Heterogeneity was significant, $\chi^2(58)$
= 204.30, \( p < .001 \); between-studies: \( I^2 = 84\% \), \( \tau^2 = 0.54 \); within-studies: \( I^2 = 2\% \), \( \tau^2 = 0.01 \). The overall effect was non-significant, SMD = 0.48, 95% CI [-0.09, 1.06].

There were 3 (\( k = 15 \)) studies examining emotional valence. Heterogeneity was non-significant, \( \chi^2(14) = 27.05, \ p = .02 \); between-studies: \( I^2 = 52\% \), \( \tau^2 = 0.11 \); within-studies: \( I^2 = 0\% \), \( \tau^2 < 0.01 \). The overall effect was not significant, SMD = 0.18, 95% CI [-0.23, 0.59]. There were 2 studies (\( k = 3 \)) examining emotional arousal ratings. Heterogeneity was non-significant, \( \chi^2(2) = 2.27, \ p = .32 \); between-studies: \( I^2 = 33\% \), \( \tau^2 = 0.03 \); within-studies: \( I^2 = 0\% \), \( \tau^2 < 0.01 \). The overall effect was significant, SMD = -0.36, 95% CI [-0.69, -0.027], indicating SF decreased arousal.

There were no SF studies available to examine effects on anxiety or depressive symptoms. Full forest plots and individual effect sizes, model estimates, and figures and statistics for all analyses are available in Supplemental File 5.

**Age, Sex, and Study Characteristics.** A series of HLMs examined between-study effects of study design, risk of bias, and sample size, as well as age and sex on emotional outcomes. Models were conducted for outcomes with at least ten effect sizes (positive affect, negative affect, general mood disturbance, and valence). There was an effect of age on negative affect, with studies using younger samples having a smaller effect size (\( b = -0.01, SE = 0.003, p = .003 \)). There was also an effect of age (\( b = 0.15, SE = 0.07, p = .03 \)) and sex (\( b = -0.05, SE = 0.02, p = .04 \)) on valence, indicating that studies with an older sample and samples with more males had larger effect sizes. Two effects were identified as outliers for general mood disturbance. After removal of these outliers, age (\( b = 0.30, SE = 0.09, p < .001 \)), sex (\( b = 0.06, SE = 0.02, p < .001 \)), study design (\( b = -14.04, SE = 3.77, p < .001 \)), risk of bias (\( b = -0.82, SE = 0.32, p = .01 \)), and sample size (\( b = 0.38, SE = 0.11, p < .001 \)) were all significant. Studies with older participants, a greater proportion of females, within-subjects designs, lower risk of bias,
and larger samples all had larger effect sizes for sleep fragmentation and general mood disturbance. The remaining 9 effects were non-significant. Full model statistics for these moderation analyses are available in Supplemental File 5.

**Variation by Type of Sleep Loss**

Separate HLMs were estimated to examine whether effect sizes differed based on whether the SR/SF occurred during REM or SWS (i.e., either by selective deprivation of those stages using forced awakenings, or by restricting sleep during the latter half of the night when most REM occurs or during the first half of the night when most SWS occurs). These models were conducted for outcomes with at least ten effect sizes (positive affect, negative affect, general mood disorders, valence). There was a significant effect of type of sleep loss on the size of the effect for valence ($b = 0.23$, $SE = 0.12$, $p = .047$). Two additional analyses were conducted with the data conditionalized on SWS or REM to test whether the effect was significantly different from zero. Results indicated that unpleasantness ratings were greater after loss of REM sleep ($SMD = 0.35$, $p = .023$, 95% CI [0.05, 0.65]) when compared to loss of SWS ($SMD = 0.09$, $p = .088$, 95% CI [-0.01, 0.12]). This is displayed in Figure 6. No differences between REM and SWS disruption were observed for positive affect, negative affect, or general mood disturbance.

**Selective Reporting Bias**

To assess for possible selective reporting in the results, we conducted Egger’s regression, 3PSM tests, and created funnel plots. There was a marginal Egger’s test ($b = 0.78$, $SE = 0.44$, $p = .08$) and a significant 3PSM test ($\chi^2(1) = 13.82$, $p < .001$) for the model examining the effect of SD on positive affect. The corrected estimated for positive affect was still significant ($SMD = -0.78$, $SE = 0.03$, 95% CI [-0.84, -0.73]). No other evidence of selective reporting (based on both
Egger’s and 3PSM tests) emerged. Full results and funnel plots are provided in Supplemental Files 3-5.

**Discussion**

Motivation to understand how various forms of sleep loss affect human emotional functioning is propelled by pervasive evidence that inadequate sleep is common and imposes major individual and public health risks. Emotions are ubiquitous to the human experience and guide attentional processes, decision making, and behavior, so therefore hold direct implications for mental and physical health. To advance understanding, inform theoretical models, and guide policy recommendations, the current systematic review and meta-analysis included studies examining the impact of different types of experimental sleep loss on emotion. The strongest and most consistent effects of sleep loss were observed for decreased positive affect, followed by increased anxiety symptoms, and blunted arousal in response to emotional stimuli. Effects on other emotional outcomes were varied and less consistent. We also examined associations regarding amount and type of sleep loss, and how age and sex may influence these effects.

**Positive Affect, Negative Affect, and General Mood Disturbance**

The effect of sleep loss on positive affect was the strongest and most robust when compared to the other emotional outcomes. All forms of sleep loss, including SD, SR, and SF reduced positive affect, with the largest effect size found for SD. These findings are in line with results from naturalistic studies (i.e., those examining day-to-day fluctuations in sleep and mood) that support the role of obtaining sufficient sleep for experiencing positive affect (Konjarski et al., 2018). Reduced positive affect is consistently linked with reduced quality of life and psychiatric difficulties, including disorders marked by anhedonia, diminished anticipation, and/or blunted responsiveness to rewards (Watson & Naragon-Gainey, 2010). Reductions in positive
affect may serve as a bridge between persistent sleep loss and the development of various psychiatric disorders due to altered functioning of neural circuitry involved in expectation and valuation of rewards (Gujar et al., 2011; Libedinsky et al., 2013), depleted cognitive energy (Kahn et al., 2013), failures to consolidate positive emotional memories following sleep loss (Walker & Stickgold, 2006), and/or effects of deficient SWS on positive emotions (Finan et al., 2015, 2016). Both linear and non-linear effects also emerged for positive affect following SR, such that positive affect decreased gradually as sleep duration decreased, with the largest deficits in positive affect occurring at about 4 hours of sleep. Interestingly, the effect of SD on positive affect was not moderated by how long participants were awake, suggesting that even when wakefulness periods are only slightly extended (i.e., 18+ hours), reductions in positive affect are evident and substantial.

The effects of sleep loss on negative affect were less consistent. Only SD and SR exerted significant effects on negative affect and these effects were small. This finding diverges somewhat from findings reported by Tomaso and colleagues (2021), where negative mood was found to increase after SD and SR with a medium effect size. Of note however, this prior meta-analysis included fewer studies assessing negative affect than our study (55 versus 98 studies, respectively). This prior meta-analysis also combined assessments of mood with psychiatric symptom measures (i.e., depressive and anxiety symptoms), which may conflate other emotional experiences (e.g., anhedonia) with negative affect and explain these larger effects. Overall, the need to clearly distinguish between various emotional constructs is critical for future sleep and emotion research. A linear and quadratic effect also emerged in our study for the impact of SD on negative affect, indicating that this effect gradually increased across hours of wakefulness.
Although this effect was small overall, this suggests that more extreme periods of sleep deprivation (e.g., 2-3 days) produce more substantial increases in negative affect.

The difference in magnitude and lack of consistent effects observed for negative affect relative to positive affect may be explained by differences in the proposed evolutionary functions of positively and negatively-valenced experiences. Whereas negative affective states such as fear or disgust may aid in survival through a coordinated set of psychological, physiological, and behavioral activations that allow an individual to respond to an immediate threat when it occurs (Tooby & Cosmides, 2008), positive affective states such as joy or interest, are proposed to broaden cognition and behavior in ways that build physical, intellectual, social, and psychological resources over longer-term periods (Fredrickson, 2001). Thus, during periods of sleep loss when physiological homeostasis is disrupted, internal positive affective experiences that do not provide immediate survival benefits may diminish more rapidly, in order to re-allocate resources toward other processes that facilitate short-term survival. In contrast, negative affective states and/or reactivity to negative stimuli may remain more stable regardless of available resources due to more immediate survival benefits. This supposition is supported by our findings that the effect of sleep loss on positive affect were generally larger when studies use the POMS vigor scale, which primarily assesses high arousal positive states (e.g., items such as lively, active, and vigorous) that may require more internal resources to maintain or generate than low arousal states (e.g. calm, content), and thus may be most susceptible to sleep loss.

Diminished positive affect may also serve a more direct protective function when internal sleep pressure is high by reducing an individual’s motivation to seek out social interactions or other pleasurable experiences at a time when cognitive processes and physical capabilities are impaired. Indeed, sleep loss has been shown to result in a desire to maintain more social distance
from others (Ben Simon et al., 2020; Palmer et al., in press). It is also possible that the influence of sleep loss on negative affect may be altered based on certain situational characteristics or may only occur for some individuals. For example, one study found that sleep restriction did not significantly alter negative affect, despite large decreases in positive affect (Cox & Olantunji, 2021). However, participants reporting an evening (or “night owl”) chronotype experienced a significant increase in negative affect, with a medium effect, whereas other participants did not experience this increase. Future studies should further explore these interindividual differences and mechanisms driving these differences that may explain these inconsistent findings in the literature.

We also considered general mood disturbance as a distinct outcome since many studies quantified ‘mood’ in a way that combines positive and negative affect. Similar to negative affect, only SD and SR exerted a significant influence on general mood disturbance, with some analyses indicating larger effects for older samples. While this finding should be interpreted with caution since this moderation only emerged for SR and SF models, ongoing neurobiological maturation and emotional changes that occur through early adulthood are likely relevant (Gogtay et al., 2004). In particular, brain regions facing extended trajectories of maturation (i.e., pre-frontal cortex) are profoundly impacted by deficient sleep and play a central role in the modulation of emotional responses (Dixon et al., 2017; Yoo et al., 2007), and younger samples may be less likely to rely on these neural resources to modulate their emotional experiences given this maturational lag. Other models of age and emotion posit that older individuals experience enhanced vulnerability to experiences that result in sustained arousal due to age-related difficulties maintaining homeostasis during times of stress (Charles, 2010), and sleep loss may be another factor that increases this age-related vulnerability. In addition, there was an effect of
sex on general mood disturbances for SF studies, with larger effects observed for female participants, in line with prior research suggesting females may be more sensitive to other neurobehavioral effects of sleep loss (Ferrara et al., 2015). However, these findings should be interpreted with caution since several other effects emerged indicating that studies with a greater proportion of males saw greater effects on other emotional outcomes (e.g., anxiety).

**Emotion Reactivity**

We found significant effects for blunted emotional arousal following SD and SF, and smaller (albeit significant) effects for SR. Although we did not include studies that examined behavioral or physiological indicators of emotional arousal, comparable results have been found; for example, in facial expressions (Alfano et al., 2020; Minkel et al., 2011) and acoustic properties of speech (McGlinchey et al., 2011). A pattern of blunted arousal likely reflects impairment in top-down emotional processing following sleep loss. That is, despite hyperactivity in emotion-based neurobiological structures, diminished functional connectivity with the pre-frontal cortex is apparent following sleep loss (Ben Simon et al., 2020; Glosemeyer et al., 2020; Motomura et al., 2013; Yoo et al., 2007) which might give rise to a general ‘de-coupling’ of emotional responses, whereby subjective appraisals and behavioral displays of emotion become disconnected from internal experiences. It is also possible that these findings reflect diminished energy levels or fatigue, which may ultimately blunt physiological responses when resources are depleted following sleep loss (Kahn et al., 2013).

We also observed that SR resulted in increased unpleasantness ratings in response to emotional stimuli, but changes in these ratings were not significant following SD or SF. These findings mirror the inconsistencies commonly found in the research literature on sleep and emotional reactivity (Tempesta et al., 2018) and suggest that findings related to these subjective
experiences may be more nuanced. For example, findings may depend on stimuli valence, intensity, and duration. Indeed, our analyses also indicate that the association between sleep loss and emotional reactivity is complex. Findings suggested that these effects may be stronger for older participants, and dose-response and moderation analyses suggest that unpleasantness may increase based on amount and type of sleep loss. Similar to positive affect, the effect of sleep loss on emotional reactivity peaked when sleep duration was reduced to 4 hours. In addition, the effect of sleep loss on unpleasantness ratings to emotional stimuli was largely negligible after loss of SWS, but the magnitude of the effect was stronger and significant after loss of REM sleep. This pattern is in line with prior work finding REM sleep in particular maintains emotional homeostasis and decreases reactivity to negative stimuli (Gujar et al., 2011; van der Helm & Walker, 2009). These shifts towards more unpleasant emotional reactivity also mirror other research demonstrating an overall bias towards threatening information after sleep loss (Krause et al., 2017).

**Anxiety and Depressive Symptoms**

Sleep loss produced positive, significant effects on anxiety symptoms, with some evidence that this effect is stronger for younger individuals. The anxiogenic effect of sleep loss is among the most consistently reported in the literature, though this relationship is likely bidirectional (Palmer & Alfano, 2017), given that the same neural alterations (e.g., excessive amygdala and insula activity) and deficient top-down emotion regulatory processes that characterize clinical levels of anxiety also typify the neurobiological consequences of sleep loss (Ball et al., 2013; Etkin & Wager, 2007; Simmons et al., 2010). Hypoactivity in the medial prefrontal cortex specifically has been directly linked with sleep-related changes in anxiety (Ben Simon et al., 2020).
The inconsistent effects for depressive symptoms were surprising given prior research showing that insomnia may contribute to the development of depressive disorders (Ford & Kamerow, 1989; Gregory et al., 2009; Perlis et al., 1997). The current study found a small effect of SR on depression, but no significant effect for SD. Sleep disturbances may influence depressive symptoms over longer periods (e.g., months), and immediate next-day effects may not be as noticeable in otherwise healthy samples. This finding should nonetheless be viewed in the context of other research suggesting that there may be beneficial effects of sleep loss on mood in a portion of clinically depressed individuals (Wu & Bunney, 1990). Yet, it is important to note that findings are mixed and mechanisms responsible for a mood enhancement effect are not well understood (Mitter et al., 2022). Although we did not include samples with known depressive disorders, it is possible that some participants had remitted or subclinical depressive symptoms and could have experienced mood enhancements following sleep loss, leading to increased variability in responses. Indeed, confidence intervals for the effects for depression, particularly for SD, were relatively large compared to other outcomes. Future research should examine possible predictors of interindividual variability in depressive symptoms after sleep loss, including baseline depressive symptoms or other indicators of risk (e.g., family or genetic influences).

Limitations of the Evidence and Future Directions

Due to the nature of sleep research, all studies were unable to fully mask participants to their condition, which may have produced expectancy effects that could not be statistically accounted for in our models. Larger effect sizes for between-subject compared to within-subject studies were also found for several emotional outcomes, in line with findings from Tomaso and colleagues (2021), but this finding is counter-intuitive given that variability in data is usually
larger for between-subject designs. We note, however, that even participants in control conditions often experience a decline in mood across the study period, likely due to extended periods in the lab (e.g., Groeger et al., 2022, Saksvik-Lehouillier et al., 2020). These declines may be heightened in within-subject designs, where participants must repeat the experience on multiple occasions, thus potentially masking differences between the control and sleep manipulation conditions. It is possible that the emotional effect of participant burden may have attenuated the effect sizes reported in the current analyses. Therefore, we encourage researchers to consider the relative participant burden (e.g., long stays in the laboratory) when designing studies with emotional outcomes, to avoid potential confounds with variables of interest.

The vast proportion of studies relied on young adult samples, which may partly explain why moderation findings related to age were somewhat inconsistent across study designs and emotional outcomes. The question of whether the emotional consequences of sleep loss varies across the life-span could not be sufficiently assessed due to age-related homogeneity across the majority of the included studies. Many studies also reported limited demographic characteristics beyond age and sex that may be important to consider in the context of sleep loss (e.g., race/ethnicity, socioeconomic status, marital/parental status) and were unable to be explored in the current study. For example, when considering the impact of sleep loss in more naturalistic settings, it is important to also understand how the cause of sleep disturbances (e.g., racial discrimination, parenting demands, inadequate sleep environment) might directly and/or indirectly influence next-day emotional experiences. Importantly, the majority of studies identified in our search include “WEIRD” samples (Henrich, Heine, & Norenzayan, 2010). Emotions experienced, expressed, and considered desirable vary dramatically across regions and cultures, in line with cultural norms and values (Chentsova-Dutton et al., 2010; De Vaus et al.,
2018; Tsai et al., 2006). As a result, understanding how sleep loss may differentially impact emotion across cultures is an important future direction, and it is possible that our meta-analytic results may not generalize to non-WEIRD populations.

The limited number of studies available for some analyses also provides insight into notable gaps in the literature. In particular, only one SF study included measures of depressive symptoms and no SF studies measured anxiety symptoms, which limited our ability to include these outcomes in the present set of analyses. There were also only a small number of studies for some analyses (e.g., SF and valence/arousal). Furthermore, we were unable to distinguish between emotional reactivity in response to positive versus negative emotional stimuli due to the small number of studies available using these methods. Therefore, results related to arousal and valence should be interpreted with caution.

In addition, while pre-registration of our study included planned analysis of physiological or behavioral indicators of emotional reactivity, search and coding procedures revealed inadequate data was available for meta-analytic techniques. Although several studies included some physiological variables such as cortisol, skin conductance, pupil responses, heart rate, and/or blood pressure, each physiological outcome was generally examined only in one or two studies. When similar indicators were used in multiple studies, the heterogeneity in methods made it difficult to examine these effects together in any meaningful way (e.g., social and non-social emotional elicitation paradigms, different sleep manipulations, pediatric vs. adult samples, measurement timing was varied). Future research should build upon these preliminary results to enhance our understanding of how sleep loss impacts objective indices of emotional reactivity in response to various stimuli, across age groups, and in response to different experimental sleep loss paradigms. Some studies also restricted sleep across multiple nights, but these studies were
too few to examine the meta-analytic effects of multiple nights of sleep loss. Given the pervasiveness of chronic sleep loss across the world (Stranges et al., 2012), future research examining multiple nights of short sleep duration seems paramount.

This meta-analysis is the first to synthesize the impact of REM sleep and SWS on emotion-based outcomes, and findings point to differential effects of REM sleep specifically on emotional reactivity. However, it is important to note that only a small number of studies in the current meta-analysis restricted REM sleep, some of which had conflicting findings. We included studies that directly targeted these stages through awakenings confirmed by polysomnography, but also studies using split-night paradigms, which likely resulted in less precise restriction of specific sleep stages. As a result, findings related to SWS and REM should be interpreted cautiously. Future research should continue to pursue questions related to sleep architecture and emotional experiences to better understand these associations and their underlying neurobiological mechanisms. In addition, the majority of experimental sleep paradigms directly manipulate sleep duration. However, sleep health is composed of many different facets, including subjective quality, daytime alertness, sleep efficiency, sleep timing, and regularity (Buysse, 2014), all of which may influence daytime emotional experiences. For example, a recent systematic review found that in addition to sleep duration, sleep quality, regularity, and sleep onset latency were all associated with both short-term changes in positive and negative affect (Konjarski et al., 2018). Included studies may have indirectly impacted other components of sleep beyond sleep duration (e.g., sleep quality, efficiency), which may have further influenced participants’ emotional state.

Importantly, influence analyses indicated that overall, the current findings were relatively robust to the influence of single studies, and when necessary, outliers were corrected for in our
models. In addition, evidence of publication bias did emerge for the effect of SD on positive affect, although the corrected effect was still significant. Although this review employed expansive search criteria and authors were contacted for unreported or unpublished data, not all data were successfully obtained during this process. Very few studies in the current meta-analyses were pre-registered or openly engaged in other open science practices (e.g., openly available materials or code). These effect sizes should be interpreted in light of this caveat.

Conclusions

This study represents the largest, most comprehensive meta-analysis to date quantifying the impact of sleep loss on domains of human emotional functioning. In light of the proportion of individuals around the world who routinely obtain inadequate sleep, findings have direct relevance for science, practice, and policy by highlighting the destabilizing effects of sleep loss on our daily emotional experiences. Results reveal sleep loss alters multiple domains of emotion, with the strongest effects observed for positive affect, anxiety, and emotional arousal. These findings should inform regulations and policies that directly impact sleep-wake patterns including maximum work periods and schedules (e.g., for surgeons, truck drivers), school start times, military/combat operations, daylight savings time, and investment in education and healthcare that prioritizes sleep.
References

*An asterisk denotes a study that was included in the meta-analysis.


https://doi.apa.org/doiLanding?doi=10.1037%2Ft00742-000


*Blagrove, M., & Akehurst, L. (2001). Personality and the modulation of effects of sleep loss on


on an evening submaximal weightlifting performance; are 1 h powernaps useful on the
https://doi.org/10.1080/07420528.2018.1552702

Burrow, A. L., Sumner, R., & Ong, A. D. (2014). Perceived Change in Life Satisfaction and
Daily Negative Affect: The Moderating Role of Purpose in Life. *Journal of Happiness


effects of acute sleep deprivation on affect and avoidance. *Journal of Behavior Therapy


neural responses to 24-hour sleep deprivation in humans: Sex differences. *American
https://doi.org/10.1152/ajpheart.01132.2011


Chentsova-Dutton, Y. E., Tsai, J. L., & Gotlib, I. H. (2010). Further evidence for the cultural
norm hypothesis: Positive emotion in depressed and control European American and
Asian American women. *Cultural Diversity and Ethnic Minority Psychology, 16*(2), 284-
295.


*Davidson, F., Rigney, G., Brine, S., Speth, T., Miller, L., Rusak, B., Chambers, C., Rajda, M.,


https://doi.org/10.1176/appi.ajp.2007.07030504

https://doi.org/10.1093/sleep/28.12.1561

https://doi.org/10.1093/scan/nsx148

https://doi.org/10.1371/journal.pone.0120029


https://doi.org/10.1093/sleep/zsw017

https://doi.org/10.5665/sleep.5154


https://doi.org/10.2147/NDT.S53633

*Kaida, K., Niki, K., & Born, J. (2015). Role of sleep for encoding of emotional memory. *Neurobiology of Learning and Memory, 121*(8), 72–79.
https://doi.org/10.1016/j.nlm.2015.04.002


https://doi.org/10.1037/a0034501

https://doi.org/10.1371/journal.pone.0116906


https://doi.org/10.1037/xge0000522


https://doi.org/10.1046/j.1525-1497.2001.016009606.x


https://doi.org/10.3389/fnbeh.2022.885302


*Psychoneuroendocrinology, 49, 34–46. https://doi.org/10.1016/j.psyneuen.2014.06.002


*Lorenzetti, M. S. (2022). The effects of sleep restriction on biological, psychological, and neurocognitive measures of health [ProQuest Information & Learning].
https://app.readcube.com/library/d39152e9-0e91-471f-bdfe-34403f94572d/item/ff5f2b7e-9133-4295-bbdf-b74353bf5032


for reporting systematic reviews. *BMJ (Clinical Research Ed.),* 372, n71. https://doi.org/10.1136/bmj.n71


https://doi.org/10.1002/jrsm.1332

https://doi.org/10.1111/jsr.12484


https://dspace.sunyconnect.suny.edu/bitstream/handle/1951/59838/Riela_grad.sunysb_0771E_10615.pdf?sequence=1&isAllowed=y

https://doi.org/10.1093/sleep/18.1.30

https://doi.org/10.1037/met0000300


adolescents. *Sleep Medicine, 16*(8), 987–993. https://doi.org/10.1016/j.sleep.2015.03.007


https://doi.org/10.1037/0022-006X.51.1.14


https://doi.org/10.3389/fnbeh.2022.880641


Acute sleep deprivation disrupts emotion, cognition, inflammation, and cortisol in young
https://doi.org/10.3389/fnbeh.2022.945661

restriction on mood, emotion, and emotion regulation: Three meta-analyses in one. *Sleep, 44*(6), zsaa289. https://doi.org/10.1093/sleep/zsaa289

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their

Triantafillou, S., Saeb, S., Lattie, E. G., Mohr, D. C., & Kording, K. P. (2019). Relationship
Between Sleep Quality and Mood: Ecological Momentary Assessment Study. *JMIR Mental Health, 6*(3), e12613. https://doi.org/10.2196/12613


Learning and Memory. *Frontiers in Psychology, 8.*

532–537.

*Vargas, I., & Lopez-Duran, N. (2017). Investigating the effect of acute sltsaieep deprivation on
hypothalamic-pituitary-adrenal-axis response to a psychosocial stressor.

http://dx.doi.org/10.1007/BF02294384

http://dx.doi.org/10.1037/1082-989X.10.4.428


https://doi.org/10.1093/jpepsy/jst033

https://doi.org/10.1126/sciadv.1501705


https://doi.org/10.1037/a0016570


Table 1
Descriptive summary of sample, study, and moderator characteristics

<table>
<thead>
<tr>
<th>Category</th>
<th>Studies</th>
<th>Effects</th>
<th>M</th>
<th>SD</th>
<th>Median</th>
<th>Mode</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Sample Female</td>
<td>51.02</td>
<td>25.25</td>
<td>50</td>
<td>50</td>
<td>0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Sample Age</td>
<td>24.16</td>
<td>8.54</td>
<td>22.35</td>
<td>33.25</td>
<td>8-67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18 years</td>
<td>16</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-29.9 years</td>
<td>117</td>
<td>900</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-59.9 years</td>
<td>11</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60+ years</td>
<td>4</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Reported</td>
<td>8</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample majority white</td>
<td>25</td>
<td>173</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample majority non-white</td>
<td>4</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>125</td>
<td>1109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>46</td>
<td>252</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia &amp; New Zealand</td>
<td>6</td>
<td>92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>75</td>
<td>787</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>25</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td>1</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>30.19</td>
<td>26.25</td>
<td>20</td>
<td>8</td>
<td>4-193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>16</td>
<td>285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11-25</td>
<td>59</td>
<td>466</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-50</td>
<td>42</td>
<td>375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-75</td>
<td>21</td>
<td>131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76-100</td>
<td>8</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101+</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1966-1979</td>
<td>4</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980-1989</td>
<td>5</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990-1999</td>
<td>14</td>
<td>101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000-2009</td>
<td>27</td>
<td>395</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-2019</td>
<td>79</td>
<td>484</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020-later</td>
<td>24</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No year (unpublished dataset)</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer-reviewed</td>
<td>146</td>
<td>1254</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Peer-reviewed</td>
<td>8</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Data Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-subjects</td>
<td>47</td>
<td>385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within-subjects</td>
<td>107</td>
<td>953</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep Manipulation and Amount of Sleep Loss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deprivation Studies</td>
<td>84</td>
<td>599</td>
<td>30.10</td>
<td>12.16</td>
<td>27</td>
<td>24</td>
<td>18-110</td>
</tr>
<tr>
<td>Hours of Wakefulness across Effect Sizes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restriction Studies</td>
<td>55</td>
<td>483</td>
<td>4.18</td>
<td>1.31</td>
<td>4.00</td>
<td>4.21</td>
<td>1.95-8.37</td>
</tr>
<tr>
<td>Sleep Duration across Effect Sizes (hours)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fragmentation Studies</td>
<td>21</td>
<td>256</td>
<td>6.21</td>
<td>1.44</td>
<td>6.72</td>
<td>6.93</td>
<td>3.82-8.5</td>
</tr>
<tr>
<td>Sleep Duration across Effect Sizes (hours)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Sleep Loss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWS</td>
<td>34</td>
<td>303</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REM</td>
<td>12</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both/Unclear</td>
<td>36</td>
<td>365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent Variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Affect</td>
<td>75</td>
<td>388</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POMS</td>
<td>47</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-POMS</td>
<td>59</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative Affect</td>
<td>81</td>
<td>623</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Mood Disturbance</td>
<td>26</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td>17</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arousal</td>
<td>6</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive Symptoms</td>
<td>10</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety Symptoms</td>
<td>29</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. POMS = Profile of Mood States, SWS = slow wave sleep, REM = rapid eye movement sleep. Some categories do not add up to the total included number of studies or number of effect sizes due to the inclusion of multiple samples, multiple dependent variables or sleep manipulations, and published and unpublished effects within the same study.
Figure 1

Flow Diagram Outlining Number of Records Identified, Included and Excluded, and Reasons for Exclusion

Note. This figure was adapted from Page et al., 2021.
Figure 2

Risk of Bias Categories

Note. Each bar represents the percentage of studies with (from left to right) high (red), low (orange), or unclear (grey) risk of bias for each category.
Figure 3

Summary Forest Plots Representing Each Meta-Analysis

Note. Each figure shows the average effect size for each sleep loss paradigm on each emotion outcome. The parentheses next to each outcome on each y-axis indicate the number of effect sizes for that criterion. An asterisk means that an outlier was removed from the model. The dotted red lines represent the conventional Cohen’s $d$ effect size interpretation, with the brightness of the line corresponding to the strength of the effect. The error bars represent the 95% CI. Full forest plots including effect sizes for each study are presented in Supplemental Files 3-5.
Figures 4a-d

Linear and Non-Linear Effects of Hours of Wakefulness in Total Sleep Deprivation Paradigms

Note. The size of each bubble corresponds with the study sample size. The teal color represents a linear function; the purple color represents a quadratic function. The shaded area represents the 95% confidence interval. The anxiety model excludes an outlier.
Figures 5a-b

Linear and Non-Linear Effects of Sleep Duration in Sleep Restriction Paradigms

Note. The size of each bubble corresponds with the study sample size. The teal color represents a linear function; the purple color represents a quadratic function. The shaded area represents the 95% confidence interval. The valence model excludes an outlier.
Figure 6

Effect of REM Sleep Loss (versus SWS Loss) on Negative Valence in Response to Emotional Stimuli

Note: The size of each dot corresponds to the effect size’s sample size. Error bar represents the 95% CI.