Consciousness and Cognition 25 (2014) 51–66

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

I know how you felt last night, or do I? Self- and external ratings of emotions in REM sleep dreams

Pilleriin Sikka [a,b,*], Katja Valli [a,b], Tiina Virta [a], Antti Revonsuo [a,b]

[a] Centre for Cognitive Neuroscience, Department of Psychology, University of Turku, Finland
[b] School of Bioscience, University of Skövde, Sweden

ARTICLE INFO

Article history:
Received 14 July 2013

Keywords:
REM sleep
Dreaming
Emotions
Content analysis
Modified Differential Emotions Scale

ABSTRACT

We investigated whether inconsistencies in previous studies regarding emotional experiences in dreams derive from whether dream emotions are self-rated or externally evaluated. Seventeen subjects were monitored with polysomnography in the sleep laboratory and awakened from every rapid eye movement (REM) sleep stage 5 min after the onset of the stage. Upon awakening, participants gave an oral dream report and rated their dream emotions using the modified Differential Emotions Scale, whereas external judges rated the participants' emotions expressed in the dream reports, using the same scale. The two approaches produced diverging results. Self-ratings, as compared to external ratings, resulted in greater estimates of (a) emotional dreams; (b) positively valenced dreams; (c) positive and negative emotions per dream; and (d) various discrete emotions represented in dreams. The results suggest that this is mostly due to the underrepresentation of positive emotions in dream reports. Possible reasons for this discrepancy are discussed.

1. Introduction

A dream is a sequence of rich subjective experiences occurring during sleep in the absence of external physical stimulation or behavioural activity. As such, dreaming reflects pure phenomenality and the importance of dream research in the scientific study of consciousness is now widely acknowledged (e.g., Hobson, 2009; Nir & Tononi, 2010; Revonsuo, 2006; Windt & Noreika, 2011).

Despite a considerable amount of research on the various phenomenal contents of dreams, several unanswered questions remain. One such question is the affective content of dream experiences. It is generally agreed that emotions are central in dreams, at least with regard to rapid eye movement (REM) sleep dreams (e.g., Hobson, Pace-Schott, & Stickgold, 2000; Nir & Tononi, 2010). However, results conflict as to the frequency with which emotions occur in dream reports (i.e., whether dreams are mostly emotional or non-emotional), the prevailing emotional valence of dreams (i.e., whether dreams are mostly negative, mostly positive or have a more balanced emotional tone) and the dominance of specific emotion categories in dreams (i.e., whether certain types of emotions, such as fear, are more prevalent than others in dreams). These inconsistencies occur due to several methodological differences in how data are acquired and analysed.

One important difference is how emotions are measured. As there is currently no reliable way to objectively measure affective experiences occurring in dreams without probing the subjects' conscious experience, studies rely on subjective

* Corresponding author at: Centre for Cognitive Neuroscience, Department of Psychology, University of Turku, Finland. Fax: +358 2 333 8770.
E-mail address: pilsik@utu.fi (P. Sikka).

http://dx.doi.org/10.1016/j.concog.2014.01.011
1053-8100/

reports. Traditionally, dream reports have been analysed by blind judges who rate the occurrence of explicitly mentioned emotions. However, it can be argued that such third-person ratings reflect only the use of affective language and not the actual phenomenal contents of the experience. Alternatively, subjects have been asked to rate their emotions either upon awakening from a dream or as expressed in their dream report using various emotion rating scales. Because such first-person ratings are performed in a different state of consciousness, it is possible that these tend to reflect the emotions experienced in the current waking consciousness rather than those experienced in the preceding dream consciousness.

The advantages and disadvantages of using self- vs external ratings for measuring dream emotions remain to be determined. It is important for both dream research and consciousness research to acquire a deeper understanding of the strengths and weaknesses of these different methodological approaches in the measurement of the subjective contents of consciousness. The conflicting results suggest that these methods involve different biases in the sampling and rating of the subjective contents of consciousness in general, and of affective content in particular. Moreover, a better understanding of these strengths and weaknesses will be useful for future studies that, for instance, investigate the neural correlates of affective contents of consciousness.

1.1. Self- vs external ratings of dream emotions

External ratings of dream reports have most often been conducted using the Emotions scale of the Hall and Van de Castle (1966) content analysis system. Following a set of clearly defined rules, external judges code only explicitly expressed emotions in dream reports. Using this scale, typically less than half of the dream reports have been rated as emotional, (e.g., 30–35% in Snyder, 1970), less than one emotion has been detected per dream report (e.g., 0.35 in Hall & Van de Castle, 1966), and negative dreams as well as negative emotions in dreams have been shown to prevail (e.g., Hall & Van de Castle, 1966; Kramer, Winget, & Whitman, 1971; Snyder, 1970).

Self-ratings of dreams have been carried out with a number of different scales with which individuals are typically asked to rate either the occurrence and/or intensity of specific emotions or of the emotional tone of the dream in general. Studies using such self-rating scales have yielded a different pattern of results. With the latter, dreamers have rated the majority of their dreams as emotional (ranging from 70% as in Foulkes, Sullivan, Kerr, & Brown, 1988; up to 98.4% as in St-Onge, Lortie-Lussier, Mercier, Grenier, & De Koninck, 2005) and have reported more than one emotion per dream (e.g., 3.7 in Merritt, Stickgold, Pace-Schott, Williams, & Hobson, 1994; or 7.8 in Nielsen, Deslauriers, & Baylor, 1991). Typically, self-ratings have demonstrated a rather balanced proportion of positively and negatively valenced dreams as well as positive and negative emotions in dreams (e.g., Blagrove, Framer, & Williams, 2004; Fosse, Stickgold, & Hobson, 2001), although in some studies negative emotions (Merritt et al., 1994) and in others positive dreams and positive emotions (St-Onge et al., 2005) have been found to dominate.

Studies directly comparing self- and external ratings are scarce. Schredl and Doll (1998) used three different approaches to measure emotions in home dream reports: content analysis by external raters using the Hall and Van de Castle (1966) coding system, and external ratings as well as self-ratings using two 4-point scales, one for measuring the overall intensity of positive tone and the other for measuring the intensity of negative emotional tone. With content analysis less than half (42.1%) and with self-ratings almost all (99.2%) of the dreams were evaluated as emotional, but when external raters used the same scale as the dreamers did, a much larger number of dream reports (86.5%) were rated as emotional. Concerning emotional valence, external ratings resulted in estimates of more than twice as many negative as positive dreams, irrespective of whether content analysis (26.3% and 9.0%, respectively) or the two 4-point scales (56.4% and 21.1%, respectively) were used. The proportion of predominantly negative (50.4%) and predominantly positive (36.8%) dream ratings was more balanced when dreams were rated by the subjects themselves, although negative dreams still prevailed.

However, it is unclear to what extent the differences in the results concerning dream emotionality in Schredl and Doll's (1998) study derived from actual differences in who rated the emotions or from differences in the particular scales used for measuring emotions. Whereas using the Hall and Van de Castle (1966) content analysis scale, emotions were divided into five discrete categories (anger, apprehension, happiness, sadness and confusion), the scales used for external and self-ratings measured overall emotional tone (how positive and negative the dream experience or report was in general). On the one hand, it can be argued that the increased proportion of emotional, especially positive, dreams, when rated with such general scales, might be due to measuring overall mood states, rather than specific emotions (Snyder, 1970; Strauch & Meier, 1996). On the other hand, due to the fact that in the Hall and Van de Castle (1966) content analysis scale the number of negative categories exceeds that of positive ones, the raters may be biased to detect negative emotions in dream reports (Mealey, 2000). It remains to be determined whether the higher proportion of emotional, and especially of positively valenced, dream ratings would also occur when a balanced number of specific positive and negative emotions is measured. Moreover, it is unclear whether and to what extent self- and external ratings differ in the frequency of discrete emotions.

Furthermore, the Schredl and Doll (1998) study was based on dream diaries collected at home. Previous studies have demonstrated that the experimental environment (home vs laboratory) can influence the emotional content of dreams with home dreams typically being rated as more emotional and negative than laboratory dreams (e.g., Domhoff & Schneider, 1999; St-Onge et al., 2005; Weisz & Foulkes, 1970). These differences may be due to the biased selection of dream reports in the home setting. Whereas home dreams are typically reported upon spontaneous morning awakenings, laboratory dreams are obtained upon controlled awakenings from pre-specified sleep stage(s) sampled throughout the night. Therefore, as pointed out by Domhoff (2005), due to the recency effect (at home people typically remember only the last dream before

spontaneous awakening) and selection bias (people tend to recall more emotionally salient, especially negative, dreams) home dreams might not reflect a representative sample of emotional dream content. Moreover, most, albeit not all, spontaneous awakenings occur after REM sleep, this being more likely during late night early morning sleep (Schulz & Bes, 1998). It has been shown that such late REM sleep dreams, as opposed to early night REM sleep dreams, are more vivid and emotional (e.g., Agargun & Cartwright, 2003; Verdone, 1965), although contrary data also exist (e.g., Fosse et al., 2001). It is debatable to what extent this early vs late REM sleep dream difference reflects a true difference in emotionality and to what extent it is accounted for by the longer length of the late REM dream reports (Casagrande, Violani, Lucidi, Buttinelli, & Bertini, 1996). Dreams containing more complex features, such as emotion, may indeed require more words to describe them but the shorter length of a dream report does not necessarily mean that the dream had less such features but may simply reflect the inability of the dreamer to recall and describe them (Hobson et al., 2000). Nevertheless, environmental setting seems to influence the emotional content of dream reports, and therefore, it is not clear whether differences between self- and external ratings as reported by Schredl and Doll (1998) would also occur with systematic awakenings in the laboratory setting.

1.2. Aim and hypotheses

The present study directly addresses the issues raised above by investigating emotions in REM sleep dreams collected in a controlled laboratory environment and by contrasting self-ratings with external ratings of the same set of dreams using identical scales with an equal number of discrete positive and negative emotions. The study aims to demonstrate whether and to what extent the self-ratings of dream emotions converge with external ratings.

Based on previous research, it is hypothesized that when the same dreams are rated with the same scale: (1) self-ratings reflect a larger number of emotional dreams than external ratings; (2) self-ratings reflect a balanced ratio of positive and negative dreams, whereas external ratings indicate a larger number of negative than positive dreams; (3) self-ratings reflect a larger number of emotions per dream than external ratings; (4) self-ratings and external ratings differ more in the number of positive emotions than negative emotions per dream.

2. Method

2.1. Participants

Participants were recruited via an advertisement sent to student mailing lists at the University of Turku, Finland. One hundred and fifty-nine volunteers responded to an online background questionnaire. From these 21 were selected, after multiple screening phases, to sleep in the sleep laboratory, and 17 of them produced data suitable for analyses. The final sample consisted of seven men and ten women. The average age was 25.76 (SD = 4.93; range = 19–39). The screening phases, inclusion and exclusion criteria as well as drop-out rates in each screening phase are depicted in Fig. 1.

First, the 159 volunteers filled in an electronic questionnaire, with items measuring handedness, native language, sleep habits and the quality of sleep (using a Finnish translation of the Pittsburgh Sleep Quality Index, PSQI; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). Healthy right-handed participants with Finnish as their native language demonstrating good sleep quality (PSQI $\leqslant$ 5) were selected. Also, spontaneous dream recall frequency was measured to obtain participants with high or low spontaneous dream recall rates. Three additional items measured the individuals' attitudes towards their dreams.

Participants who met the above-mentioned criteria (n = 53) were asked to keep dream diaries at home for seven consecutive days to gain more objective information of their dream recall rates. Paying attention to dreams improves recall (Cohen, 1969), or at least for those having low or medium frequency of dream recall (Schredl, 2002), and writing a dream diary encourages people to report dreams (Revonsuo & Salmivalli, 1995). The instructions were derived from previous studies (e.g. Revonsuo & Salmivalli, 1995) where they have proven to work well.

Based on the number of dream reports recorded in home dream diaries per week as well as the word count of those dream reports as calculated according to the Total Recall Frequency (total number of dream-related words minus utterances, fillers, repetitions, corrections, commentaries; Antrobus, 1983), two recall groups (high and low recallers) were formed. These groups did not differ in their waking memory functions as measured with the Visual Puzzles from the Wechsler Adult Intelligence Scale, 3rd edition (Wechsler, 2005), the Logical Memory I and II from Wechsler Memory Scale, 3rd edition (Wechsler, 2007) and free oral recall of a 3-min video (Heinilä, 2008). From the 53 participants, 22 completed both the dream diary and the memory tests. With one participant excluded, the number of dreams reported in home dream diaries became the only statistically significant difference between the low (M = 4.82, SD = 1.25, range = 3–6) and high (M = 10.30, SD = 3.49, range = 7–17) recallers, Mann–Whitney U = 0.00, Z = −3.91, p < .001, r = .85. Thus, 21 subjects were invited to sleep in the laboratory.

The study protocol was approved by the Ethical Board of the University of Turku. Subjects participating in the laboratory nights signed an informed consent form prior to their first night and were informed that they could discontinue the experiment at any time. A monetary compensation of 100 Euros was provided for subjects when they had completed the whole procedure.

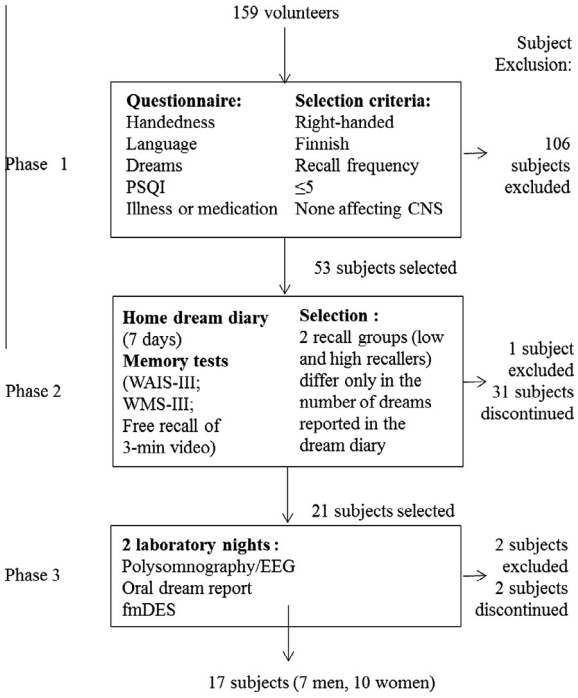

Fig. 1. Participant flow chart for the study. PSQI = Pittsburgh Sleep Quality Index; WAIS-III = Wechsler Adult Intelligence Scale; WMS-III = Wechsler Memory Scale; fmDES – Finnish translation of the modified Differential Emotions Scale.

2.2. Laboratory procedure

Of the 21 participants invited to continue with the study, 19 accepted and spent two non-consecutive nights in the sleep laboratory. The time interval between the two nights was, in most cases, a week. A recovery period between sessions ensured that there was no REM sleep deprivation that could have affected the results.

In the laboratory sleep was monitored by polysomnography with 24 electrodes (Fp1/2, AF3/4, AF7/8, F7/8, F3/4, Fz, T7/8, C3/4, Cz, P7/8, P3/4, Pz, O1/2, Oz) positioned according to the standard 10 /10 system. Four electrodes were used to monitor vertical and horizontal eye movements, and two to assess chin EMG. All electrodes, except bipolar EOG and EMG electrodes, were referenced to the right ear mastoid. The EEG signal was amplified (SynAmps Model 5083), recorded and manually sleep stage scored using NeuroScan equipment and software.

After the electrodes were attached, subjects were allowed to fall asleep. A tone signal was used to awaken the subjects every time REM sleep had lasted continuously for 5 min and was in a phasic stage determined visually by a judge according to scoring manuals (Iber, Ancoli-Israel, Chesson, & Quan, 2007; Rechtschaffen & Kales, 1968). As the phasic stage of REM sleep seems to lead to a higher dream recall rate than the tonic stage (Pivik, 1991), awakenings were always performed from the phasic stage.

First, after the awakening, subjects reported the last image they recalled from the dream, and then the whole dream in as much detail as possible. Oral dream reports were recorded and transcribed for external ratings and the total dream-related word count calculated according to the Total Recall Frequency (Antrobus, 1983). After the oral dream report, subjects rated their dream emotions using the self-rating scale and were then allowed to continue their sleep. The scale was filled in electronically using a mouse and a computer screen above the bed. In case the subjects reported "no dreams" the laboratory assistant asked whether they had not had a dream or they felt like they had had a dream but could not recall any specific content. The latter response was labelled as a 'white dream' (Strauch & Meier, 1996). In these two cases no scale was filled in. During the night, the room was entered only when the subject needed to leave the bed, or electrodes had become detached and needed replacing. The study was finished according to an agreement with the subject sometime between 5:30 and 8:30 a.m.

Two subjects were excluded from the analyses. One lacked the results of the emotion measurement scale because of an Internet connection failure, and the other had a very poor quality EEG. Also, three dreams of different subjects were omitted because of an experimenter error. Thus, 17 subjects and all but three of their dreams were included in the analyses.

2.3. Measures

2.3.1. Self-ratings

A Finnish translation of the modified Differential Emotions Scale (fmDES; Fredrickson, 2013) was used to measure positive and negative emotions occurring in dreams experienced by the dream self. The fmDES consists of 10 items for positive

emotions and 10 items for negative emotions, each described by three adjectives. Subjects rated the emotions they had had in the dream on a Likert scale from 0 ("I did not experience any of these feelings at all") to 4 ("I experienced one or more of these feelings extremely much").

To have data comparable to those collected with external ratings (see Section 2.3.2), the frequency of occurrence, rather than the intensity, of the emotions was analysed. Thus, the results obtained with the fmDES were used in a dichotomous manner with a cut point between 0 and 1, that is, between emotion experienced "not at all", and emotion experienced "a little bit".

In addition to evaluating discrete emotions, aggregate subscales were formed separately for positive and negative emotions. The Positive Emotions (PE) subscale consisted of nine positive emotions (all but awe/wonder/amazement), with coefficient $\alpha = .85$. The Negative Emotions (NE) subscale consisted of nine negative items (all but embarrassed/self-conscious/blushing) with coefficient $\alpha = .84$. These two items were omitted due to poor item-to-total correlations, as in Fredrickson, Tugade, Waugh, and Larkin (2003). Although these two items were not considered as part of their respective subscales, they were included in the analysis of the overall emotionality of dreams. Hence, the PE and NE reflect the sum of the occurrence of each of the nine positive and negative emotions, respectively. These subscales were used to calculate the overall emotional valence of dreams as well as the number of positive and negative emotions per dream.

An additional set of exploratory analyses based on the intensity ratings (actual scores from 0 to 4) of the emotions was conducted (see Section 3.5.3 for a more detailed description).

2.3.2. External ratings

The same fmDES scale described above was used to rate emotions in the transcribed dream reports by two judges. First, the judges, working blindly and independently, identified all cases of emotions when (1) an emotion was explicitly expressed (e.g., "I noticed that there were two shockingly big dogs and *I was afraid* of what was going to happen"); (2) an emotion was present but its target was unclear and could not be attributed to any particular person besides the dream self (e.g., "Three of the puppies jumped on me which was *terribly funny* as they began biting each other's tails"); (3) the dream character exhibited behaviour that clearly depicted an emotional state, and the emotional state was explicitly inferable from the behaviour, i.e., only one prominent emotion could be interpreted from the outside as underlying the behaviour (e.g., "He was quite a joker so we were *laughing*").

Altogether, 54 emotions were identified. The judges agreed on 44 emotions and disagreed on 10, thus identification inter-rater agreement per cent was 81.48%. In case of disagreement, the judges discussed the case, and if they reached an agreement, the emotion was included in the subsequent rating process ($n = 7$) while all ambiguous cases ($n = 3$) were excluded. Next, the judges rated the 51 (36 expressed and 15 inferred) emotions using the 20 different categories of the fmDES. Only the occurrence, and not the intensity, of each of these was rated. The inter-rater reliability for the coding of emotions, evaluated with Cohen's Kappa, was strong ($\kappa = .84$).

To be comparable to the self-ratings, the frequency of occurrence of each of the 20 emotion categories per dream was counted. Thus, it was not taken into account whether the same emotion appeared in the dream report more than once or not. As with the self-ratings, separate aggregate subscales for PE (9 items except awe/wonder/amazement) and for NE (9 items except embarrassed/self-conscious/blushing) were created to calculate the overall emotional valence of dreams and the number of positive and negative emotions per dream.

3. Results

Statistical analyses were carried out using the IBM SPSS statistics software (version 20). For analyses, aggregate scores of every variable being studied across the dreams for each individual were calculated. The normality assumption was tested with the Shapiro–Wilk test (Shapiro & Wilk, 1965). As the majority of the variables were not normally distributed, most comparisons were conducted with nonparametric tests (Mann–Whitney *U* test or Wilcoxon signed rank test). Only when both variables in the comparison were normally distributed, were parametric tests used (independent samples *t*-test or paired-samples *t*-test). Correlation analyses were carried out using the Spearman's rank correlation coefficient ($r_{s)}$. All statistical tests were two-tailed. Effect sizes were calculated using Cohen's *d* for *t*-tests, and Pearson's correlation (*r*) for nonparametric tests.

To account for the possible time of the night effect, dreams from early night REM sleep were compared with those from late night REM sleep. The cut point for defining early and late REM was determined as in Casagrande et al. (1996): The first two REM periods were grouped as early REM and from the third REM period onward as late REM.

3.1. Preliminary analyses: distribution of awakenings and dreams

Before testing specific hypotheses, the overall number of awakenings, reported dreams, and dreams from early and late REM stages were examined to see whether they differed between the first and second laboratory night, or between men and women, or high and low recallers. Due to experimenter error in observing arousal states, or due to spontaneous awakenings during REM sleep, the length of continuous REM stage did not always reach the intended 5 min; the average duration of REM sleep per subject before awakening occurred was 4.65 min ($SD = 3.21$).

The total number of awakenings was 126 (Night 1, $n = 60$; Night 2, $n = 66$), and dreams were reported in 115 awakenings (see Table 1). Eleven awakenings (all from early REM sleep) did not result in dream recall. The mean dream recall rate was

Table 1
Descriptive statistics for the distribution of awakenings, dreams and dream recall percentage.

	All		Night				Gender				Recall group			
			1st		2nd		Men		Women		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Awakenings	7.41	3.02	3.53	1.74	3.88	1.58	7.29	3.64	7.50	2.72	7.70	2.67	7.00	3.65
Dreams	6.76	3.05	3.24	1.79	3.53	1.50	6.86	3.53	6.70	2.87	6.80	2.78	6.71	3.64
Early REM dreams	2.29	1.21	1.12	0.70	1.18	0.81	2.29	1.38	2.30	1.16	2.40	1.17	2.14	1.35
Late REM dreams	4.47	2.50	2.12	1.50	2.35	1.27	4.57	2.44	4.40	2.67	4.40	2.46	4.57	2.76
Recall%	90.75	10.05	87.89	25.69	91.76	11.95	94.39	7.43	88.20	11.19	87.05	10.50	96.03	6.96
Early REM recall%	76.15	27.80	77.78	31.29	72.62	37.33	86.11	15.52	70.17	32.38	70.17	30.16	73.81	38.32
Late REM recall%	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00

90.75%, ranging from 70.17% to 100%. There were no significant differences between the two laboratory nights in the mean number of awakenings, $t(16) = -1.07$, $p > .05$, $d = -0.21$; dreams, $t(16) = -0.96$, $p > .05$, $d = -0.18$; early REM dreams, Wilcoxon $Z = -0.28$, $p > .05$, $r = -.05$; late REM dreams, $t(16) = -0.81$, $p > .05$, $d = -0.17$; or in dream recall percentage in general, Wilcoxon $Z = -0.06$, $p > .05$, $r = -.01$; early REM dream recall percentage, Wilcoxon $Z = -0.09$, $p > .05$, $r = -.02$; or late REM dream recall percentage, Wilcoxon $Z = 0.00$, $p > .05$, $r = .00$.

Men and women did not differ in the number of awakenings, $t(15) = -0.14$, $p > .05$, $d = -0.07$; dreams, $t(15) = 0.10$, $p > .05$, $d = 0.05$; early REM dreams, $t(15) = -0.02$, $p > .05$, $d = -0.01$; late REM dreams, $t(15) = 0.14$, $p > .05$, $d = 0.07$; or in dream recall percentage in general, Mann–Whitney $U = 23.00$, $Z = -1.24$, $p > .05$, $r = -.21$; early REM dream recall percentage, Mann–Whitney $U = 20.50$, $Z = -1.08$, $p > .05$, $r = -.19$; or late REM dream recall percentage, Mann–Whitney $U = 35.00$, $Z = 0.00$, $p > .05$, $r = .00$.

Also, there were no significant differences between low and high recallers in the mean number of awakenings, $t(15) = 0.46$, $p > .05$, $d = 0.22$; dreams, $t(15) = 0.06$, $p > .05$, $d = 0.03$; early REM dreams, $t(15) = 0.42$, $p > .05$, $d = 0.20$; late REM dreams, $t(15) = -0.14$, $p > .05$, $d = -0.07$. Importantly, low recallers proved to be just as good in recalling their dreams in the laboratory setting as high recallers, Mann–Whitney $U = 17.50$, $Z = -1.81$, $p > .05$, $r = -.31$. This applied to both early REM dream recall percentage, Mann–Whitney $U = 21.00$, $Z = -1.02$, $p > .05$, $r = -.18$, as well as to late REM dream recall percentage, Mann–Whitney $U = 35.00$, $Z = 0.00$, $p > .05$, $r = .00$. Therefore, data from the two nights, from men and women, and from high and low recallers were pooled together in subsequent analyses.

As to the time of the night, significantly more late REM ($M = 4.47$, $SD = 2.50$) than early REM ($M = 2.94$, $SD = 1.34$) awakenings were conducted, $t(16) = 2.38$, $p < .05$, $d = 0.76$. As a result, there were significantly more late than early REM dream reports, $t(16) = 3.62$, $p < .01$, $d = 1.11$. There was also a significant difference between the recall percentage of dreams from early REM awakenings and late REM awakenings, Wilcoxon $Z = -2.68$, $p < .01$, $r = -.47$ (see Table 1).

3.2. Hypothesis 1: Do self-ratings reflect a larger number of emotional dreams than external ratings?

The overall emotionality of dreams was calculated using all 20 items of the fmDES. A dream was considered emotional when at least one of the 20 emotions was reported to occur at least once (i.e., with self-ratings received any score above 0; with external ratings was detected at least once in the dream report).

With self-ratings (SR), every dream ($N = 115$) was rated to contain at least one type of emotion. Therefore, all subjects ($N = 17$) rated their dreams as emotional with an average of 6.76 ($SD = 3.05$) emotional dreams per subject across the two nights.

When measured with external ratings (ER), emotions were detected in 33 dream reports (28.7%).[1] Eleven subjects had at least one emotional dream report. Subjects whose dream reports were coded as non-emotional differed significantly from subjects whose dream reports were coded as emotional in the length of the dream reports, $t(15) = -3.67$, $p < .01$, $d = -2.09$. Subjects with non-emotional dream reports ($n = 6$) used on average 58.14 words ($SD = 22.68$) in their dream reports, whereas subjects with emotional dream reports ($n = 11$) used on average 166.79 words ($SD = 69.66$).

With ER subjects had, on average, 1.94 ($N = 17$; $SD = 1.78$) dream reports rated as emotional across the two nights. When taking into account only subjects who had at least one emotional dream report (leaving out the six subjects with non-emotional dream reports), the average number of emotional dream reports per subject was 3.00 ($N = 11$; $SD = 1.26$). There was a significant difference in the number of dream reports rated as emotional from different times of night: on average a subject had 1.65 emotional late REM dream reports ($N = 17$; $SD = 1.62$) as compared to 0.33 emotional early REM dream reports ($N = 15$; $SD = 0.62$), $Z = -2.55$, $p < .01$, $r = -.47$. This difference remained when controlling for the number of early and late REM dream reports with 0.39 emotional late REM dream reports per late REM dream ($N = 17$; $SD = 0.38$) as compared to 0.13 emotional early REM dream reports per early REM dream ($N = 15$; $SD = 0.23$), Wilcoxon $Z = -2.68$, $p < .01$, $r = -.49$. This difference can be explained by the fact that, on average, subjects used significantly more words ($N = 17$; $M = 141.28$;

[1] Only in 6 (out of 33) dreams were the external ratings based on inferred, rather than expressed, emotions. As there were no differences with regards to whether only dreams with expressed emotions or dreams with expressed and inferred emotions were included in any of the analyses, the results are based on all detected emotions.

SD = 91.51) in late REM dream reports than in early REM dream reports (*N* = 15; *M* = 93.24; *SD* = 71.47), $t(14) = -2.39$, $p < .05$, $d = 0.59$. This was corroborated by the positive correlation between the length of the dream report and the number of emotional dream reports from both early REM, $r_s(15) = .52$, $p < .05$, and late REM, $r_s(17) = .78$, $p < .001$, sleep.

When directly comparing the two measures, with self-ratings a significantly larger number of dreams was rated to be emotional than with external ratings, Wilcoxon $Z = -3.63$, $p < .001$, $r = -.62$. Thus, hypothesis 1 was confirmed.

3.3. Hypothesis 2: Do self-ratings reflect a balanced ratio of positive and negative dreams, whereas external ratings indicate a larger number of negative than positive dreams?

The overall emotional valence of dreams was calculated with the 18 reliable items of the fmDES with both the positive emotions (PE) and negative emotions (NE) subscales consisting of 9 items or categories of emotion (excluding the two items/categories with poor item-to-total correlations with the respective subscale). When the frequencies of positive and negative emotion categories in a given dream were equal, the dream was referred to as a balanced dream, otherwise it was classified as either positive (more positive than negative emotion categories) or negative (more negative than positive emotion categories). With SR, all the dreams were included in the analyses. With ER, 7 out of the 33 emotional dream reports were excluded as they were rated to only contain either of the two items not considered as part of the PE or NE subscales (referred to as undetermined dreams).

The proportion of dreams with different emotional valence, that is, with positive, balanced or negative overall contents for both SR and ER can be seen in Table 2.

With SR, subjects (*N* = 17) had significantly more positive than negative ratings of dreams, Wilcoxon $Z = -3.63$, $p < .001$, $r = -.62$, and more positive than balanced ratings of dreams, Wilcoxon $Z = -3.63$, $p < .001$, $r = -.62$. However, there were no differences between the number of negative and balanced ratings of dreams, Wilcoxon $Z = -0.66$, $p > .05$, $r = -.11$. When controlling for the number of dreams, early and late REM dreams did not differ in the number of positive, Wilcoxon $Z = -0.04$, $p > .05$, $r = -.01$, negative, Wilcoxon $Z = -0.29$, $p > .05$, $r = -.05$, or balanced, Wilcoxon $Z = -0.09$, $p > .05$, $r = -.02$, ratings.

With ER, subjects (*N* = 17) had significantly more negative than balanced ratings of dream reports, Wilcoxon $Z = -2.59$, $p < .01$, $r = -.45$. There were no differences between the number of positive and balanced, Wilcoxon $Z = -2.11$, $p > .05$, $r = -.36$, or positive and negative ratings of dream reports, Wilcoxon $Z = -.57$, $p > .05$, $r = -.09$. When controlling for the number of dreams, early and late REM dream reports did not differ in the number of positive, Wilcoxon $Z = -2.02$, $p > .05$, $r = -.37$, negative, Wilcoxon $Z = -1.97$, $p > .05$, $r = -.36$, or balanced, Wilcoxon $Z = -0.45$, $p > .05$, $r = -.08$ ratings.

When directly comparing the two measures, significantly more dreams were rated to be positive with SR than with ER, Wilcoxon $Z = -3.64$, $p < .001$, $r = -.62$. However, there were no differences in the number of negative, Wilcoxon $Z = -0.30$, $p > .05$, $r = -.05$, or balanced, Wilcoxon $Z = -2.06$, $p > .05$, $r = -.35$, ratings of dreams (see Table 3 for descriptive statistics).

When including only those subjects who were rated to have emotional dreams with both methods (i.e., 11 subjects out of 17) the same pattern of results remained: a significant difference between self- and external ratings in the number of positive (SR: *M* = 5.63; *SD* = 2.42; ER: *M* = 1.00; *SD* = 1.00), Wilcoxon $Z = -2.94$, $p < .01$, $r = -.63$, but not negative (SR: *M* = 1.09; *SD* = 0.94; ER: *M* = 1.18; *SD* = 0.75), Wilcoxon $Z = -0.33$, $p > .05$, $r = -.07$, or balanced (SR: *M* = 0.55; *SD* = 0.69; ER: *M* = 0.18; *SD* = 0.40), Wilcoxon $Z = -1.63$, $p > .05$, $r = -.35$ ratings of dreams (see Fig. 2).

In sum, with self-ratings a significantly larger number of dreams was rated to be positive than negative, whereas with external ratings a rather balanced ratio of positive to negative dream estimations was obtained. Thus, in contrast to original predictions, both types of measurements resulted in a higher proportion of positively rated dreams. However, as expected, externally rated dreams were evaluated as relatively less positive than self-rated dreams. As such, hypothesis 2 was partly confirmed.

3.4. Hypotheses 3 and 4: Do self-ratings reflect a larger number of emotions per dream than external ratings? Do self-ratings and external ratings differ more in the number of positive emotions than negative emotions per dream?

The overall number of positive (PE) and negative (NE) emotions per dream was calculated by summing up the occurrences of the 9 positive and the 9 negative emotion categories. As such, the maximum number of both, PE and NE was 9 and the maximum number of all different emotions per dream 18.

Table 2
Proportion of positive, negative, balanced, and non-emotional dreams as measured with self-ratings and external ratings (*N* = 115) based on intensity and frequency ratings of emotions.

	Intensity	Frequency	
	Self-ratings (%)	Self-ratings (%)	External ratings (%)
Non-emotional dreams	0.0	0.0	71.3
Positive dreams	82.6	79.1	9.6
Negative dreams	13.9	12.2	11.3
Balanced dreams	3.5	8.7	1.7
Undetermined[a]	0.0	0.0	6.1

[a] Dreams rated to contain only either of the two items (awe/wonder/amazement; embarrassed/self-conscious/blushing) were not considered as part of the positive emotion (PE) or negative emotion (NE) subscales.

Table 3
Mean number of emotionally valenced dreams (out of the total of 115) and emotions in a dream per subject ($N = 17$) as measured with self-ratings and external ratings based on intensity and frequency ratings of emotions.

	Intensity		Frequency					
	Self-ratings		Self-ratings		External ratings		Wilcoxon signed rank test (two-tailed)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*p*	*r*
Emotional dreams	6.76	3.05	6.76	3.05	1.94	1.78	<.001	−.62
Positive dreams	5.59	2.71	5.35	2.71	0.65	0.93	<.001	−.62
Negative dreams	0.94	0.83	0.82	0.88	0.76	0.83	>.05	−.05
Balanced dreams	0.24	0.56	0.59	0.87	0.12	0.33	>.05	−.35
Undetermined[a]	0.00	0.00	0.00	0.00	0.41	0.62	N/A	N/A
Emotions in a dream	11.94	5.46	7.24	2.91	0.31	0.34	<.001	−.62
Positive emotions	8.94	4.45	5.26	1.92	0.13	0.19	<.001	−.62
Negative emotions	3.00	2.32	1.97	1.51	0.18	0.22	<.001	−.62

[a] Dreams rated to only contain either of the two items were not considered as part of the positive emotion (PE) or negative emotion (NE) subscales.

With SR, a subject reported, on average, 7.24 ($SD = 2.91$) different categories of emotions when measured with PE ($M = 5.26$; $SD = 1.92$) and NE subscales ($M = 1.97$; $SD = 1.51$) of the fmDES. The larger number of different positive emotions present per dream compared to negative emotions was statistically significant, Wilcoxon $Z = -3.62$, $p < .001$, $r = -.62$. Early and late REM dreams did not differ in the number of positive, $t(14) = -0.62$, $p > .05$, $d = -0.10$, or negative, Wilcoxon $Z = -0.97$, $p > .05$, $r = -.18$, emotion categories per dream.

When measured with ER, only 0.31 ($SD = 0.34$) different types of emotion categories per dream report were detected. The difference between the number of positive ($M = 0.13$; $SD = 0.19$) and negative ($M = 0.18$; $SD = 0.22$) emotion categories was not statistically significant, Wilcoxon $Z = -0.87$, $p > .05$, $r = -.15$. As with SR, there were no differences between early and late REM dream reports in the number of positive, Wilcoxon $Z = -1.16$, $p > .05$, $r = -.21$, and negative, Wilcoxon $Z = -1.41$, $p > .05$, $r = -.27$, emotion categories per dream.

When comparing the two measures, SR reflected a significantly larger mean number of different emotion categories in general, Wilcoxon $Z = -3.62$, $p < .001$, $r = -.62$, as well as of positive, Wilcoxon $Z = -3.62$, $p < .001$, $r = -.62$, and negative, Wilcoxon $Z = -3.62$, $p < .001$, $r = -.62$, emotion categories in a dream per subject than ER (see Table 3 for descriptive statistics). This was also the case when only those 11 subjects who were evaluated to have emotional dreams with ER were analysed, Wilcoxon $Z = -2.94$, $p = .001$, $r = -.63$. With SR there were significantly more estimates of different types of positive emotions ($M = 4.98$; $SD = 1.28$) and negative emotions ($M = 2.02$; $SD = 1.03$) than with ER ($M_{PE} = 0.20$; $SD_{PE} = 0.20$; $M_{NE} = 0.27$; $SD_{NE} = 0.22$) (PE, Wilcoxon $Z = -2.94$, $p = .001$, $r = -.63$; NE, Wilcoxon $Z = -2.93$, $p = .001$, $r = -.63$) (see Fig. 3).

Exploring the relationship between the two measures using Spearman's rank correlation coefficient revealed that the NE subscales of self- and external ratings were positively correlated, $r_s(17) = .52$, $p < .05$, whereas there was no significant relationship between the PE subscales, $r_s(17) = .18$, $p > .05$.

As predicted, self-ratings resulted in larger estimates of the number of emotions per dream than external ratings. Moreover, the difference was more pronounced with respect to positive than negative emotions. The correlation analysis indicates

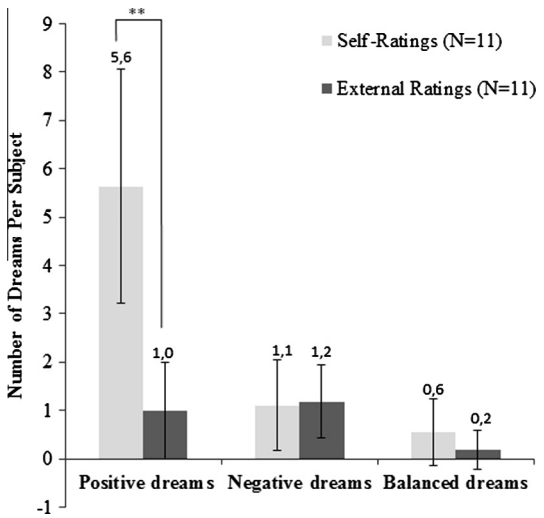

Fig. 2. Average number of emotionally valenced dreams per subject as measured with self-ratings (SR) and external ratings (ER). Includes only those subjects whose dreams were rated as emotional by external raters. $^{**}p < .01$.

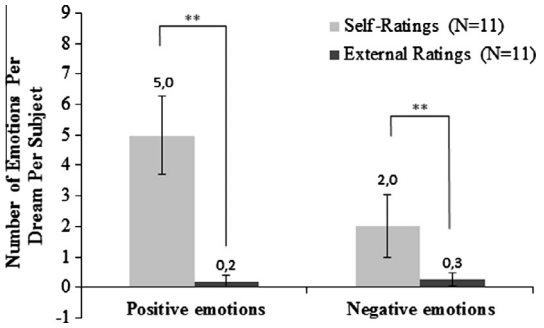

Fig. 3. Average number of different categories of positive and negative emotions per subject as measured with the positive emotion (PE) and negative emotion (NE) subscales of the fmDES, respectively. Includes only those subjects who dreams were rated as emotional by external judges. $^{**}p < .001$.

that the ratings of negative emotions of the two measures correspond better to each other than the ratings of positive emotions. As such, hypotheses 3 and 4 were confirmed.

3.5. Exploratory analyses

3.5.1. Relationship between positive and negative emotion subscales

Irrespective of the measure used, positive and negative emotion subscales have been typically found to be structurally independent or weakly negatively correlated (Schimmack, 2008). In accordance with this, in the current study the PE and NE subscales were not significantly correlated, r_s (17) = .28, $p > .05$ when measured with SR. With ER, however, the PE and NE subscales were moderately positively correlated, r_s (17) = .49, $p < .05$. This indicates that the results obtained with self-ratings here correspond better to previous findings than those obtained with external ratings.

Both, PE and NE subscales were positively correlated with the length of the dream report when analysed with ER, indicating that the longer the dream report the more positive, r_s (17) = .65, $p < .01$, and negative, r_s (17) = .78, $p < .001$, emotions could be detected. With SR, on the other hand, only NE was positively correlated with the length of the preceding dream report, r_s (17) = .58, $p < .05$. Hence, the longer the dream report the more emotions were detected with ER. With SR, dreams rated to contain more negative emotions were also described in more words than those rated to contain less negative emotions.

3.5.2. Frequency and distribution of discrete emotions in dreams

The occurrence of discrete emotions per emotional dream was calculated using all the 20 items of the fmDES according to whether a particular emotion category was rated to occur at least once with self-ratings (irrespective of the intensity ratings) or with external ratings (irrespective of how many times it was coded in one dream). With SR, all of the 115 dreams and with ER the 33 emotional dream reports were included in the analysis and an aggregate score across all dreams for each individual for each discrete emotion calculated.

The average occurrence of each of the discrete emotion categories per emotional dream, measured with both self- and external ratings, is depicted in Table 4. When looking at all the emotional dreams as measured with SR (115 dreams of 17 subjects), almost all of the various positive emotions were more frequently reported than the various negative emotions. In the emotional dream reports as measured with ER (33 dreams of 11 subjects), only twelve (five positive and seven negative) of the 20 emotion categories were detected.

To directly compare the two measures, the same 33 dreams of the 11 subjects rated to contain emotions with both SR and ER were analysed. Eight positive emotions and five negative emotions were rated to occur more often with SR than with ER (see Table 4 for descriptive statistics and significance tests). The occurrence of other discrete emotions did not differ between the two measures and these were typically among the fewest to be reported or detected with either measure.

3.5.3. Intensity of self-ratings

All the above analyses are based on the frequency of occurrence, rather than the intensity, of dream emotions. It is possible, however, that some emotions (e.g., positive emotions) may be experienced frequently but mildly and others (e.g., negative emotions) infrequently but intensely. As such, analyses based on the frequency and intensity ratings may yield different results. Therefore, additional analyses were conducted in which the intensity of self-rated emotions was investigated. First, the intensity ratings of the PE and NE for each dream was obtained by summing up the actual scores (ranging from 0 to 4) given to each of the nine positive and negative emotion categories, respectively. As such, the maximum intensity score for both PE and NE was 36. When the intensity scores of PE and NE in a given dream were equal, the dream was referred to as a balanced dream, otherwise it was classified as either positive (PE > NE) or negative (PE < NE). Then, an aggregated mean of the sum scores of PE and NE across all the dreams for each individual was calculated.

The proportion of dreams with different emotional valence, that is, with positive, balanced or negative ratings of overall contents can be seen in Table 2. Across the two nights and irrespective of the time of the night, the 17 subjects had

Table 4
Mean number of discrete emotions per emotional dream per subject as measured with self- and external ratings based on intensity and frequency ratings of emotions.

	115 Dreams of 17 subjects				33 Dreams of 11 subjects					
	Self-ratings				Self-ratings		External ratings		Wilcoxon signed rank test (two-tailed)	
	Intensity		Frequency		Frequency		Frequency			
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*p*	*r*
Positive emotions										
Interested/Alert/Curious	1.61	0.63	0.87	0.14	0.95	0.10	0.18	0.32	<.01	−.61
Joyful/Glad/Happy	1.53	0.64	0.81	0.24	0.79	0.33	0.05	0.10	<.01	−.61
Serene/Content/Peaceful	1.25	0.58	0.75	0.28	0.64	0.39	0.02	0.08	<.01	−.58
Amused/Fun loving/Giggly	1.00	0.68	0.59	0.33	0.63	0.32	0.19	0.20	<.01	−.61
Proud/Confident/Self-assured	0.89	0.57	0.56	0.30	0.52	0.38	0.00	0.00	<.01	−.54
Hopeful/Optimistic/Encouraged	0.82	0.60	0.54	0.29	0.36	0.29	0.00	0.00	<.01	−.54
Inspired/Uplifted/Elevated	0.81	0.78	0.48	0.30	0.32	0.36	0.00	0.00	<.05	−.47
Love/Closeness/Trust	0.65	0.67	0.40	0.34	0.35	0.37	0.00	0.00	<.05	−.47
Awe/Wonder/Amazement[a,b]	0.57	0.44	0.42	0.30	0.52	0.39	0.28	0.31	>.05	−.39
Grateful/Appreciative/Thankful	0.37	0.49	0.28	0.32	0.18	0.32	0.00	0.00	>.05	−.35
Negative emotions										
Angry/Irritated/Annoyed	0.64	0.42	0.39	0.23	0.52	0.31	0.19	0.20	<.05	−.51
Stressed/Nervous/Overwhelmed	0.57	0.42	0.39	0.25	0.59	0.34	0.12	0.18	<.01	−.57
Hate/Distrust/Suspicion	0.40	0.41	0.25	0.25	0.43	0.37	0.05	0.12	<.01	−.54
Disgust/Distaste/Revulsion	0.36	0.38	0.22	0.26	0.45	0.44	0.11	0.30	>.05	−.44
Embarrassed/Self-conscious/Blushing[b]	0.29	0.45	0.21	0.29	0.14	0.21	0.02	0.08	>.05	−.35
Contemptuous/Scornful/Disdainful	0.28	0.38	0.20	0.26	0.32	0.39	0.00	0.00	<.05	−.47
Sad/Downhearted/Unhappy	0.26	0.32	0.19	0.21	0.29	0.37	0.00	0.00	<.05	−.52
Scared/Fearful/Afraid	0.22	0.29	0.13	0.15	0.17	0.21	0.05	0.10	>.05	−.39
Ashamed/Humiliated/Disgraced	0.14	0.25	0.09	0.13	0.19	0.32	0.00	0.00	>.05	−.39
Guilty/Repentant/Blameworthy	0.11	0.21	0.12	0.21	0.24	0.34	0.02	0.08	>.05	−.38

[a] With external ratings includes the emotion of *surprise*.
[b] Omitted from positive emotion (PE) and negative emotion (NE) subscales due to poor item-to-total correlations.

significantly more positive than negative ratings of dreams, Wilcoxon $Z = -3.64$, $p < .001$, $r = -.62$, more positive than balanced ratings of dreams, Wilcoxon $Z = -3.63$, $p < .001$, $r = -.62$, and more negative than balanced ratings of dreams, Wilcoxon $Z = -2.76$, $p < .001$, $r = -.47$ (see Table 3 for descriptive statistics). Positive emotions were rated to be significantly more intense ($M = 8.94$; $SD = 4.45$) than negative emotions ($M = 3.00$; $SD = 2.32$), Wilcoxon $Z = -3.62$; $p < .001$, $r = -.62$. As was the case for the frequency data, there was a positive correlation between the length of the dream report and the intensity of the negative emotions rated to be present in the preceding dream, $r_s(17) = .57$, $p < .05$. This was not observed with respect to positive emotions, $r_s(17) = -.06$, $p > .05$. The PE and NE subscales were not correlated with each other, $r_s(17) = .28$, $p > .05$.

Analysis of the relationship between the frequency and intensity measures demonstrated a near perfect correlation with respect to both, for PE, $r_s(17) = .96$, $p < .001$, and for NE, $r_s(17) = .97$, $p < .001$, subscales. The same applied to discrete emotions with correlations between the frequency and intensity measures of all the emotion categories ranging from $r_s(17) = .77$ (interested/alert/curious) to $r_s(17) = .982$ (grateful/appreciative/thankful), $ps < .001$, for positive emotions and from $r_s(17) = .835$ (angry/irritated/annoyed) to $r_s(17) = .1000$ (guilty/repentant/blameworthy), $ps < .001$, for negative emotions. As can be seen from Table 4, the overall distribution of discrete emotions was similar regardless of the analysis method used, that is, regardless of whether only the frequency of different emotion categories was used in the analyses or whether intensity was also taken into account.

4. Discussion

In the present study, self-ratings of emotions occurring in REM sleep dreams were compared to ratings performed by external judges to determine whether and to what extent the results concerning the overall emotionality and emotional valence of dreams as well as the number of positive and negative emotions and the occurrence of specific emotions in dreams differ depending on who rates the emotions. Below, the results are discussed separately for each of these aspects.

4.1. Emotionality of dreams

With self-ratings, every dream was rated to contain at least one type of emotion, whereas with external ratings, only approximately one third of the dream reports were evaluated as emotional. The proportion of emotional dreams when measured with self-ratings (100%) is similar to the figures reported in a study by St-Onge et al. (2005) in which subjects rated 90.6% of laboratory and 98.4% of their home dream reports as containing at least one type of emotion. The proportion of emotional dreams as measured with external ratings (29%) is in agreement with studies using traditional dream content analysis methods according to which 30–40% of dream reports are rated as emotional (e.g., Schredl & Doll, 1998; Snyder, 1970).

However, this figure is much lower when compared to the external ratings using a scale identical to self-ratings in the study by Schredl and Doll (1998), as in the latter case the proportion of emotional dreams as measured by external judges (85.6%) was similar to that obtained by self-ratings (92.7%). In Schredl and Doll (1998) judges rated the overall intensity of positive and negative emotions, but in the current study the expression of specific emotions was rated. As such, the Schredl and Doll (1998) method most likely included mood states which might explain the increased dream emotionality they observed.

Therefore, the results obtained with the two methods corroborate previous findings that self-ratings of dream emotions produce significantly larger estimates of dream emotionality than external ratings based on dream reports.

4.2. Emotional valence of dreams

When emotions were externally rated based on dream reports given by the subjects, the proportions of positive and negative dreams were rather similar (9.6% and 11.3%, respectively). However, when emotions were rated by the subjects themselves upon awakening from the dream, positive dreams were more than six times more prevalent than negative dreams (79.2% and 12.2%, respectively). This was corroborated at the subject level of analysis which demonstrated that when self-rated, there were more than five times more positive dreams per subject than when emotions were externally rated, whereas the number of negative dreams was basically the same. Hence, the two methods differed as to the proportion of positive dreams.

Although the ratio of positive to negative dreams, as measured with both self- and external ratings, is different from earlier studies (Kramer et al., 1971; Schredl & Doll, 1998), the relationship between the two methods is in accordance with previous results according to which externally rated dreams are evaluated to be relatively more negative than self-rated dreams (e.g. Schredl & Doll, 1998).

4.3. Emotions in dreams

4.3.1. Total number of emotions

The overall number of emotions was significantly smaller when external judges evaluated the dream reports as opposed to when subjects themselves rated their dreams. Out of the 18 emotion categories, self-rated dreams had, on average, seven different emotion categories represented in one dream. In the externally rated dream reports, less than one emotion category per dream was detected. These findings are in agreement with previous results demonstrating a more than tenfold difference between the two methods in detecting the overall number of emotions per dream (Hall & Van de Castle, 1966; Merritt et al., 1994; Nielsen et al., 1991; Schredl & Doll, 1998). This difference is due to the fact that external raters can reliably evaluate only the emotions explicitly mentioned in dream reports.

4.3.2. Number of positive and negative emotions

When looking separately at the positive and negative emotion subscales, self-ratings reflected almost three times as many different positive than negative emotions per dream. When dream reports were externally rated, a similar number of positive and negative emotions were detected. Thus, whereas with self-ratings positive emotions outnumbered negative emotions in a dream, with external ratings the number of positive and negative emotions did not differ. Although self-ratings indicated an increased number of both positive and negative emotions per dream, it was especially the number of positive emotions that was underrepresented in dream reports and thus remained undetected by external raters. The latter was corroborated by the finding that whereas the negative emotion subscales of the self- and external ratings were moderately correlated, there was no such relationship between the positive emotion subscales. This indicates that dream emotions are differently represented in a verbal dream report than in the self-rating of emotions of the same dreams when subjects are specifically probed about the discrete emotions, and that it is the detection of positive emotions where the two measures diverge most notably. The latter was also found to be the case in the study by Schredl and Doll (1998).

4.4. Discrete emotions

All of the 20 different emotion categories were reported by the subjects to be present in their dreams. External judges, on the other hand, detected only slightly more than half (i.e., 12) of all the possible emotion categories in the dream reports. Whereas with self-ratings a positive emotion (interested/alert/curious) was the most frequently reported specific emotion, with external ratings a positive (amused/fun-loving/giggly[2]) and a negative (angry/irritated/annoyed) emotion were both among the two most frequently rated categories. Sadness, guilt and shame were rarely rated by subjects and never or almost never detected by external judges. These results are well in line with previous studies (Fosse et al., 2001; Hall & Van de Castle, 1966; Merritt et al., 1994; Schredl & Doll, 1998; Strauch & Meier, 1996). Only the findings concerning the frequency of the emotion category scared/fearful/afraid is incompatible with previous research. Whereas anxiety and fear have often been found to

[2] The emotion awe/wonder/amazement was the most frequently detected category with external ratings but external judges used this particular category for rating the emotion of surprise (i.e., meaning "I wondered about. . ." rather than "I was in wonder"). Depending on the context, surprise can have a negative or a positive valence (Watson & Clark, 1994) and hence cannot be considered as clearly belonging to either category.

dominate in dream reports (Hall & Van de Castle, 1966; Merritt et al., 1994; Nielsen et al., 1991; Snyder, 1970), in the present study it was among the least frequently rated with both self- and external ratings.

4.5. Possible explanations for the positivity of dream ratings

The question arises as to why dreams were evaluated to be so positive and to contain so many positive emotions, especially with self-ratings. This contradicts hypothesis 2, and the findings of a number of previous studies demonstrating that dreams are typically negatively biased (e.g., Hall & Van de Castle, 1966; Merritt et al., 1994; Snyder, 1970; Strauch & Meier, 1996; Valli & Revonsuo, 2009; Valli, Strandholm, Sillanmäki, & Revonsuo, 2008).

One possible reason may stem from the experimental environment, procedure and the timing of the REM awakenings. First, studies comparing the content of dreams in the home vs laboratory setting have often found home dreams to contain more physical and verbal aggression (e.g., Domhoff & Schneider, 1999; Weisz & Foulkes, 1970) and reported a lower prevalence of negative dreams as well as negative emotions in laboratory dreams (St-Onge et al., 2005). As discussed previously, spontaneously recalled home diary dreams may be negatively biased (Domhoff, 2005), whereas controlled laboratory awakenings from the same sleep stage throughout the experimental night used in the current study may have produced a more representative sample of REM dreams and their features.

Additionally, it may be argued that a dream typically evolves from being relatively positive to being relatively more negative (Merritt et al., 1994) and the timing of awakenings (5 min from the beginning of each REM stage) chosen for this study may have led to a biased sample of dreams. Letting dreams unfold naturally without interrupting them might result in an increased proportion of negative emotions, as is usually the case with dreams reported in the home setting. The results of Merritt et al. (1994) support this possibility, even though these researchers reported that a dream progresses "from bad to worse" (p. 56), that is, from being negatively toned to being even more negatively toned. A replication of the procedure with awakenings at the end of each REM stage could shed some light on this.

Another possible explanation for the decreased negativity and increased positivity of dream ratings in this study, as compared to previous studies, may have to do with the emotion rating scales used. First, the current study measured a balanced number of specific positive and negative emotions. The smaller number of negative dreams obtained with external ratings may reflect the negativity bias inherent in the traditional content analysis methodology itself in that the latter includes a larger number of negative than positive emotion categories. Therefore, it is not surprising that a preponderance of negative dreams and negative emotions is observed. Moreover, it may well be the case that a few negative emotions, such as fear, occur several times in a dream report as a response to different dream events, and measuring how many times the specific emotion occurs in total (rather than whether it is present in a dream or not, as done in the current study) may result in an increased number of negative emotions. However, in the present study the same emotion occurred more than once only in four (out of 33) dreams[3] and, moreover, this was not specific to negative emotions. Thus, the smaller number of negative emotions detected with external ratings in the present study cannot be explained by the underrepresentation of recurring negative emotions in dream reports.

Second, in the current study the frequency, rather than the intensity, of emotions was analysed. It can be argued that as in the waking state phenomena known as *the positivity offset* and the *negativity bias* (Cacioppo, Gardner, & Berntson, 1999), in dreams mild positive emotions prevail in frequency but negative emotions dominate when intensity is taken into account. However, as the exploratory analyses demonstrated, the results concerning the intensity of self-rated emotions were similar to those reflecting the frequency of self-rated emotions. Hence, such an argument is not supported.

Third, the scales applied in previous studies might have not enabled the detection of positive emotions in as much detail as the fmDES used in this study. It has been suggested that, compared to negative emotions, positive emotions are more diffuse and less differentiated (Fredrickson, 1998). This means that there are more ways to express negative than positive emotional experiences (Ben-Ze'ev, 2000), a bias reflected in the unequal number of words denoting positive and negative emotions in both the English (Averill, 1980) and the Finnish (Tuovila, 2005) language. Therefore, unless explicitly probed, individuals may underreport positive emotions not only in their dream reports but also when rating their general mood. Indeed, in the current study, when externally rated, dream reports that contained more words were also rated to be more emotional and to contain more of both, positive and negative emotions. When dreams were self-rated, however, the length of the dream report correlated only with the number of reported negative emotions. This suggests that subjects are good at expressing negative emotional states, whereas positive emotions remain more ineffable for verbal reporting. In a similar vein, due to the more diffuse nature of positive emotions, it may well be the case that when individuals report the occurrence of one positive emotion they are more inclined to report the existence of other types of positive emotions as well. As discussed above, negative emotions are more distinct and hence, the reporting of their co-occurrence less likely. Such a carry-over effect of positive emotions, however, was not supported by the results of the current study as the inter-item correlations of the self-reported positive emotions remained moderate (ranging from 0.2 to 0.7) and did not exceed those of negative emotions (ranging from 0.2 to 0.6).

[3] In the majority of dreams the same emotion category was rated to occur only once per dream report. Only in four (out of 33) dreams was the same emotion category coded more than once (twice in four dreams and three times in one dream). Two of these emotion categories were positive and two negative.

A further possible explanation concerns the small number of subjects in the study. Because the subjects ($N = 17$) were selected out of the initial population of $N = 159$ through various phases, the selection criteria may have resulted in a group that does not represent the average of the general population in terms of their emotional experiences. Due to the sample consisting of young adults with a particular set of characteristics initially selected for, the results of the study apply to populations with similar demographics. It is not known whether similar results would have been obtained with individuals belonging to a different age, socioeconomic and/or country/cultural group. Moreover, subjects with poor sleep quality are more likely to suffer from negative emotional states, e.g., depression or daytime distress, and therefore also display more negatively toned dreams (Selvi et al., 2012). Thus, the selection criteria (e.g., good sleep quality) may have inadvertently resulted in a small group of exceptional subjects who in general, also during wakefulness, experience more positive and less negative emotionality than the average population, which is consequently reflected in their emotional dream content. Additionally, it can be argued that the individuals who composed the final sample were highly motivated to complete the study and displayed a generally positive attitude toward dreams. Analysis of pre-experimental attitudes, however, demonstrated that the 17 participants did not display a more positive disposition toward dreams than those selected out of the study.[4]

The higher positivity of self-ratings, as compared to external ratings, may also reflect differences in our linguistic tendency to attribute positive emotions to external factors, rather than to ourselves. Dream reports that were rated as non-emotional by external judges but positive by the dreamers themselves often contained characters (e.g., parents, siblings, friends, pets), actions (e.g., being at a concert, at a party) and other situational features (e.g., beautiful day, sunshine) without direct references to emotional experiences despite these presumably having been experienced as positively toned. In fact, in our waking life we seem to describe positive events often in an impersonal manner (e.g., *It was such a beautiful day*; *It was such a great party*; *It was a really good concert*; *The kittens were so cute*) rather than referring to how these make us feel. As a result, external judges rating the dream reports may not have been able to detect emotions despite the subjects having experienced them in dreams.

In sum, the positivity of dream ratings in the current study challenges not only the study conducted by Schredl and Doll (1998) but also several previous studies and theories based on the assumption of *the negativity bias* in dreams (Cartwright, 1996; Flanagan, 2000; Hartmann, 1995; Kramer, 1991; Revonsuo, 2000) and calls for further research to clarify this contradiction.

5. Limitations of present study and suggestions for future research

A systematic order effect might have influenced the results. As the self-ratings of emotions were always conducted after the oral dream report, it is possible that if the order had been reversed, dream reports might have contained more emotions due to the priming effect of focusing on a list of specific emotions. At the same time, it can be argued, that if the latter had been the case, the participants would have displayed a learning effect over the course of one and/or two laboratory nights. However, the dream reports obtained during the second laboratory night did not differ from those collected during the first night in the characteristics of emotionality.[5] As it is not possible to refute either of the claims without direct experimental verification, future studies could counterbalance the order of giving a verbal dream report and self-rating the dream emotions using a scale in the experimental design. It has to be noted though, that counterbalancing the order poses additional problems in that if the dream is not reported immediately upon awakening, intervening tasks and/or cognitive processes can easily interfere with the dream memory and result in an alteration or decay of the dream experience (Parke & Horton, 2009).

Another issue has to do with the instructions concerning the dream reports. Whereas the external judges were asked to rate emotions that were directly expressed or could be unambiguously inferred from the behaviour of the dream self or from the dream plot, the instructions given to the subjects were not specific to emotions but the participants were simply asked to describe the content in as much detail as possible. Research demonstrates that without being directly probed, people focus on the story line of the dream (who, what, when, where), rather than on the process (emotional, perceptual, cognitive) aspects (Kahan & Horton, 2012). Therefore, despite having experienced emotions, participants may not have expressed them in their dream reports, unless the dream contained very vivid and intense emotions. This might have led to the underestimation of emotions with external ratings in general, and due to the *negativity bias*, the underestimation of positive emotions in particular. On the other hand, given that the participants went through the same procedure of reporting the dreams and rating the dream emotions several times a night and during two experimental nights, it is possible that the participants understood the underlying idea of the study after the first laboratory night. As a result, they might have tended to report more emotions in their oral dream report during the second experimental night. As stated above, the two laboratory nights did not differ in

[4] The original background questionnaire included three items measuring the individuals' attitudes towards their dreams in general. The Mann-Whitney test showed no significant differences between the final sample of 17 subjects and the rest of the 142 individuals who filled in the questionnaire in their answers to the questions "How often do you discuss your dreams with family or friends?" (0 = *never*, 6 = *always*), $U = 1175.00$, $Z = -0.09$, $r = -.01$ ($M_{17} = 3.29$; $SD_{17} = 1.36$; $M_{140} = 3.32$; $SD_{136} = 1.31$); "How much attention do you usually pay to your dreams?" (0 = *none*, 6 = *a lot*), $U = 932.50$, $Z = -1.53$, $r = -.12$ ($M_{17} = 3.00$; $SD_{17} = 1.28$; $M_{141} = 3.54$; $SD_{136} = 1.31$); "How much significance do you usually attach to your dreams?" (0 = *none*, 6 = *a lot*), $U = 991.00$, $Z = -1.19$, $r = -.09$ ($M_{17} = 2.06$; $SD_{17} = 1.25$; $M_{141} = 2.54$; $SD_{136} = 1.37$).

[5] There were no significant differences between the two laboratory nights in the number of emotional dreams, $Z = -0.32$, $p > .05$, $r = -.06$; positive dreams, $Z = -0.32$, $p > .05$, $r = -.08$; negative dreams, $Z = -1.00$, $p > .05$, $r = -.18$; balanced dreams, $Z = -0.00$, $p > .05$; undifferentiated dreams, $Z = -1.34$, $p > .05$, $r = -.24$; different types of emotion categories per dream, $Z = -0.05$, $p > .05$, $r = -.01$; different types of positive emotions per dream, $Z = -0.00$, $p > .05$; or different types of negative emotions per dream, $Z = -0.25$, $p > .05$, $r = -.04$.

the emotionality of dream reports (footnote 5), thus not offering support for such a possibility. Nevertheless, studies investigating the emotional content of dreams could benefit from explicitly directing the participants' attention to the emotional content of the dream by providing more emotion-specific instructions. Moreover, future studies could consider using participants highly trained in reporting emotional experiences (both in dreams and waking life) (Kahan, 2012). That way dream reports could perhaps be more comprehensive and accurate, and, as such, external ratings might yield similar results as self-ratings.

The finding that external raters were more likely to detect emotions in longer dream reports refers to the possibility that the process aspects of subjective experience, emotion in this particular case, are more likely to be reported late in the narrative (Kahan & Horton, 2012). Whether this is the case with regards to dream reports, that is, whether emotions are (or are not) more likely reported towards the end of the dream report remains an open question and future studies are needed to shed light on this.

Additionally, whereas the external ratings were based on oral *dream reports*, that is, on the verbal descriptions of the immediate dream experience, the self-ratings were essentially retrospective evaluations of *dream experiences*. It can be argued that evaluating dream reports vs dream experiences are not directly comparable and may explain the different findings obtained with the two methods. On the other hand, due to the fact that self-ratings were conducted subsequent to giving the oral dream report it is likely that the subjects based their ratings on the preceding dream report rather than the actual experience. Moreover, self-ratings of previous dream experiences are essentially ratings of reports (even in the absence of an overt oral report the subjects may nevertheless generate an internal narrative as a basis for their ratings) as the dream experience itself is not accessible anymore. In fact, only by using lucid dreamers, who rate their experiences while the dream unfolds, could the dream emotions be evaluated as they are experienced. Hence, future studies could ideally include a group of lucid dreamers and compare their self-ratings of emotions carried out while having the dream *experience* to those reported upon awakening. Of course, it has to be taken into account that becoming lucid in a dream alters the dream experience, including the emotional content of the dream (LaBerge & DeGarcia, 2000), and the results obtained using lucid dreams may not apply to non-lucid dreams. Alternatively, studies could add a further self-rating condition in which the subjects retrospectively rate their transcribed *dream reports* similarly to the external judges.

A related issue has to do with the fact that the subjects rated their emotions after the dream experience and in a different state of consciousness and, as such, the waking state might have influenced their results. It can be questioned to what extent the emotions rated by the subjects using the fmDES corresponded to the emotions originally *experienced* in the dream, and to what extent they were induced only in the waking state by cognitively *evaluating* the preceding dream (report). It has been argued that, despite being correlated, the "experiencing self" does not perfectly correspond to the "remembering and evaluative self" (Kahneman & Riis, 2005, p. 285). In order to know whether the differences in the emotional characteristics obtained with the different rating methods are state-dependent (dreaming vs waking state) or dependent on the immediacy of the reporting to the experience (experiencing self vs remembering self), a similar procedure using self-ratings and verbal reports of waking experiences could be employed and compared to those of dream experiences. Although there are indications that a similar discrepancy between self- and external ratings can be observed when evaluating waking reports (Kahan & LaBerge, 1996), it is unclear to what extent this applies to the various characteristics of emotional experiences.

6. Conclusion

Our findings demonstrate that self-ratings, as compared to external ratings, result in greater estimates of (a) emotional dreams; (b) positively valenced dreams; (c) both positive and negative emotions per dream; and (d) various discrete emotions in dreams. Thus, the results question the convergent validity between self- and external ratings of dream emotions and highlight the importance of implementing carefully controlled study designs, data collection and analysis methods in the study of affective contents of (dream) consciousness.

Acknowledgments

The authors would like to thank Nils Sandman for his invaluable help in this study as well as the editor and the three anonymous reviewers for their insightful comments and useful suggestions.

This research was supported by the Academy of Finland (Project 266434), the Signe and Ane Gyllenberg Foundation (Pilleriin Sikka) and the Turku Institute for Advanced Studies (Katja Valli).

References

Agargun, M. Y., & Cartwright, R. (2003). REM sleep, dream variables and suicidality in depressed patients. *Psychiatry Research, 119*, 33–39. http://dx.doi.org/10.1016/S0165-1781(03)00111-2.

Antrobus, J. (1983). REM and NREM sleep reports: Comparison of word frequencies by cognitive classes. *Psychophysiology, 20*, 562–568. http://dx.doi.org/10.1111/j.1469-8986.1983.tb03015.x.

Averill, J. R. (1980). On the paucity of positive emotions. In K. R. Blankstein, P. Pliner, & J. Polivy (Eds.), *Advances in the study of communication and affect. Assessment and modification of emotional behavior* (Vol. 6, pp. 7–45). New York: Plenum.

Ben-Ze'ev, A. (2000). *The subtlety of emotions*. Cambridge, MA: MIT Press.

Blagrove, M., Framer, L., & Williams, E. (2004). The relationship of nightmare frequency and nightmare distress to well-being. *Journal of Sleep Research, 13*, 129–136. http://dx.doi.org/10.1111/j.1365-2869.2004.00394.x.

Buysse, D. J., Reynolds, C. F., III, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*, 193–213.
Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*, 839–855. http://dx.doi.org/10.1037/0022-3514.76.5.839.
Cartwright, R. D. (1996). Dreams and adaptation to divorce. In D. Barrett (Ed.), *Trauma and dreams* (pp. 173–185). Cambridge, MA: Harvard University Press.
Casagrande, M., Violani, C., Lucidi, F., Buttinelli, E., & Bertini, M. (1996). Variations in sleep mentation as a function of time of night. *International Journal of Neuroscience, 85*, 19–30.
Cohen, D. B. (1969). Frequency of dream recall estimated by three methods and related to defence preference and anxiety. *Journal of Consulting and Clinical Psychology, 33*, 661–667. http://dx.doi.org/10.1037/h0028438.
Domhoff, G. W. (2005). The content of dreams: Methodologic and theoretical implications. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (4th ed., pp. 522–534). Philadelphia: W.B. Saunders.
Domhoff, G. W., & Schneider, A. (1999). Much ado about very little: The small effect sizes when home and laboratory collected dreams are compared. *Dreaming, 9*, 139–151. http://dx.doi.org/10.1023/A:1021389615347.
Flanagan, O. (2000). Dreaming is not an adaptation. *Behavioral and Brain Sciences, 23*, 936–939. http://dx.doi.org/10.1017/S0140525X00404024.
Fosse, R., Stickgold, R., & Hobson, J. A. (2001). The mind in REM sleep: Reports of emotional experience. *Sleep, 24*, 947–955.
Foulkes, D., Sullivan, B., Kerr, N. H., & Brown, L. (1988). Appropriateness of dream feelings to dreamed situations. *Cognition & Emotion, 2*, 29–39. http://dx.doi.org/10.1080/02699938808415227.
Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*, 300–319. http://dx.doi.org/10.1037/1089-2680.2.3.300.
Fredrickson, B. L. (2013). Positive emotions broaden and build. In P. G. Devine & E. A. Plant (Eds.). *Advances in experimental social psychology* (Vol. 47, pp. 1–53). San Diego, CA: Academic Press.
Fredrickson, B. L., Tugade, M. M., Waugh, C. E., & Larkin, G. R. (2003). What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *Journal of Personality and Social Psychology, 84*, 365–376. http://dx.doi.org/10.1037/0022-3514.84.2.365.
Hall, C. S., & Van de Castle, R. L. (1966). *The content analysis of dreams*. New York: Appleton-Century-Crofts.
Hartmann, E. (1995). Making connections in a safe place: Is dreaming psychotherapy? *Dreaming, 5*, 213–228.
Heinilä, A. (2008). *3-9-vuotiaiden lasten unennäkö ja unisisällöt: unipäiväkirjatutkimus kotioloissa* (Master's thesis). University of Turku, Finland.
Hobson, J. A. (2009). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience, 10*, 803–813. http://dx.doi.org/10.1038/nrn2716.
Hobson, J. A., Pace-Schott, E. F., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences, 23*, 793–1121. http://dx.doi.org/10.1017/S0140525X00003976.
Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications* (1st ed.). Westchester, Illinois: American Academy of Sleep Medicine.
Kahan, T. L. (2012). Cognitive expertise and dreams. In D. Barrett & P. McNamara (Eds.), *Encyclopedia of sleep and dreams* (pp. 135–139). Santa Barbara, CA: Greenwood Publishers.
Kahan, T. L., & Horton, C. L. (2012). Methodological challenges in dream science. In D. Barrett & P. McNamara (Eds.), *Encyclopedia of sleep and dreams* (pp. 418–421). Santa Barbara, CA: Greenwood Publishers.
Kahan, T. L., & LaBerge, S. (1996). Cognition and metacognition in dreaming and waking: Comparisons of first and third-person ratings. *Dreaming, 6*(4), 235–249.
Kahneman, D., & Riis, J. (2005). Living, and thinking about it: Two perspectives on life. In F. A. Huppert, N. Baylis, & B. Keverne (Eds.), *The science of well-being* (pp. 285–304). Oxford, UK: Oxford University Press.
Kramer, M. (1991). The nightmare: A failure in dream function. *Dreaming, 1*(4), 227–285.
Kramer, M., Winget, C., & Whitman, R. M. (1971). A city dream: A survey approach to normative dream content. *American Journal of Psychiatry, 127*, 1350–1356.
LaBerge, S., & DeGarcia, D. J. (2000). Varieties of lucid dreaming experience. In R. G. Kunzendorf & B. Wallace (Eds.), *Individual differences in conscious experience* (pp. 269–307). Amsterdam: John Benjamins.
Mealey, L. (2000). The illusory function of dreams: Another example of cognitive bias. *Behavioral and Brain Sciences, 23*(6), 971–972. http://dx.doi.org/10.1017/S0140525X00003976.
Merritt, J. M., Stickgold, R., Pace-Schott, E., Williams, J., & Hobson, J. A. (1994). Emotion profiles in the dreams of men and women. *Consciousness and Cognition, 3*, 46–60. http://dx.doi.org/10.1006/ccog.1994.1004.
Nielsen, T. A., Deslauriers, D., & Baylor, G. W. (1991). Emotions on dream and waking event reports. *Dreaming, 1*, 287–300.
Nir, Y., & Tononi, G. (2010). Dreaming and the brain: From phenomenology to neurophysiology. *Trends in Cognitive Sciences, 14*, 88–100. http://dx.doi.org/10.1016/j.tics.2009.12.001.
Parke, A. R., & Horton, C. L. (2009). A re-examination of the interference hypothesis on dream recall and dream salience. *International Journal of Dream Research, 2*(2), 60–69.
Pivik, R. T. (1991). Tonic states and phasic events in relation to sleep mentation. In S. J. Ellman & J. S. Antrobus (Eds.), *The mind in sleep: Psychology and psychophysiology* (2nd ed., pp. 214–247). New York, NY: John Wiley & Sons Inc.
Rechtschaffen, A., & Kales, A. (Eds.). (1968). *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Washington: U.S. Government Printing Office.
Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences, 23*, 877–901. http://dx.doi.org/10.1017/S0140525X00003976.
Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon*. Cambridge, MA & London: MIT Press.
Revonsuo, A., & Salmivalli, C. (1995). A content-analysis of bizarre element in dreams. *Dreaming, 5*, 169–187.
Schimmack, U. (2008). The structure of subjective well-being. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being* (pp. 97–123). New York, NY: The Guilford Press.
Schredl, M. (2002). Questionnaires and diaries as research instruments in dream research: Methodological issues. *Dreaming, 12*, 17–26. http://dx.doi.org/10.1023/A:1013890421674.
Schredl, M., & Doll, E. (1998). Emotions in diary dreams. *Consciousness and Cognition, 7*, 634–646. http://dx.doi.org/10.1006/ccog.1988.0356.
Schulz, H., & Bes, F. W. (1998). The temporal distribution of awakenings during bed rest. *Journal of Sleep Research, 7*(Suppl. 2), 244.
Selvi, Y., Aydin, A., Gulec, M., Boysan, M., Besiroglu, L., Ozdermir, P. G., et al (2012). Comparison of dream anxiety and subjective sleep quality between chronotypes. *Sleep and Biological Rhythms, 10*, 14–22. http://dx.doi.org/10.1111/j.1479-8425.2011.00511.x.
Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*, 591–611. http://dx.doi.org/10.1093/biomet/52.3-4.591.
Snyder, F. (1970). The phenomenology of dreaming. In L. Madow & L. H. Snow (Eds.), *The psychodynamic implications of the physiological studies on dreams* (pp. 124–151). Springfield, IL: Charles S. Thomas.
St-Onge, M., Lortie-Lussier, M., Mercier, P., Grenier, J., & De Koninck, J. (2005). Emotions in the diary and REM dreams of young and late adulthood women and their relation to life satisfaction. *Dreaming, 15*, 116–128.
Strauch, I., & Meier, B. (1996). *In search of dreams: Results of experimental dream research*. New York: State University of New York.
Tuovila, S. (2005). *Kun on tunteet. Suomen kielen tunnesanojen semantiikka* (Doctoral thesis). Oulun Yliopisto, Oulu, Finland. <http://urn.fi/urn:isbn:9514278070>.

Valli, K., & Revonsuo, A. (2009). Sleep: Dreaming data and theory. In W. P. Banks (Ed.). *Encyclopedia of consciousness* (Vol. 2, pp. 341–355). Oxford, UK: Elsevier.

Valli, K., Strandholm, T., Sillanmäki, L., & Revonsuo, A. (2008). Dreams are more negative than real life: Implications for the function of dreaming. *Cognition & Emotion, 22*, 833–861. http://dx.doi.org/10.1080/02699930701541591.

Verdone, P. (1965). Temporal reference of manifest dream content. *Perceptual and Motor Skills, 20*, 1253–1268.

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the positive and negative affect schedule—Expanded form* (2nd ed.). Iowa City, IA: University of Iowa.

Wechsler, D. (2005). Osatestien esitys- ja pisteitysohjeet: Kuvien järjestäminen. In *WAIS-III käsikirja: Wechsler adult intelligence scale*, pp. 154–156). Helsinki: Psykologien Kustannus Oy.

Wechsler, D. (2007). Esitys- ja pisteitysohjeet: Looginen muisti I and Looginen muisti II. In *WMS-III käsikirja: Wechsler memory scale* (3rd ed., pp. 108–109 & pp. 128–130).

Weisz, R., & Foulkes, D. (1970). Home and laboratory dreams collected under uniform sampling conditions. *Psychophysiology, 6*, 588–596. http://dx.doi.org/10.1111/j.1469-8986.1970.tb02247.x.

Windt, J. M., & Noreika, V. (2011). How to integrate dreaming into a general theory of consciousness – A critical review of existing positions and suggestions for future research. *Consciousness and Cognition, 20*, 1091–1107. http://dx.doi.org.ezproxy.utu.fi:2048/10.1016/j.concog.2010.09.010.