

Measuring emotions in dreams: Effects of dream length and personality

Silke Aline Röver & Michael Schredl

Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Germany

Summary. There are different methods to measure emotions in dreams: self-ratings and external ratings by blind judges. Previous research indicated that the measurement technique has a strong effect on the findings, e.g., the ratio of negative to positive dreams is much higher for findings based on external ratings compared to findings based on self-ratings. In the present study, 1207 diary dreams reported by 413 participants were included. The participants rated the intensity of positive and negative dream emotions on two four-point scales; a blind external judge applied the same two emotion scales. The results confirmed previous findings that external judges underestimate emotional intensity in general – more so for positive emotions – but the correlations between self-ratings and external ratings are satisfactory. Higher mean word count, extraversion, and neuroticism are related to smaller differences between self-ratings and external ratings. Future studies should investigate how the instruction of reporting all experienced emotions explicitly, especially positive emotions, could influence the difference between self-ratings and external ratings.

Keywords: Dream emotions, self-rating, external ratings, personality

1. Introduction

Dream content analysis is a widely used tool in modern dream research (Schredl, 2010). A large number of scales and coding systems have been developed to measure different dream content characteristics, e.g., number of dream persons, presence of threats, friendly interaction, and aggression (Hall & Van de Castle, 1966; Valli & Revonsuo, 2000; Winget & Kramer, 1979). Research has also focused on reliability and validity indices of this paradigm (Schredl, 2010). Whereas reliability, most often determined as inter-rater reliability (congruence between the rating of two independent judges), is very often high (Domhoff, 1996; Schredl, Burchert, & Grabatin, 2004), validity studies did yield conflicting results, e.g., for measuring dream emotions and dream bizarreness using self-ratings as criteria for the dream content analytic scale (Schredl & Doll, 1998; Schredl & Erlacher, 2003).

As emotions are also central part of dreams (Kramer, 2007), the following review is focusing on this topic. There are different methods to measure dream emotions (Schredl & Doll, 1998): (1) self-rating by the dreamer himself/herself, (2) external rating by a blind external judge who is coding dream reports with regard to explicitly mentioned emotions as well as emotions derived from the dream action, and (3) external rating of explicitly mentioned emotions, e.g., using the classification system by Hall and Van de Castle (1966). The differences between the external ratings are illustrated by the following fictive dream example “I see a monster and

run away as fast as possible.” given by Schredl (2010): An external judge would code fear based on the dream action whereas no emotions would be coded according to Hall and Van de Castle (1966) as no explicit emotions were mentioned. I. e., the problem of validity does not concern the content analytic scale itself but the question whether the dream report which is the basis for the judge to code the dream according to the scales is including all information that is needed, as the primary aim is not to analyze what kind of emotions are reported but what kind of emotions are experiences within the dream (Schredl, 2010). Another source of error might be that situations in dreams evoke different emotions compared to the emotions experienced in a waking-life situation but Foulkes, Sullivan, Kerr, and Brown (1988) were able to demonstrate that this occurs very rarely, below 5% of the analyzed dream situations.

Schredl and Doll (1998) demonstrated that external ratings underestimate dream emotions compared to self-ratings using the same four-point scales measuring intensity of negative and positive emotions ranging from 0 = none to 3 = strong, particularly for positive emotions. If only explicitly mentioned emotions were considered, the underestimation was even stronger: According to the classification system by Hall and Van de Castle (1966) more than half of the dream reports did not include explicitly mentioned emotions even if self-ratings showed at least some form of emotion was experienced within the dream (see Table 1). Also depicted in Table 1, the ratio of positive and negative emotions was almost balanced for the self-ratings findings, whereas negative emotions outweighed positive emotions if external ratings were considered. This difference can be explained by the underestimation of positive emotions by the external judges (Schredl & Doll, 1998; Sikka, Valli, Virta, & Revonsuo, 2014).

Sikka et al. (2014) attributed the general underestimation of emotions to the fact that external judges can reliably assess only the emotions that were explicitly named in dream reports but not emotions that are implicit or not

Corresponding address:

Michael Schredl, Sleep laboratory, Central Institute of Mental Health, PO Box 12 21 20, 68072 Mannheim, Germany.
Email: Michael.Schredl@zi-mannheim.de

Submitted for publication: December 2016

Accepted for publication: April 2017

Table 1. Emotions in diary dreams (Schredl & Doll, 1998) (N = 133 dream reports)

Category	Self-ratings	External ratings	Hall & Van de Castle
No emotions	0.8%	13.5%	57.9%
Balanced emotions	12.0%	9.0%	6.8%
Predominantly negative emotions	50.4%	56.4%	26.3%
Predominantly positive emotions	36.8%	21.1%	9.0%

mentioned. Schredl and Doll (1998) hypothesized that negative emotions more probably affect the mood of the subsequent waking period than positive ones and, therefore, are more likely to be explicitly reported. Interestingly, the correlations between external rated and self-rated emotions is quite high: $r = .557$ (positive emotions) and $r = .669$ (negative emotions), despite the considerable underestimation by the external judge (Schredl & Doll, 1998).

The aim of the present study is to replicate previous studies concerning the underestimation of dream emotions by external judges. In addition, possible effects of dream length and personality-related factors on the difference between self-ratings and external ratings were examined in an exploratory manner.

2. Method

2.1. Participants

Overall, 425 participants who were recruited at the universities of Mannheim, Heidelberg and Landau took part in the present study. The total sample consisted of 361 women and 64 men, mainly psychology students, with a mean age of 23.40 ± 5.41 years (range: 16 to 61 years; two missing values). Participants were paid or received course credit. For the present analysis, participants who reported more than one dream per night or had missing values in the self-ratings were excluded from data processing (see procedure section). The resulting sample included 413 participants (350 women and 63 men) with a mean age of 23.19 ± 4.83 years (range: 16 to 55 years, two missing values). In total, they reported 1207 dreams with a mean word count of 137.63 ± 111.98 words. The mean number of reported dreams per participant was 2.92 ± 1.28 .

2.2. Dream diary

A standardized dream diary was given to the participants. They were asked to keep the diary for 14 consecutive days with the instruction to record all dreams of the previous night as completely as possible (to a maximum of 5 mornings with dream recall). After recording the dream(s), participants were asked to rate the intensity of positive and negative emotions in their dreams on two four-point rating scales (0 = none, 1 = mild, 2 = moderate, 3 = strong); the same scales were used by Schredl and Doll (1998).

2.3. Dream content analysis

For the external ratings, the same two four-point rating scales (0 = none, 1 = mild, 2 = moderate, 3 = strong) as used for measuring self-rated emotions. The external judge was instructed to consider any explicitly mentioned emotions as well as emotions that can be inferred from the dream action and to score the most intensive one if various positive or negative emotions appeared. Interrater reliabilities for these scales in previous studies were $r = .825$ for negative emotions and $r = .642$ for positive emotions (Schredl et al., 2004).

2.4. Personality measure

To assess the personality of the participants, the German version of the NEO PI-R (Ostendorf & Angleitner, 2004) was applied. It contains 240 five-point items (coded: 0–4) and measures interindividual personality differences on the main scales of neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. The internal consistencies of these scales are high ($r = 0.89–.92$) as reported by the test authors. Confirmatory multitrait–multimethod analyses replicated the findings for the English version (Ostendorf & Angleitner, 1994).

2.5. Procedure

Originally, the study entitled “Sleep, dreams, and personality” was carried out to investigate factors of home dream recall (Schredl, Wittmann, Ciric, & Götz, 2003). Participants completed different questionnaires investigating personality, sleep quality, stress, and creativity. Dream diaries were handed to the participants with oral instructions how to fill in the diary. To prevent loss of motivation, participants should only report their dreams on the first five mornings with successful dream recall and afterward just state if they recalled a dream or not. Morning reports that contained more than one dream per night were not included because in such cases it was not possible to match the subjective ratings to the corresponding dream reports of this night unambiguously. Dreams with missing self-ratings were also excluded.

The collected reports were typed to facilitate external rating and randomized to ensure blind rating. All information not describing the dream experience was deleted. The intensity of positive and negative dreams was estimated by an external judge (see dream content analysis section). Dreams were classified into four categories: (1) neutral dreams, if dreams included neither positive nor negative emotions, (2)

Table 2. Intensity of emotions in dreams measured by self-ratings and external ratings (N = 1207 dream reports)

Category	Positive dream emotions		Negative dream emotions	
	Self-ratings	External ratings	Self-ratings	External ratings
Strong	16.98%	2.65%	22.78%	6.05%
Moderate	24.19%	10.60%	26.18%	26.18%
Mild	29.91%	24.44%	27.34%	37.45%
None	28.91%	62.30%	23.70%	30.32%

Table 3. Emotions in dreams measured by self-ratings and external ratings (N = 1207 dream reports)

Category	Self-ratings	External ratings
No emotions	4.72%	17.15%
Balanced emotions	16.16%	10.60%
Predominantly negative emotions	44.41%	53.94%
Predominantly positive emotions	34.71%	18.31%

balanced dreams, if the intensity of positive and negative dream emotions was balanced, (3) predominantly positive dreams, if the intensity of positive dream emotions outweighed the intensity of negative dream emotions, and (4) predominantly negative dreams, if the intensity of negative dream emotions outweighed the intensity of positive dream emotions. For each participant who could report up to five dreams (two participants reported more than five dreams), average values for dream emotions were calculated. The difference between self- and external ratings was calculated by subtracting the value of the external rating from the value of the self-rating.

To control for potentially confounding variables, regression analyses included age, gender, number of dreams per person, and mean word count per report. Statistical comparisons for differences within the participants paired t-tests were computed. Statistical analyses were conducted using the SAS software for Windows 9.4.

3. Results

For all 1207 dreams, the external ratings and the self-ratings are depicted in Table 2. Self-ratings included moderate and strong emotions more often than external ratings. Table 3 shows that for external ratings the ratio of predomi-

Table 4. Intensity of positive and negative emotions in dreams (N = 413 participants)

Category	Self-ratings	External ratings	Effect size	t-test
Positive Emotions	1.26 ± 0.76	0.51 ± 0.51	d = 1.134	t = 23.1 p < .0001
Negative Emotions	1.51 ± 0.77	1.08 ± 0.64	d = 0.614	t = 12.4 p < .0001

nantly negative dreams to predominantly positive dreams was much higher compared to the ratio if self-ratings were analyzed. The correlations between self-ratings and external ratings were as followed: positive emotions ($r = .480$, $p < .0001$, $N = 1207$) and negative emotions ($r = .567$, $p < .0001$, $N = 1207$). With regard to self-ratings, dreams were rarely categorized as neutral compared to the classification based on external ratings.

Using the averages per participant, the external judge rated the dreams less intense than the dreamer (see Table 4). The difference between self-ratings and external ratings was higher for positive emotions (0.76 ± 0.67) than for negative emotions (0.43 ± 0.70) with an effect size of $d = 0.351$ ($t = 7.1$, $p < .0001$).

Table 5 summarizes the two regression analyses. Concerning negative emotions, a higher mean word count, as well as a higher degree of extraversion and neuroticism (marginally significant) is related to a smaller difference between self-ratings and external ratings, whereas a higher mean self-rating is associated with a higher difference between self-ratings and external ratings. In respect to positive emotions, the analysis only indicated that a higher mean self-rating is related to a higher difference between external ratings and self-ratings. Regarding all other variables, no significant effects were found.

Table 5. Effect of personality variables on the difference between self-ratings and external ratings of negative and positive emotions in dreams

Variable	Negative emotions		Positive emotions	
	Standardized estimate	Statistical test t =... p =...	Standardized estimate	Statistical test t =... p =...
Age	-.0650	-1.7 .0998	-.0083	-0.2 .8094
Gender	-.0700	-1.8 .0812	.0371	1.1 .2847
Mean word count	-.1299	-3.3 .0011	-.0204	-0.6 .5541
Number of dreams	-.0198	-0.5 .6052	-.0608	-1.9 .0658
Neuroticism	-.0898	-1.9 .0591	-.0223	-0.6 .5808
Extraversion	-.0978	-2.2 .0311	-.0446	-1.1 .2588
Openness to experience	-.0121	-0.3 .7705	-.0475	-1.3 .1873
Agreeableness	.0506	1.3 .1947	-.0398	-1.2 .2420
Conscientiousness	.0036	0.1 .9292	.0229	0.7 .5115
Mean self-rating	.6528	16.4 <.0001	.7738	22.2 <.0001

4. Discussion

In line with the studies of Schredl and Doll (1998) as well as Sikka et al. (2014) the present findings showed that by using the same scales external judges underestimate emotional intensity in general but especially for positive emotions. Although the underestimation is considerable, the correlation between self-ratings and external ratings is satisfactory. For negative emotions, a higher mean word count, extraversion, and neuroticism are related to smaller differences between self-ratings and external ratings.

From a methodological view point, the strong effect of emotional intensity levels on the difference between self-ratings and external ratings is very plausible as underestimations are only possible if dreams include moderate to strong emotions. Comparing the present findings to the study of Sikka et al. (2014) which used another method for measuring dream emotions, it seems unlikely that the type of scale plays an important role as the findings regarding underestimating dream emotions are comparable. The current study used only one external rater but previous studies have shown that for samples of students' dreams and different raters the interrater reliabilities were comparable (Schredl et al., 2004), so it is not necessary to obtain rating from a second judge every time. Although the two variables (differences between self-rating and external rating) were not normally distributed (due to the small number of distinct values), the shape of the distribution was symmetrically around the mean value. One would assume that the findings of this exploratory study have not been affected by this methodological issue. Nevertheless, in future studies this issue should be taken into consideration, e.g., by applying scales with much more than four categories.

The negative correlation between mean word count and the difference between self-ratings and external ratings (found for negative emotions) support the idea that the most probable explanation for this difference is the incomplete description of the dream experience since a more detailed dream report makes it easier for an external judge to rate dream emotions adequately. To test this hypothesis, it would be interesting to instruct participants to describe dream emotions as fully as possible in the dream report and test whether the difference between self-ratings and external ratings decrease. By comparison, the instruction in this study was to note the dream as completely as possible without indicating the necessity of mentioning emotions explicitly.

Concerning the fact that the underestimation of emotions is stronger with regard to positive emotions we hypothesized that negative emotions more probably affect the mood of the subsequent waking period and, thus, are more likely to be mentioned in the dream report. One possibility to test this assumption is to ask participants to rate the intensity of the dream emotions persisting upon awakening during the process of dream recording. If negative dream emotions are more persistent than positive emotions the above mentioned hypothesis would be supported.

An explanation for the finding that higher degrees of extraversion and neuroticism are related to smaller differences between self-ratings and external ratings of negative emotions, might be that extraverted persons, as well as persons with high neuroticism, tend to communicate negative emotions more explicitly. One might assume that the narrative style not only affect findings obtained from dream reports but also from reports of waking experiences. In a future

study, participants could be asked, for example, to retell the emotions of an emotional story in order to analyze whether there is a preference of telling negative emotions more often than positive emotions is also present in waking life and whether this preference is related to dream reporting. If personality measures are also included one could investigate whether extraversion and neuroticism are related to describing more explicitly mentioned negative emotions in these waking reports.

To summarize, the measurement method has a strong influence on the results regarding dream emotions, e.g., the ratio between positive and negative dreams. Future studies should investigate whether the instruction of reporting all experienced emotions explicitly, especially positive emotions, could influence the difference between self-ratings and external ratings. Another option would be to provide the judge with verbal records of dreams instead of written dream reports; this might reduce the underestimation since the judge is able to perceive the emotional state of the dreamer while s/he is reporting the dream emotions, especially if the dream is collected in the sleep laboratory directly upon awakening.

References

- Domhoff, G. W. (1996). *Finding meaning in dreams: a quantitative approach*. New York: Plenum Press.
- Foulkes, D., Sullivan, B., Kerr, N. H., & Brown, L. (1988). Appropriateness of dream feelings to dreamed situations. *Cognition and Emotion*, 2, 29-39.
- Hall, C. S., & Van de Castle, R. L. (1966). *The content analysis of dreams*. New York: Appleton-Century-Crofts.
- Kramer, M. (2007). *The dream experience: a systematic exploration*. New York: Routledge.
- Ostendorf, F., & Angleitner, A. (1994). A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology*, 44, 45-53.
- Ostendorf, F., & Angleitner, A. (2004). NEO-PI-R - NEO Persönlichkeitsinventar nach Costa und McCrae - Revidierte Fassung (Vol. 44). Göttingen: Hogrefe.
- Schredl, M. (2010). Dream content analysis: Basic principles. *International Journal of Dream Research*, 3, 65-73.
- Schredl, M., Burchert, N., & Grabatin, Y. (2004). The effect of training on interrater reliability in dream content analysis. *Sleep and Hypnosis*, 6, 139-144.
- Schredl, M., & Doll, E. (1998). Emotions in diary dreams. *Consciousness and Cognition*, 7, 634-646.
- Schredl, M., & Erlacher, D. (2003). The problem of dream content analysis validity as shown by a bizarreness scale. *Sleep and Hypnosis*, 5, 129-135.
- Schredl, M., Wittmann, L., Ciric, P., & Götz, S. (2003). Factors of home dream recall: a structural equation model. *Journal of Sleep Research*, 12, 133-141. doi:http://dx.doi.org/10.1046/j.1365-2869.2003.00344.x
- Sikka, P., Valli, K., Virta, T., & Revonsuo, A. (2014). I know how you felt last night, or do I? Self- and external ratings of emotions in REM sleep dreams. *Consciousness and Cognition: An International Journal*, 25, 51-66.
- Valli, K., & Revonsuo, A. (2000). Threatening events in dreams. *Psychologia*, 35, 485-492.
- Winget, C., & Kramer, M. (1979). *Dimensions of dreams*. Gainesville: University of Florida Press.