SOMNIEVE: ADVANCING DREAM RESEARCH THROUGH AN OPEN, MULTIMODAL DATASET

Valentina Elce[1], Bianca Pedreschi[1], Giorgia Bontempi[1], Michele Bellesi[2], Giacomo Handjaras[1] & Giulio Bernardi[1]

[1]MoMiLab Research Unit, IMT School for Advanced Studies Lucca, Italy;
[2]School of Bioscience and Veterinary Medicine, University of Camerino, Italy

Grant 91/20

Background: Dreaming provides a privileged access point to the study of sleep-dependent processes such as learning and memory consolidation, as well as to the broader investigation of human consciousness. Changes in the frequency or content of dreams are also linked to psychophysical health, often accompanying symptoms of psychiatric and neurological conditions. Yet, dream research is resource-intensive, underscoring the value of open, shared resources.

Aims: We created *Somnieve*, a multimodal open-source database collecting dream reports alongside demographic, psychometric, cognitive, and EEG measures from a representative sample of healthy adults. This resource enables testing hypotheses about the relation between dreams and individual state or trait variables and provides a reference for clinical studies. As an initial application, we used *Somnieve* to examine determinants of physiological dream content and recall frequency.

Methods: Analyses included 217 Italian native speakers (age: 18–70y). Participants wore an actigraph and reported their last dream experience each morning for 15 days. They also completed questionnaires and cognitive tasks. Reports were classified as contentful dreams, white dreams (contentless recall), or no dream. Natural language processing was applied to characterize the semantic features of the reports. Generalized Mixed-Effect Models assessed predictors of dream content and recall ($p<0.05$, FDR corrected).

Results: The probability of reporting a dream upon morning awakening was associated with attitude toward dreaming, proneness to mind wandering, and sleep patterns. The likelihood of recalling specific dream content was predicted by age and susceptibility to interference. Dream recall was also modulated by night-to-night variations in sleep and showed seasonal fluctuations. Dream content itself — and especially features such as visual qualities, vividness, and bizarreness — was influenced by several individual traits, again including attitude towards dreaming and proneness to mind wandering. Longitudinal analyses showed that major external events, such as the COVID-19 pandemic, may affect dream content and leave enduring traces.

Conclusions: Open, multimodal resources like Somnieve will be key to enhancing reproducibility in dream research and clarifying how state and trait factors shape dream recall and content in health, while also serving as a benchmark to detect clinical alterations.

Keywords: Sleep, Dream, Sleep mentation, Natural language processing

Publications:

Bertolini, L., Elce, V., Michalak, A., Widhoelzl, H. S., Bernardi, G., & Weeds, J. (2024). *Automatic annotation of dream report's emotional content with large language models*. In Proceedings of the 9th Workshop on Computational Linguistics and

Clinical Psychology (CLPsych 2024) (pp. 92–107). Association for Computational Linguistics. https://aclanthology.org/2024.clpsych-1.7.pdf

Elce, V., Bergamo, D., Bontempi, G., Pedreschi, B., Bellesi, M., Handjaras, G., & Bernardi, G. (2025). The individual determinants of morning dream recall. *Communications Psychology*, *3*(1), 25. https://doi.org/10.1038/s44271-025-00191-z

Elce, V., Handjaras, G., & Bernardi, G. (2021). The language of dreams: Application of linguistics-based approaches for the automated analysis of dream experiences. *Clocks & Sleep*, *3*(3), 495–514. https://doi.org/10.3390/clockssleep3030035

Wong, W., Herzog, R., Andrade, K. C., Andrillon, T., Barros de Araujo, D., Arnulf, I., Ataei, S., Avvenuti, G., Baird, B., Bellesi, M., Bergamo, D., Bernardi, G., Blagrove, M., Decat, N., Demirel, Ç., Dresler, M., Eichenlaub, J.-B., Elce, V., Gais, S., ... Tsuchiya, N. (2025). A dream EEG and mentation database. *Nature Communications*, *16*, 7495. https://doi.org/10.1038/s41467-025-61945-1

E-mail contact: giulio.bernardi@imtlucca.it