Cross-Cultural Analysis of Dream Reports and Mental Health

Xiangyu Wang, Sizhe Chen, Zhanylai Turatkhan, & Paul McKee, Ph.D.

Duke MIDS
Master in Interdisciplinary Data Science

Introduction and Purpose

This study investigates the relationship between **dream structure and mental health** across cultures. Using a **cross-cultural survey (**U.S. and Argentina), we analyze dream reports alongside measures for depression (PHQ-9) and anxiety (GAD-7).

To determine how computational **linguistic features** - such as lexical diversity, syntactic complexity, and graph connectivity - alongside self-reported nightmare frequency and distress, can predict clinical mental health scores.

Methods

Data Collection: The study analyzes existing data from **1,975 participants** across the **United States and Argentina**. Participants provided free-text dream reports and completed standardized mental health assessments, specifically the **PHQ-9** (depression) and **GAD-7** (anxiety) scales.

Text Analysis (NLP): Researchers apply **Natural Language Processing (NLP)** to the dream reports to extract structural and emotional features. This includes measuring:

- **Sentiment**: The emotional tone of the dream.
- **Structure**: Lexical diversity (vocabulary range) and syntactic complexity.
- **Connectivity**: Using graph theory to measure how ideas are linked and the overall coherence of the narrative.

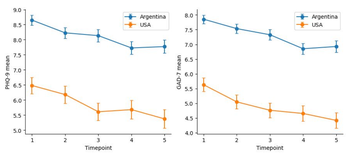
Statistical Modeling: The team uses **Linear Mixed-Effects Models** and **Supervised Machine Learning** (such as logistic regression) to determine if **specific dream features can predict clinical levels of depression or anxiety.**

Nightmare Classification: Utilized the **Empath** library to identify the presence and frequency of fear, violence, and threat-related lexical terms.

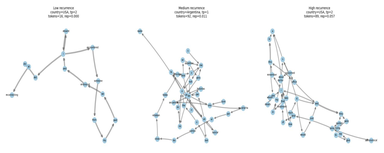
Nightmare Intensity: Applied **multilingual zero-shot** semantic alignment to calculate a proximity score between the dream's narrative meaning and core constructs of fear or distress.

Cross-Cultural Comparison: To account for cultural differences, "Country" is treated as a fixed effect in the models. This allows researchers to see if the relationship between dreams and mental health remains consistent or changes between the two nations.

Mental Metrics Across Culture and Time

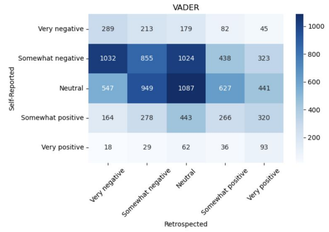


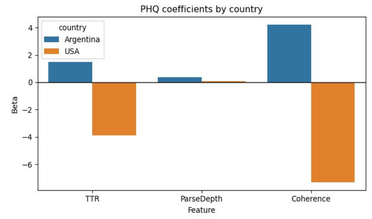
Visualized Lexical Structure of Dream Reports



Results

- People tend to recall dreams in the same emotional direction they report. This is significant for both methods on the primary metric (non-neutral same-direction rate > 50%)
- Structural features extracted : TTR (lexical diversity), ParseDepth (syntactic complexity), and Coherence (adjacent-sentence overlap).
- These structural features showed weak predictive signal for PHQ-9 and GAD-7 (tiny (R^2), non-significant coefficients, near-zero/negative CV (R^2), and minimal gain over baseline covariates).





- A one-unit increase in measured nightmare intensity is associated with a ~0.6 point increase in clinical depression and anxiety scores.
- The "nightmare–symptom slope" is statistically identical in both the U.S. and Argentina. There is no evidence that culture moderates how nightmares reflect mental health.
- Nightmare "Intensity" (meaning) consistently outperformed "Content" (word counts) as a predictor of distress. This association was cross-culturally stable (p > .4), suggesting a universal psychological mechanism linking nightmare intensity to mental health

Mixed-Effects Regression of Nightmare Language Predictors on Mental Health Symptoms

Predictor (from Dream Text)	Country	Outcome (Mental Health)	Effect (β)	Significance (p-value)
Nightmare Intensity	U.S.	PHQ-9 (Depression)	~0.60	P < .001 (Robust)
Nightmare Intensity	Argentina	PHQ-9 (Depression)	~0.60	P < .001 (Robust)
Nightmare Content	Both	GAD-7 (Anxiety)	Moderate	Significant
Intensity X Country	Interaction	PHQ-9 / GAD-7	N/A	**Non-Significant**

Conclusions

The research explores how dreams relate to mental health:

- **Emotional Connection:** Dream reports are a strong reflection of a person's actual emotional state. If you are feeling a certain mood in your waking life, it is very likely to show up consistently in your dreams over time.
- **Quality over Structure:** When analyzing dreams, **what** you feel matters more than **how** you tell the story. The study found that "structural" features do not really help predict the severity of mental health issues.
- **The "Nightmare" Pattern:** Nightmare intensity is a reliable "behavioral indicator" for digital population-level mental health monitoring.
- **Cross-Cultural Universality:** The nightmare–symptom relationship is statistically consistent across the U.S. and Argentina. Non-significant interaction terms suggest a universal threat-sensitivity mechanism .
- **Key Takeaway for Future Research:** Structural features show limited value for distress, while affective content is more sensitive. Future work should use LLMs to identify specific semantic themes (e.g., helplessness, grief, social isolation).

Acknowledgments

This research was conducted as part of a **Duke MIDS Capstone project** in collaboration with the **Department of Psychology & Neuroscience**. We sincerely thank the participants for their voluntary contributions to the longitudinal surveys. This work was supported by Duke University and approved by the Campus IRB (Protocol #2024-0062).

This project adheres to **Open Science principles**, including pre-registration of hypotheses, the use of synthetic data for model rehearsal, and a commitment to reproducible data workflows using open-source libraries

References

Primary Project References

1. Gallo et al. (2023): Used to refine methodological choices and ensure alignment with Open Science standards.
2. Mota et al. (2014): how dream reports can reveal clinically relevant structural features of thought, specifically noting that bipolar participants showed significantly lower connectivity in dream reports.
3. Herrero et al. (2025) Preprint: Part of the relevant papers reviewed to refine the project's analytical approach.

Clinical & Supplemental Sources

1. PHQ-9 & GAD-7: Standardized mental health questionnaires used to assess depression and anxiety severity.