Consciousness and Cognition 106 (2022) 103428

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Mapping dreams in a computational space: A phrase-level model for analyzing *Fight/Flight* and other typical situations in dream reports

Maja Gutman Music [a,*], Pavan Holur [b], Kelly Bulkeley [c]

[a] *Department of Anthropology, University of California, Los Angeles, United States*
[b] *Department of Electrical and Computer Engineering, University of California, Los Angeles, United States*
[c] *The Sleep and Dream Database, Portland, OR, United States*

ARTICLE INFO

Keywords:
Typical situation
Pattern quantification
Dreams
Natural Language Processing (NLP)
Network analysis
Semantic structure
Representation learning
Fight/Flight
Flying/Falling

ABSTRACT

This article demonstrates that an automated system of linguistic analysis can be developed – the *Oneirograph* – to analyze large collections of dreams and computationally map their contents in terms of typical situations involving an interplay of characters, activities, and settings. Focusing the analysis first on the twin situations of *fighting* and *fleeing*, the results provide densely detailed empirical evidence of the underlying semantic structures of typical dreams. The results also indicate that the *Oneirograph* analytic system can be applied to other typical dream situations as well (e.g., *flying, falling*), each of which can be computationally mapped in terms of a distinctive constellation of characters, activities, and settings.

1. Introduction

> "As in a dream a man is not able to follow one who runs From him, nor can the runner escape, nor the other pursue him, So he [Achilles] could not run him down in his speed, nor the other [Hector] get clear." Homer, *The Iliad,* XXII, 199–201

A long and widely-held observation in dream research is that typical dream themes can be identified across cultures and populations and statistically measured for their occurrence (Calkins, 1893; Seligman, 1923; Eggan, 1952; Griffith, Miyagi, & Tago, 1958). For example, an extended study of common topics in dreams conducted by Nielsen et al. in 2003 among Canadian University students (N = 1,348) showed that the most reported themes were: being chased or pursued, sexual experiences, falling, arriving too late, being on the verge of falling, trying again and again, and flying or soaring through the air (Nielsen et al., 2003). However, while there seems to be an emerging empirical consensus on the most common topics in human dreaming (Schredl & Hofmann, 2003a; Mathes, Schredl, & Göritz, 2014; Maggiolini, Di Lorenzo, Falotico, & Morelli, 2020; Yin, Shen, He, Wei, & Cao, 2013; Yu, 2008), less is known about the distinctive semantic structure of each typical dream theme, or even if they have any semantic structure at all. It is possible that so-called typical dreams have no such deeper structure, and they are united only by superficial linguistic labels (e.g., "chasing", "flying") that mean different things to different people. Perhaps what makes a dream "typical" is nothing more than the presence of one or two particular words, and everything else in the dream's content is random and variable. If this were true, the concept of a typical

* Corresponding author.
E-mail address: gutman.maja@gmail.com (M. Gutman Music).

https://doi.org/10.1016/j.concog.2022.103428
Received 10 March 2022; Received in revised form 27 September 2022; Accepted 11 October 2022
Available online 27 October 2022
1053-8100/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

dream would still remain, but its substance and relevance to dream research would be greatly diminished.

This article introduces a new automated system to facilitate dream analysis, *Oneirograph*, to help determine the presence or absence of underlying patterns of semantic structure in typical dreams. To identify these patterns using traditional methods of human reading and manual coding is possible with a small number of dreams, but larger data sets are necessary when studying consensus effects such as consistency in themes, sentiments, recurring actors and their inter-activity. Emerging tools of automated linguistic analysis can work well with large data sets, and the *Oneirograph* was designed with the specific goal of enabling the large-scale study of recurrent meaningful patterns — what we here refer to as typical situations — in dreams. Drawing upon the emergence of novel Natural Language Processing (NLP) architectures including the Transformer (Vaswani et al., 2017), and its adeptness at Transfer Learning (Azunre, 2021), the *Oneirograph* is able to determine with a high level of precision the presence or absence of a distinctive set of characters, settings, and actions that define the semantic contents of a typical dream situation.

Psychologists and anthropologists have long noted the occurrence of typical situations in dreams, but their clinical and ethnographic observations have been difficult to validate with systematic empirical analyses. There is a gap on this point between what dream researchers have observed and what they can convincingly demonstrate, a gap the *Oneirograph* method described in this article seeks to bridge. Two primary hypotheses will be proposed and tested:

Hypotheses. **H1:** Each typical dream situation is constituted by a semantic structure involving a distinctive set of characters, settings, and actions whose relations can be mapped in computational space.

H2: An automated system of analysis can be developed with the general capacity to identify and computationally map the semantic structures of a variety of typical dream situations.

The analysis here will focus on identifying the distinctive semantic structures in the typical dream situations of *fighting*, *fleeing*, *flying*, and *falling*. The results will be discussed in relation to possible applications of the *Oneirograph* for studying these and other typical dream situations in individuals and groups. The *Oneirograph's* findings and availability for use by other researchers will also be discussed in relation to contemporary theories of dreaming, including the continuity hypothesis (Hall & Van de Castle, 1966; Domhoff, 2003; Schredl, 2010) and the immersive spatio-temporal hallucination model (Windt, 2010).

2. Related works

There have been many studies on the lexico-statistical analysis of dreams and other major attempts where machine learning tools were employed to detect semantic or non-semantic structures in dream reports. For example, Schwartz (Schwartz, 2004) provided a valid statistical approach to pattern-seeking in dream content by utilizing the Factorial Correspondence Analysis (CoA) and clustering method in 1,713 dream reports. Five semantically discrete clusters were identified, suggesting that the word pattern corresponds to cognitive categories, such as fear, social events, and locomotion. Mota et al. (Mota, Furtado, Maia, Copelli, & Ribeiro, 2014) applied the speech graph attributes (SGA) method to study the non-semantic structure of temporal connections between consecutive words in the speech of psychotic patients and non-psychotic control groups as they reported waking and dream contents. One of many promising outcomes of this study provided empirical evidence of similarity between dreaming and psychosis, along with a set of behavioral markers that could contribute to the more objective assessment of psychiatric diagnosis for schizophrenia and bipolar disorders. Similarly, Martin et al. (Martin et al., 2020) used a non-semantic graph analysis to investigate the structural differences between REM and non-REM dream reports collected in the sleep laboratory. The study showed that the REM reports were typically longer than non-REM and characterized by higher structural connectedness than non-REM reports. These findings are consistent with previous studies, which concluded that non-REM dreams are less visual and memorable, shorter, and more conceptual than REM dreams (Bulkeley, 2016).

Another substantial achievement at the intersection of extensive dream data and NLP tools can be seen in the recent work by Fogli et al. (Fogli, Maria Aiello, & Quercia, 2020). In this study, the researchers designed the processing tool that automatically scores dream reports using the Hall and Van de Castle coding system and two external knowledge bases of Wikidata and Wordnet. Based on 24,000 dream reports captured from the extensive online repository DreamBank[1], the analysis provided supporting empirical evidence for the continuity hypothesis, which posits that patterns of dream content reflect the dreamer's most important waking life concerns, interests, and activities (Bulkeley, 2016; Domhoff & Schneider, 2008; Schredl & Hofmann, 2003b; Schredl, 2010). Moreover, the continuity hypothesis suggests that the content of dreams, when studied longitudinally (for example, in dream series or dream diaries of an individual), can reveal references to various themes that play crucial roles over the course of the waking life of a subject (Bulkeley, 2016).

In addition, Pesonen et al. (Pesonen et al., 2020) investigated over 4,200 sleep and dream reports to examine the effects on dream content during the lockdown in Finland. Unsupervised computational network and cluster analysis of word association methods were used in the study, which provided yet another supporting evidence of the continuity hypothesis, with over 55% of clusters reflecting distressing content related to the pandemic. Finally, Elce et al. (Elce, Handjaras, & Bernardi, 2021) observe that graph analysis in dream reports seems to be particularly useful in studying structural differences between clinical groups and healthy control groups, while dictionary-based methods compare data with predefined lexical categories to account for the occurrence frequency of unique words. Elce and colleagues also describe the distributional semantics technique, which is based on word embedding and requires an

[1] Link to DreamBank: dreambank.net.

extensive corpus of dream reports to produce reliable and stable results.

It is worth emphasizing that this review study by Elce et al. briefly touches upon the latest techniques, "such as the recently introduced machine learning approach named 'Bidirectional Encoder Representations from Transformers' (BERT [79]), which may be able to analyze complicated syntactic constructions efficiently. Indeed, BERT allows for a sophisticated analysis of texts by encoding sentence-level properties of linguistic units and exploring the hierarchical relationship between different words within the same syntactic context." This mention is significant for the present study, as our conceptual framework for modeling the typical situations in dreams builds on the BERT model. Thus the Elce et al. review underscores the novelty and potential value of the present article's methodological approach.

The largest corpus of dream content is in the form of *dream reports*: text segments written by the dreamer that describe the encountered situation during the activity of dreaming. As a result, the primary mode of communicating dream experiences is written language, which makes any extraction of structured information from such data amenable to study using the methods of Natural Language Processing (NLP). In general, many of these methods involve large attention-based Deep Learning (DL) models (Devlin, Chang, Lee, & Toutanova, 2019; Vaswani et al., 2017; Liu et al., 2019) – spurred by recent breakthroughs in NLP – and their relative success at classical language classification and generation tasks – information extraction tasks in their own right – such as Semantic Role Labeling (SRL) (Palmer, Gildea, & Xue, 2010) and Question Answering (QA) (Rajpurkar, Zhang, Lopyrev, & Liang, 2016). Attention-based DL models are highly effective on these tasks because they have been *pre-trained* on internet-scale language data. Pre-training is the process of allowing the model to learn language patterns from extensive unlabeled data streams so that it may be highly effective when oriented toward a smaller-scale fine-tuning task. The success of such pre-trained models is well-documented (Raffel et al., 2020; Brown et al., 2020). In this work, we employ two such pre-trained and then fine-tuned models to extract situational parts from dream reports: (a) TransformerQA (Liu et al., 2019) in order to extract information in the form of answers to questions that describe a situation; and (b) Sentence-Transformer (Reimers & Gurevych, 2019) in order to effectively aggregate answers in a numerical vector space – see Sections 3.4 and 3.5 respectively.

Extraction of situational parts from text using a method of question-driven information extraction from text as we propose here, is an emerging field of interest. A recent work (He, Lewis, & Zettlemoyer, 2015) attempted to generate answers to questions from a template of possible questions as a proxy task for SRL span classification. Another work (Buck et al., 2017) extended this template-of-questions framework to drive a sequential question–answer-question game, where the collection of answers generated at the end of a trajectory of sequential questions provides a rich extraction of information chunks from an input context. A parallel stream of information extraction models finds pairwise entity relationships (Mausam, 2016): Using the features presented by dependency parse trees, larger hard-coded information databases such as YAGO (Suchanek, Kasneci, & Weikum, 2007) and FrameNet (Ruppenhofer, Ellsworth, Schwarzer-Petruck, Johnson, & Scheffczyk, 2016) tags, information is structured as a set of relationship tuples (subject phrase, verb phrase, object phrase). The nature of dream reports (e.g., mostly reported in the first-person, with poor punctuation and idiosyncratic word usage) and inadequate control provided by these existing models to specify the *information-of-interest* that is extracted from the text, limits its direct use for identifying situational parts.

3. Methodology

3.1. Selected situations, definitions, and sources of data

On a conceptual level, the typical situation can be defined as a set of factors that constitute a common theme in dreams, such as *flying*, *falling*, *escaping*, or *fighting*. Note that their core activity defines the labeling of these situations. This activity, however, has many accompanying constituents – other actants, settings, and even trajectories – that appear to be variable. For example, one can be escaping from a snake or missiles or fighting a virus or a fictional character in a video game. The first example can be characterized as a typical *fleeing* dream, whereas the second describes a typical *fighting* dream. While the main activity is fixed, the corresponding constituents (snakes, missiles, viruses, fictional characters) can vary. We analyzed these corresponding constituents to understand the semantic structure for typical situations.

Two pairs of typical dream situations are examined here: (i) the primary situation analyzed in this paper, *Fight/Flight*, and (ii) the secondary situations analyzed more briefly, *Flying (levitating)/Falling*. It can be seen that our selected situations have been configured as binary oppositions. By following this line of reasoning in our study, we sought to accomplish several goals: preserving the overarching concept of each situation (it is difficult to think of flying without the gravitational inevitability – falling), or any conflict situation without two options (like fleeing or fighting); being able to represent the situations in 2-dimensional and in 3-dimensional space; and juxtaposing the opposites to measure and interpret their differences and similarities.

We defined the situations as follows: (i) *Flight/Fight* is a situation in which the dreaming subject is physically or non-physically (i.e. verbally) confronted with the real or imaginary threat posed by external forces (either fictional or non-fictional characters, natural or cosmic events, or other fictional or non-fictional entities from waking or imaginary world). Another important feature of this situation is the outcome which consists of two possible options: fleeing (escaping from the threat) and fighting (an immediate physical or non-physical confrontation with the object of a threat), or both. (ii) *Flying/Falling* situations in dreams are defined as any first-person experience that includes flying, levitating and hovering in closed or open spaces at any altitude or speed. Falling is defined as a

situation where the dreamer is experiencing a sudden or gradual fall from any height due to gravity.

Sources of Dream Reports The dream reports used in the analysis were provided by the Sleep and Dream Database (SDDb)[2]. The SDDb is an online, open-access archive of dream reports and related information about sleep and dreams, gathered from a variety of sources: individual dream journals, demographic surveys, collections from other researchers, and historical and anthropological texts. Inspired by the Dreambank of G. William Domhoff and Adam Schneider, the SDDb includes more than 35,000 dream reports from more than 10,000 different people. All of the dream reports available in the SDDb have been gathered by researchers using systematic, transparent methods, from participants who have consented to their dreams serving as a source of academic study. The dreamers providing these reports are demographically diverse in terms of gender, age, race/ethnicity, education, and other frequently employed variables. However, the current collection of dreams in the SDDb does not provide a representative sample of all human dreaming. Rather, it offers a relatively large source of data, large enough to yield meaningful results when studied by current technologies of linguistic analysis.

3.2. *Hierarchical Dream Network: A generic template for situations in dreams*

Hypothesis 1 (see Section 1) attempts to show that dream reports together comprise a *consensus summary structure* while Hypotheses 2 rely on this structure to identify parts of specific situations. In order to validate Hypothesis 1 and to facilitate Hypotheses 2, we define a general 2-level template for situations in dreams and measure the accuracy-of-fit to this template from text-based dream reports: This template is an accommodating structure motivated by recent work – discussed in Section 2.

- **First Level - Roles:** Each situation is comprised of multiple distinct roles that interact with one other to create a situation. A role describes the broad purpose of a character, entity or more generally a text phrase, in supporting a situation in a dream report. Computationally, we model a role as being attached to a text phrase if the phrase is the answer to a motivating question that describes the role's intent. For example, while describing the situation of *Fighting*, a dream report contributor often includes text segments to satisfy the role of *antagonist* by answer the implicit question "Whom am I fighting against?". Other reports may include segments that specify the *reason*: answers to the question "Why am I fighting?"
- **Second Level - Contextual Groups:** A single role can be represented by different groups of semantically similar phrases. For example, the role motivated by a question "Whom am I fighting against?" can include entities as diverse as "demons", "dragons", "mother" and "sister". Broadly, these entities can be grouped into {familial} entities: "mother" and "sister", and {mythical} entities: "demons" and "dragons". Such contextual groups serve to highlight the semantic similarity present *across* dream reports and facilitates the aggregation of similar entities into semantic categories across the dream corpus.

In the following sections, we discuss the *automated* pipeline of interlocking Natural Language Processing (NLP) tools that together help extract situational information from dream reports and that fit the *Oneirograph* template. Success in modeling dreams in this way is tantamount to creating interpretable consensus representations of dream reports that provide a stable structured view of the common situations that are (a) originally obfuscated within the dream report text; and (b) distributed across multiple reports. The broader utility of these network representations in the field of Dream Studies, and potentially psychotherapeutic environments, are discussed in Section 7.

The computational pipeline can be decomposed into 4 sequential stages: (1) **Role Identification:** Identifying the relevant roles in a situation; (2) **Information Extraction:** Extracting the constituent actors, entities or other general text spans that satisfy these roles; (3) **Actor Group Role-Tagging:** Creating contextual groups of entities and assigning roles; and (4) **Situation Estimation:** Connecting contextual groups via co-occurrence of roles in dream reports to create a network representation that is situation-specific. A schematic overview of the pipeline is presented in Fig. 1. The input into the pipeline is a large database of dream reports, the output is a network of interconneted actor groups that we term a *Situation Network*. *Henceforth, all the computational parts are described per situation.*

Using the proposed *Oneirograph* methodology and situation template, we obtain results that include: (a) A qualitative analysis of the obtained network that reveals situation-conditioned contextual groups and the roles within the situation that they embody; (b) Metrics that specify situations (as a fit in the proposed template) generated from dream reports to a random baseline. This later point demonstrates our raw dataset of dream reports abides by a necessary structure in order to facilitate the modeling of situations in dreams.

3.3. *Module 1: Role identification*

With our proposed computational definition of a role – the answers to questions that describe a role's intent – it is first paramount to define these questions. This is a challenging task, because different situations emphasize different roles. Our method starts with the classic 5Ws1H questions: *Who*,*What*, *Where*, *Why*, *When*, and *How*? Then we refine these questions so they apply more directly to our selected situations. For example, in a *Fleeing* situation, a dream may provide more information that can answer the question "Where is the [dreamer] fleeing to?" better than "When was the [dreamer] fleeing?" since the perception of time is often distorted or omitted in dreams. To put the method into practice, we aggregated a seed set of these questions *per situation* in the following way: (1) Three independent English-speakers were asked to come up with distinct questions that they would like answered given a situation. (2) Any

[2] https://sleepanddreamdatabase.org/.

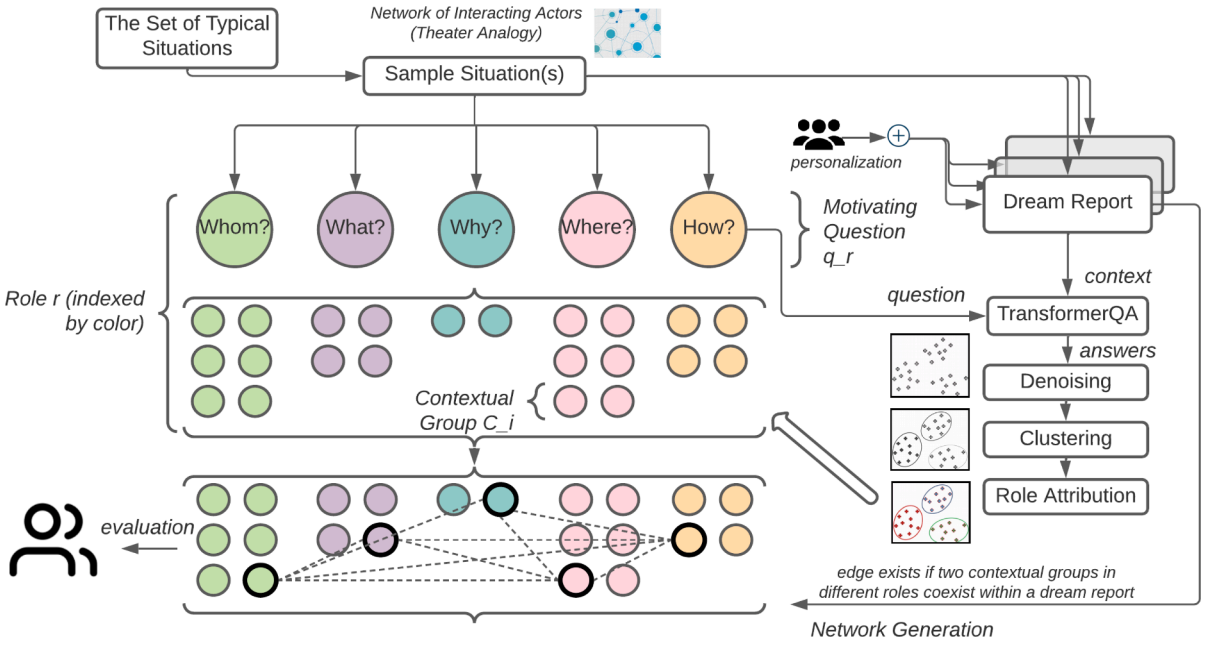

Fig. 1. An Outline of the *Oneirograph* pipeline and the hidden generative process involved in the creation of dream content: The pipeline involves (i) Selecting a set of situations to be featured in a dream (and then later recapped in a dream report); (ii) Individuals recounting dream experiences and the embedded situations using variable language constructs in dream reports; (iii) *Oneirograph* reconstructing the latent situations in the dream reports using state-of-the-art attention-based Question–Answering (QA) model on the resultant text samples.

question that had at least two votes was considered *eligible* and each semantically distinct question q_r is affiliated to a role r. From each question q_r, we create a larger set of questions in order to increase the questions' relevance to dream data:

- q_r is converted to the first person: Since ⩾90% of dreams are written in the first-person, the questions must reflect the same perspective.
- A preposition that occurs in q_r is compared to other prepositions from a popular preposition database [3] and questions with semantically similar prepositions are manually augmented to the existing question. For example, "Who am I fighting?" is semantically very similar to "Who am I fighting against?"
- Dream reports are written in variable tense. Most write in past continuous ("I was going to the store in my car…"), others write in present continuous ("I am standing in a graveyard in the middle of the night…") and few write in absolute past tense ("I walked home and saw a ghost"). As a result, we extend each question to incorporate these variable tenses.

The larger augmented set of questions generated for each role r is Q_r. The supplementary data file (Gutman, HOLUR, & Bulkeley, 2022) [4] contains the table of (Situation, Role, Questions) tuples assembled in this section.

3.4. Module 2: Information extraction

The initial goal of creating the question sets Q_r per role r was to facilitate the extraction of the answers to these questions, the fundamental units that support a situation and constitute roles. In our setting, answers are *spans* (a contiguous window of text) within the dream report and we seek to identify the index of the start and end character in the text that constitutes the span. Note that the question set Q_r was aggregated without analyzing individual dream reports, so the Question–Answering (QA) model must be robust enough to answer indirect questions given a single dream post.

For example a QA model posed with a dream report (context): "I escaped from a dragon and got home safe." and a question: "What was I fleeing from?" – a typical question in the situation of *Fleeing* – must yield the answer span "a dragon" *despite there not being a perfect match in the language used to express the idea of "running away" in the context ("escaped from") and question ("fleeing from") pair.*

Such a model then requires a contextual understanding, one that has embedded within it the semantic similarities (that includes synonyms) between phrases in the English language, a must-have when the inputs do not employ identical vocabularies and sentence structure.

[3] https://inspirassion.com/en/prep/.
[4] Data and Code Repository: https://osf.io/e9jnv/.

The latest advancements in Natural Language Processing (NLP) are capable of such generalization, as discussed in the Related Works (Section 2): In particular, attention models (BERT, Transformers and their variants) have demonstrated the ability to *infer* semantic similarity between phrases of text by relying on word-piece-level co-occurrence. The pre-trained RoBERTa-large QA model (TransformerQA) (Liu et al., 2019) in particular, a question–answering model built atop the attention backbone, is well-suited to our task: Its performance on the SQuAD dataset (Rajpurkar et al., 2016) is near-state-of-the-art and qualitatively appears to provide generalization across context-question pair inputs at the semantic level.

For a context (dream) $d \in D$, and for every question that applies to a particular role $q \in Q_r$, we query an answer a_r: an inference through TransformerQA. Note that the applied model returns an answer as well as its confidence of the answer. If the answer does not exist, the model is expected to output nothing. A summary of the information extraction pipeline is provided in Algorithm 1. Note that answers are also indexed by the role: i.e. a question q_r generates an answer a_r.

Algorithm 1. Information retrieval from dreams for a specific situation

Data: The set of dreams, D, and
question sets per role, $[Q_{r_1}, Q_{r_2}, \ldots, Q_{r_N}]$
Result: The set of answers, A
for $d \in D$ **do**
  for $Q_r \in [Q_{r_1}, Q_{r_2}, \ldots, Q_{r_N}]$ **do**
    for $q \in Q_r$ **do**
      $a_r, conf \leftarrow \texttt{TransformerQA}(d; q)$;
      if $conf \geq \tau$ **then**
        append a_r to A;
      end
    end
  end
end

Observed Limitations of Algorithm 1: Since not all dreams have the answers to every question for every role posed in a situation, Algorithm 1 has the potential to return answers to questions that are incorrect – this despite the output confidence metrics, high F1-performance on SQuAD and the pretraining of the model to not return answers when they do not exist. We observe that these answers can be wrong in 2 significant ways:

- **Answers are nonsensical:** An answer a_r may not be interpretable. This is especially common when the dream report is poorly punctuated and/or the answer to a question does not exist in the context. Then, TransformerQA is coaxed into returning a false positive.
- **Answers are not affiliated to the relevant role:** Despite an answer span a_r being interpretable on its own, it may not be the correct answer to the intended role *r*. This confusion is especially apparent in *Who* and *Where* roles: TransformerQA appears to regularly confuse the answers generated in response to "*Who* am I fleeing from?", and "*What* am I fleeing from?"

In order to address these limitations, we devise a simple means to "de-noise" the results by using the consensus information *across dream reports*. This is based on two observations: (a) Nonsensical spans do not repeat as frequently across dream reports as interpretable ones do; and (b) Similar answers across dream reports would more likely affiliate to the intended role in aggregate despite any confusion per dream report. Such aggregation would also help contextualize individual dream reports within the broader community of dream reports. The next sections describe the method to generate this consensus structure.

3.5. Module 3: Contextual group generation and labeling

Answers that are semantically similar are likely to belong to the same role. We exploit this fact and group the resultant answers – *we call these groups contextual groups* – from Algorithm 1 using *clustering*. We will provide a way to use these groups to (a) remove nonsensical answers; and (b) improve the reliability of answer-role mapping.

Step 1 - Constructing numerical representations: A first-order requirement for clustering is to define a semantically-rich vector space that has embedded within it, the phrases that need to be clustered. Grouping phrases in such a space by using a reliable distance measure would facilitate the formation of a set of high-quality contextual groups. We encode each answer a_r into a D-dimensional ($D = 768$) pre-trained BERT representation (Reimers & Gurevych, 2019). This encoding has been widely used in comparable downstream applications (Grootendorst, 2020) and has been shown exhaustively to consist of substantial representative information of the original phrase.

Validation: To verify the richness of these embeddings and their use on text specific to dream reports, we encode sentences from dreams corresponding to the *Fight* and *Fight* situations [5] and visualize (in 1- and 3-dimensions) the semantic separation between the two opposing situations in the vector space (see Fig. 2). The separation between the *Fight* and *Flight* embeddings indicated in these figures qualitatively demonstrates the effectiveness of the proposed library of numerical encodings to represent semantic information in dream reports.

Algorithm 2. Identifying actor groups within the set of answers

Data: A
Result: The clusters of answers, C_A
for $a \in A$ **do**
  append BERT(a) to R_A
end
$\hat{R}_A \leftarrow \texttt{UMAP}(R_A)$;
$C_A \leftarrow \texttt{HDBSCAN}(\hat{R}_A)$;

Step 2 - Dimensionality Reduction and Clustering: In order to construct contextual groups, we cluster these BERT representations in a UMAP (McInnes, Healy, & Melville, 2020)-dimensionality reduced 5-dimensional space using HDBSCAN (McInnes & Healy, 2017). The end result is a set of clusters C_A derived from the set of answers A where each cluster $C_i \in C_A$ comprises a set of answers that are semantically similar to one another (see Algorithm 2). This clustering framework is widely applied, most notably in BERTopic (Grootendorst, 2020), a state-of-the-art library for topic modeling.

Step 3 - Associate role to contextual group: So far, grouping similar answers has resulted in denoising – answers that are not semantically similar to others or are ill-structured are filtered out by HDBSCAN into the *noise* cluster marked -1. Each contextual group represents a group of answers that are similar and many times, we are able to qualitatively infer the *theme* of a group. For example, one cluster hosting the set of answers, {"sister","brother"}, indicates a broader theme of {siblings}.

Algorithm 3. Assigning roles to each actor group using the KS-test

Data: Set of answers, A, Contextual group C_i
Result: Role, r^*, and Confidence (p-value) p
/* Estimate Prior Distribution */
for *answer* $a_r \in A$ **do**
  $\pi(r) += 1$
end
Normalize π;
;
/* Estimate Posterior Distribution */
Initialize $f_R(r) = 0 \quad \forall r \in R$;
for *answer* $a_r \in C_i$ **do**
  $f_R(r) += 1$
end
Normalize f_R;
;
/* Measure Significance */
Confidence, $p = \texttt{KS-test}(f_R, \pi)$;
Role, $r^* = \arg\max_r \frac{f_R(r) - \pi(r)}{\pi(r)}$;

We improve the reliability of the answer-role pairing by associating a role to each contextual group. Recall that each answer a_r was generated in association to a role. For a set of answers we use a relative majority vote on the available roles – with respect to a non-uniform prior to label each cluster – a confidence measure is obtained via the Kolmogorov-Smirnov test (KS-test). Details of this labeling process are provided in Algorithm 3. *This method mitigates role misattribution by TransformerQA: A misclassified answer from a {dream report,question} pair would be reclassified correctly when considered jointly alongside semantically-similar answers in a group that is correctly attributed to a role.*

[5] We considered this pair of opposite situations for visualization because of their popularity in our dataset, but the observed properties hold for different pairs of seemingly opposing situations.

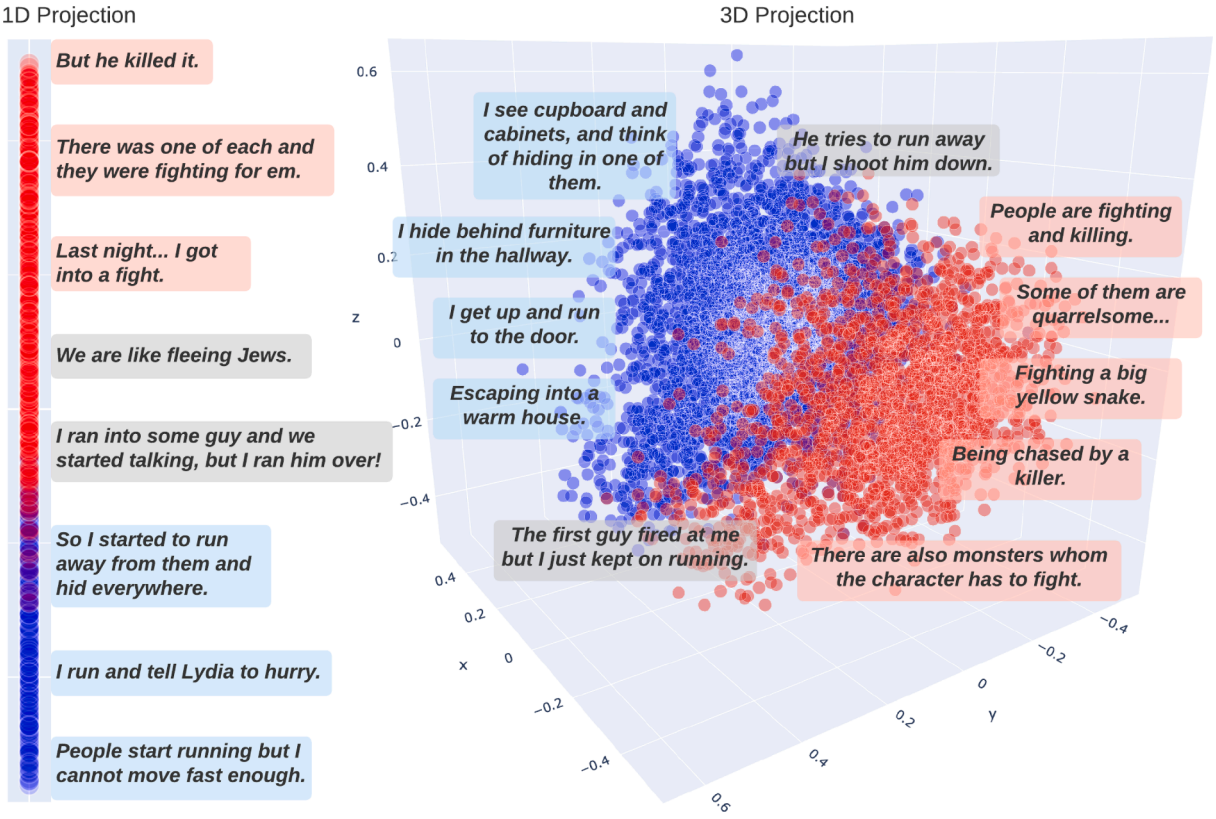

Fig. 2. A qualitative experiment that visualizes the semantic separation in the embedding space between real dream report sentences from SDDb about *Fighting* (in red) and *Fleeing* (in blue): Few neutral sentences from dream reports that included aspects of both *Fighting* and *Fleeing* are in grey. The smooth color gradient suggests that the pretrained numerical representations used in the *Oneirograph* framework contain information that can inform about situations. Dimensionality reduction was *unsupervised* as a result of Principal Component Analysis (PCA). We visualize sentences, as opposed to entire dream reports, because a dream report often consists of multiple situations.

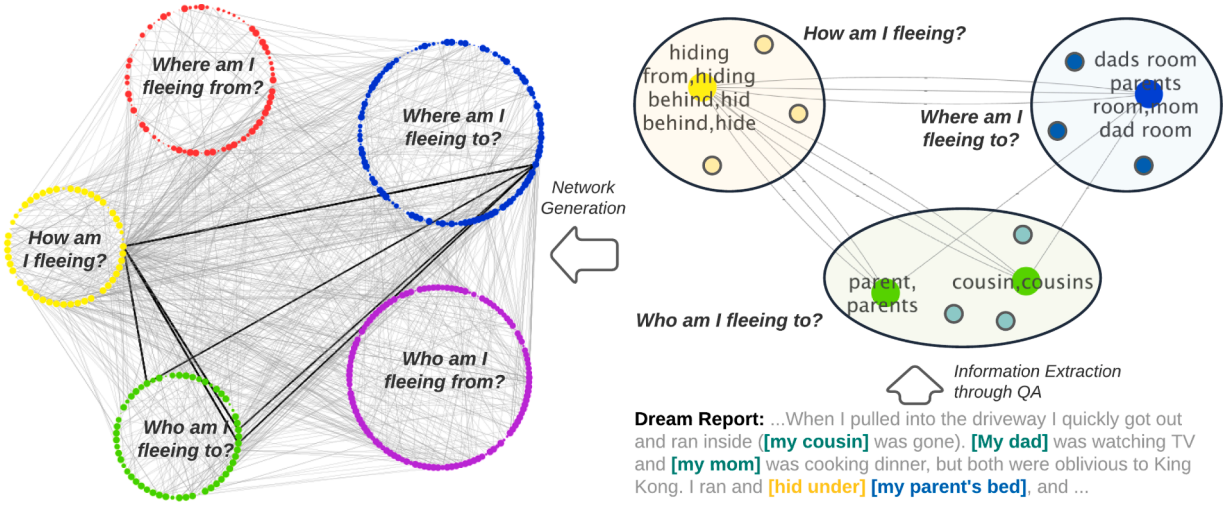

Fig. 3. An overview of the network generated for the *Fleeing* situation and the influence of 1 real dream report sample from SDDb in influencing the final network: The mentioned dream report clearly insinuates the situation of *Fleeing* and the roles that describe the situation are labeled in bold and in italics. The spans from the dream report that map to these roles then contribute to larger groups of actors to provide additional context to the situation described in the specific dream. This way, *every dream report is embedded within a consensus situational framework that provides both lower-order information about specific dreams as well as higher-order information about shared context across dream reports.*

3.6. Module 4: Hierarchical dream network generation

The network comprises nodes and edges organized in the hierarchical layout described in Section 3.2.

- **Nodes:** Each contextual group forms a node in the network. The label of the node is the concatenation of the top 4 N-grams ($N \leqslant 2$) with respect to their inter-cluster TF-IDF scores (cTF-IDF). Stopwords are excluded. The node's color corresponds to its role. Its size is proportional to confidence as output by the KS-test.
- **Edges:** A pair of nodes (contextual groups, say $C_1, C_2 \in C_A$) are connected by a directed edge e_d where $d \in D$ is a dream report, if a pair of answers that originate from d say $a_{r_1}, a_{r_2} \in A$ are such that $a_{r_1} \in C_1$ and $a_{r_2} \in C_2, r_1 \neq r_2$.

As a result, the generated networks are multi-edge undirected graphs, and we generate one such network per situation. Nodes that do not have any emergent or incident edges – this happens when none of the answers that constitute a node co-occur with others – are removed. These nodes are termed *Isolates* (see Section 8 an extended discussion). Networks are visualized with Cytoscape and D3 (Gustavsen et al., 2019; Bostock, 2012). An example of the network generated for the situation of *Fleeing* and a demonstration of the means in which a single dream report *contributes* to this network is presented in Fig. 3.

4. Quantitative summary statistics

We discuss quantitative results summarizing the extracted networks for the *Fight/Flight* and *Flying/Falling*. Additional results are provided in the Appendix A and interactive networks are provided as a .cys file in the Supplementary Material (Gutman et al., 2022).

Across the 6 situations, the degree distributions of the extracted networks are presented in histograms (Fig. 4). *The evidence of non-zero degree is a first step in verifying that a dream report can contain multiple connected semantic constituents to jointly model a typical situation.*

In order to get a better sense of the size of the support data employed to generate these networks, we provide Table 1. The table lists two features for each situation: (left) The relative number of *unique* dream reports that form the support set for generating the situation network. *The same dream report can facilitate the creation of more than one edge: For every pair of roles identified within a dream report, there is an edge between them*; (right) The number of words in the set of unique dream reports that forms the support set. There is a correlation between the size of a dream report and the likelihood of the report representing a situation. The word count histogram provides a reference to interpret the post count histogram.

Oneirograph **confidence measures and adversarial examples:** To ensure that the proposed method is robust, the *Oneirograph* pipeline adopts several confidence tests: (1) Answers as a result of querying TransformerQA are rejected if the confidence level of the answer is too low ($\tau < 0.7$) or if the answer does not exist; (2) The clustering of the answers using UMAP and HDBSCAN provides a measure to capture the likelihood of an answer belonging to a cluster. We realize only those *edges* between a pair of contextual groups, if both the source and target answers belonging to separate contextual groups have a high likelihood ($\geqslant 0.2$) in belonging to the respective clusters. Furthermore, the auto-detected *noise* cluster (labeled by -1) is rejected; (3) Last, the histogram of p-values across the aggregated contextual groups provides a summary view of the metric to measure the bias modeled by our system – a measure-of-fit to the question bank template. A histogram of this metric across situations is provided in Fig. 5. Since most contextual groups (nodes in the situation network) have a *small* p-value, it implies that the groups (nodes) are indeed role-specific and not aggregated at random from across the template. The histogram shows the distribution of the p-values. We would like to see a bias toward 0.

These thresholds together provide a sense of robustness to *Oneirograph* as it extracts networks that match a situation template:

- If all the answers to the questions (per role) are *noise* yet somehow generated by TransformerQA, the HDBSCAN clustering would be diffuse and the grouping process would return all the answers in the noise cluster and thus the *final network would be empty*.
- If the answers were, on the other hand, sensible, but were *shuffled* with respect to the roles, i.e each answer was affiliated to a different role than the one that generated it, clustering would work as usual, but the p-values for each cluster would be unusually high. *The nodes on the network would appear very small.*

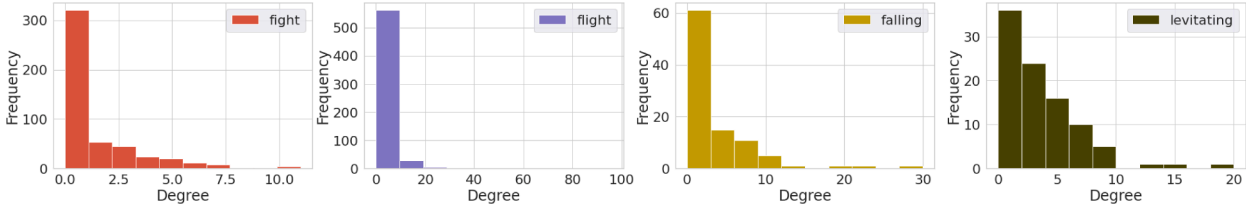

Fig. 4. Summary of the degree distribution of the generated network per situation: The degree of 0 correlates to *Isolates* (Section 8). *A higher degree implies that the contextual group with this high degree is (a) more frequently discussed in dream reports; (b) connected with other roles within the situation*. For example, in the situation of *Fleeing*, a contextual group labeled mainly with *hiding* – belonging to the role of *How am I fleeing?*is observed to have a high degree. Attached files in the Supplementary Material (Gutman et al., 2022) include the entire networks for interactive human evaluation.

Table 1
The number of samples of text data *estimated* by *Oneirograph* to be useful in defining the structures of the different situations: This table verifies that the SDDb database is indeed large-scale – *Oneirograph* has identified thousands of words that are eligible to define a situation. Generally, longer posts provide more situational parts since the likelihood of finding answer spans in a longer piece of text is higher.

Situation	Number of posts that constitute network	Total word length of unique posts that constitute network
fighting	209	27922
fleeing	606	87137
falling	108	13387
levitating	117	15671

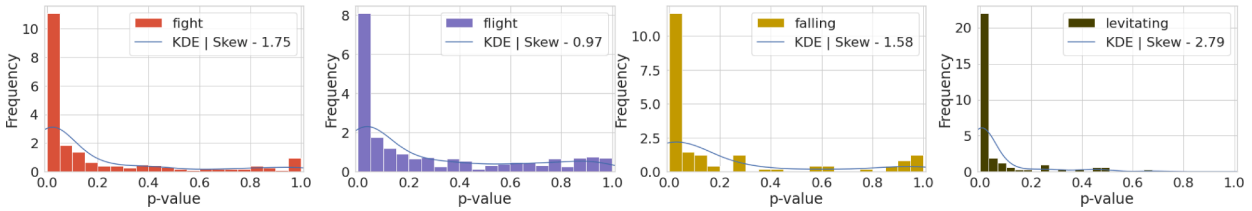

Fig. 5. A histogram of p-values across the contextual groups generated per situation in accordance to Algorithm 3: Lower values are better (typically $p \leqslant 0.1$ hypotheses are considered significant). A lower p-value suggests that the contextual group is comprised of text spans from dream reports that are the result of querying the *same* role. A contextual group formed with answers derived as a response to different roles is unlikely to actually belong to a single role. This results in a larger p-value. The imbalance is also highlighted by the skew - a higher positive value implies more weight in the left tail of the distribution.

- If dream reports did not contain enough information to describe the situation, i.e 0 or only 1 role was discussed, edges cannot be generated on the network and the aggregated contextual groups are classified as *isolates*. *The greatest connected component would have 1 node.*

5. Qualitative results

While the large-scale quantitative measures provide a high-level understanding of the structure of dream networks and our

Table 2
Network-supported discovery of abstract categories, based on automated definition of contextual groups in the *Fighting* situation: These abstract semantic categories are identified by manually exploring the network representation of the collection of *Fighting* dream reports. The structured aggregation process of *Oneirograph* facilitates easier summmarization of recurring themes and actants in a dream situation. More examples can be accessed in Appendix A.1.

Sample Role	Manually abstracted categories	*Oneirograph*: Contextual groups automatically discovered for each category
How am I fighting?	*Tools and Weapons*	knife, rifle, gun, shot gun, pulled revolver pocket shot gun
	Physical Confrontation	stab, grab, push, slap, beat, violently kick, start throwing, attack, stomp
	Voiced Confrontation	argue, accuse, cry, yell, scream, anger
Where am I fighting?	*In Transport*	bus, train, car, ship, raft, fighter jet
	Battle Sites	vietnam, iraq, combat area, big war
	Urban Spaces	high school, restaurant, mall, store, street, top floor, pub, medieval castle
	Natural Environments	jungle, water, rainforest, beachfront, lake
Who/What am I fighting?	*First-name Persons*	marc, tommy, jess, emily, hannah, alec
	Family/ Close Relations	cousin, son, parents, sister, daughter, nephew, dad, father, stepfather, brother, monther-in-law, mother, wife, husband, ex girlfriend, ex husband, girlfriend, boyfriend
	Fictional Characters	vampires, creatures, witch, monsters, ghosts, demons, spirit
	Animals	bear, panther, dog, wild animals, cat, horses, sharks, wolf, snakes, spider, dinosaurs, alligators, bees, rats
	Groups of Power	security officers, police, cop, boss, gangs, savages, pirates, ninjas, rebels, enemy, left wing utopians
Why am I fighting?	*Accusations and Resentment*	still angry, got mad, confronted, scolding
	Discomfort and Disagreement	expensive coat, broken hip, racist vendor
	Mistreatment	prude, insecure, yell dad, angry persecuted
	Physical Threats	biting, beating, bloody mouth, huge blob monster, hit my mom, wants harm

confidence in it, we are more interested in identifying key semantic features brought forth by a close, i.e. human reading of the generated representations. *Oneirograph* provides a means to not only visualize the situation network at scale but also to examine the individual dreams that constitute the nodes and edges that form the network in the first place. As seen in Fig. 3, the nodes and the connecting edges that constitute a situation can be traced directly back to the answer spans and their parent dream reports. This feature allows for the qualitative examination of the different contextual groups discovered by *Oneirograph*, verify their correctness and their relevance to the considered situation.

We thus want to discover groups of entities who commonly appear in situations and the contexts in which these entities interact with one another. Exploring these network parts and the dream reports that constitute these parts forms the Qualitative Results described in this Section. We highlight the insights for the *Fighting/Fleeing* situation.

5.1. Discovered structure of the Fight Situation

The quantitatively derived semantic structure in the *Fight* situation can be explained with automatically discovered contextual groups that correspond to the most prevalent *strategies* of fighting, the common *locations* of fights, the information about the *opponents*, and the *rationale* for fighting. Note that all contextual groups were abstracted into semantic categories manually.

Fighting Strategies: As shown in Table 2, we were able to manually identify specific fighting strategies obtained by *Oneirograph* in its aggregation process from dream reports: The use of *Tools and Weapons* – for example,

> "I have **a sword** and the means to win the fight.".

– suggests that the dreaming subjects use various types of weapons to confront the opponent. *Note that these individual dream reports are the result of the processing from Oneirograph. The NLP pipeline facilitates not only the discovery of the network structure but also transparency to directly access the subjects' dream reports that constitute the nodes and edges.* In addition to *Tools and Weapons*, we observe other strategies to deal with *Fight*s including *Physical confrontation*, and *Voiced confrontation* examples of which can be found in Table 2.

Locations: Computationally generated contextual groups show diverse locations, ranging from *Urban Spaces*, *Battle Sites*, *Transport*, – for example,

> "/f/ighting on a moving**train** with the group of women."

and *Natural Environments*:

> "I am on a wooden raft in **a lake**. A huge black man and I are fighting."

Identity of the opponent: (queried as *Who or What am I fighting?*) This contextual group appears to have the majority of hits. We observe that the associated spans obtained from *Oneirograph* can be manually partitioned into 5 distinctive semantic categories: (a) Various *familiar names* in the first category suggest that the dreamers often fight with friends/acquaintances familiar to them; (b) The second category includes *Close family-members* as the opponents. For example:

> "**My mom** came out and the more I faded and tried to talk the stronger her energy was towards me. She was beating me down…"

(c) The third category consists of various *Fictional characters* with either cinematic or folkloric references; (d) Many times dreamers confront *Animals*, and last (e) Many opponents hail from *Groups of Power* such as the *police*, *boss*, *gangs* or *pirates*.

Motives behind fighting: (queried as *Why am I fighting?*) The answer spans returned for this contextual group can once again be summarized into the following categories: *Accusations and Resentments*, *Discomfort and Disagreement*, *Physical Threats* and *Mistreatment*. A representative example of a dream that features mistreatment:

> "At one point we were in a car accident, and I tried to help him, **but he spat at me**. Finally I pushed him hard enough, and he admitted that he hates me because I had a charmed life…"

5.2. Discovered structure of the Fleeing Situation

Oneirograph appears to extract equally convincing representations of the *Fleeing* situation as the *Fighting* situation and qualitatively, we observe fascinating groups of answer spans across dream reports that mention among other concepts – semantic categories: (a) The *trajectory of escape* (initial and end location); (b) The *subjects and objects of a threat*; and (c) The various *modes of escaping*. We highlight a few interesting groups below:

The subjects and objects of a threat: (queried as *Who or What am I fleeing from?*) These spans appear to have clear semantic sub-categories of *Family & Close Relations*, *Distant Non-fictional Characters*, *Fictional Characters*, *Animals*, and*Groups of Power*. A threatening appearance of close relatives can be often found in many posts. For example:

> "I had a dream that my **grandmother** was a ghost. **She chased me** into the living room of a three bedroom apartment where my family and I had once lived. I tried to run to my parents but she caught."

Animals, too, regularly appear hazardous, for instance:

> "I was being chased by a mean **grizzly bear** that was trying to kill me and my sons and I was running everywhere trying not to let the grizzly bear get me or my sons."

The variety of fictional characters suggests that the identity of the threat can be imaginary, but nonetheless well-defined and purposeful:

"**Aliens** secretly harvesting humans under the guise of bring in utopia but I get suspicious and run away with the girls."

Strategy of escape: (queried as *How am I fleeing?*) We further divided this role into two escaping strategies: *Primary movement* which mainly consists of bodily reactions that support fast physical response, such as:

"Gunmen came in and started shooting everyone in sight, I **ran** as fast as I could."

The *secondary movement* consists of popular man-made modes of transportation.

People or Persons who are safe: (queried as *Who am I fleeing to?*) Many dreamers recount escaping to the homes that belong to families, particular family members, or friends. For example:

"I have divorced parents Mother is mean in the first day if school so **I run to my fathers house.**"

Locations of initial Escape: (queried as *Where am I fleeing from?*) *Oneirograph* identifies *Public Spaces*, for example:

"I was running away from bad guys **at a hospital,"**

followed by *Battle Sites*, *Natural Places*, and *Natural Phenomena*, for example:

"I was running away **from a thunder cloud**with a perpetual streak of lightning coming out of it and also a tornado. I was all by myself and scared to death."

Note that some groups of answers that are observed the trajectory of escape are not strictly related to locations. We find many *Animals*, *Fictional Characters & Creatures*, *Non-Fictional characters*, and *Relatives* and indeed these groups are similar to the groups obtained for the query *Who or What am I fleeing from?* Given that both queries inquire about the threat that can potentially represent the initial place of escape, the partial overlap in categories is expected. For example:

"At that time, I saw a **big scary snake** that was coming towards me. I started running away **from there** and the snake was chasing me..."

demonstrates that the QA model used in *Oneirograph* accurately identified the answer to the question *Where am I fleeing from?*, despite

Table 3
Network-supported discovery of abstract categories, based on automated definition of contextual groups in the *Fleeing* situation: Similar to the observations listed in Table 2, these groups of phrases are identified for the *Fleeing* situation. More examples can be accessed in Appendix A.2.

Sample Role	Manually Abstracted categories	*Oneirograph*: Contextual groups automatically discovered for each category
Who/What am I fleeing from?	Family/ Close relations	dad, mother, aunt, son, college ex-boyfriend, husband, mother, grandmother
	Distant Non-fictional Characters	guy random homeless, evil men, bad guys, chased somebody something, bad guys
	Fictional Characters	monster, darth vader, evil monster, devil, aliens, witch, giant teddy, big monster creature, godzilla, mutant zombie, vampire
	Animals	black bear, snake, guenia pig, bees, fox, leopard, tiger, shark, wolf, panther spiders, rabbits, gorilla, grizzly bear, panther, bat, alligator
	Groups of Power	police, boss, mayor county, sheriff, officers, policemen, teacher, soldiers
	Popular Figures	mel gibson, hugh jackman
	Violent Envoronments	holocaust, nazi checkpoint, streets war
Where am I fleeing from?	Public Spaces	store, mall, college, disneyland, temple, hospital, prison, train station
	Battle Sites	air force base, military
	Natural Places	woods bunch, outside, night
How am I fleeing?	Primary Movement	escapes fly, jumped, get away, ran away, running top, tree, climbed tree, river swam, hiding from, hiding behind
	In Secondary Transport	drove away, car escaped, riding car sidewalk, boats leave flights stairs, train
Who/What am I fleeing to?	Friend's or Family's homes	fathers house, house dad house, neighbor house, stranger house, parents, grandparents, mom house, friend house
Where am I fleeing to?	Public urban Places	neighborhood street, theater, school lab, school, old neighborhood, park, corner building, library, church, café, allies café
	Places in Nature	mountains, cave, lake, large forest, valley woods, wood, mountains, mountains, top field across
	Cities and Countries	chicago, grenada, new york, england, mexico, italy rome, tokyo, jamaica, australia, ireland, europe, new zealand, california, city, sao paulo
	Private Places	home, friends apartment, mothers house, backyard, bedroom window, grandparent home, loft attic, house bathroom, dad['s] room, parents' room, friend house, treehouse, house, kitchen, dorms, hotel room, farm house
	Vehicles	truck, armored car, camp trailer, bus

the indirect description of the initial point of escape as an *animal*.

Place of Final Escape: (queried as *Where am I fleeing to?*) This group illustrates the variety of locations that represent places of safety, such as *Public Places*, for example:

> "I'm in **a school, in a classroom**, hiding from the Devil amongst the students."

Other locations include *Natural Places*, *Cities and Countries*, *Private Places/ Homes* – for example,

> "I was in bed in my room and then a bunch of statues that looked like the oscars appeared and started toppling over toward me. I ran out **to my mom's room** barely being missed by each one".

and *Vehicles* (see Table 3).

6. *Flying/Falling* situation: Flexibility of the *Oneirograph* Framework to extend to other Situations

We mentioned that the *Oneirograph* has the potential to be expanded on other typical situations in dreams. To further support this hypothesis, we applied our framework to another binary situation of *Flying/Falling* and examine its performance. In our experimental findings, we limited the extraction of the *Flying situation* to three questions in which we explored: the *locations* of flying, *how* the dreamers are flying, and, finally, if the flying is a *social or individual activity*. After reviewing the contextual groups with locations, we manually abstracted two semantic categories:

(i) *Urban spaces*; for example:

> "Flying **over the city** in my pajamas. Once I took off so fast my bottoms were coming off from the force of the wind..."

(ii) *Natural environments*, with examples, such as:

> "I am able to fly, and I fly all over the globe, especially**to exotic islands**, where I stay most of the time."

How dreamers are flying was mainly limited to *bird-like imitations*, for example:

> "I am near a cliff and think that it would be fun to fly. So I jump off the cliff. At first, I am terrified, but then I start to **move my arms…**"

In the last question, queried as *Who am I flying with?* we implied that flying is a collaborative practice. The posts from this role, however, showed a different connotation of flying, where dreamers were air traveling with others, for example:

> "I was traveling in an airplane to the state of Rio Grande de Sulin southern Brazil. With me was **my aunt."**

The change in the connotation of flying enriched our observations that taking a commercial flight can be categorized as a particular journey situation, which, as such, involves a company others. On the other hand, flying with our bodies did not reveal posts that included others. This 'conspicuous by its absence' reasoning suggests that the typical flying situation is *a solitary activity*. To validate this reasoning further, we manually extracted 541 dream reports on flying from the SDDb that were limited to flying with one's body. The close reading suggests that dreamers seldom mention anyone else joining them in their flying adventure. The typical flying situation thus seems to be a solitary experience, often described as:

> "I was flying over the ocean. I looked down, and there was a huge, square hole in the ocean. Water was pouring down into it like a waterfall, and it went down for miles. It was deep blue at the bottom. Dolphins and colorful tropical fish were leaping across the opening, from one wall of the hole to the other."

As mentioned earlier, we tested the performance of *Oneirograph* on the binary opposition axis, which also includes the aspect of *Falling*. In this particular situation, we were mostly interested in *trajectory* (*Where am I falling from and to?*) and *how* the falling occurs. After reviewing the contextual groups with the initial location of the fall, we abstracted two categories that can be broadly segmented into:

- *High urban structures and Aeroplanes*, for example:

  > "Falling from **a ferris** wheel into the river at our local amusement park…" or,

  > "Falling off a**high building**, hitting the ground. I woke up when I hit the ground."

- *Highlands* in natural environments, for example:

  > "…was falling down a **really long mountain** on my white horse unicorn and landing in water. I was falling forever and it was a freeing but terrifying feeling…" or,

  > "Falling **off a steep hill** and waking up terrifed."

Lastly, we examined the contextual groups with the final locations of the fall. Two categories emerged from these groups, namely:

- *Water* – many posts mention water as the final destination of the fall, for example:

  "I suddenly felt myself falling into space without even coming to a halt. Suddenly I plunged into the **icy cold water**. I could not swim. I cried for help…" or,

  "I fell into **the ocean** off of the porch and the ocean was really choppy."

- The second category includes posts in which dreamers are either *hitting the floor or fall deeper*, for example:

  "The only part of it that I remember is me in midair falling from the second floor (inside) of my childhood home**to the ground floor**, with me waking up as I hit the ground" or,

  "I am running along a path when I suddenly **fall into a hole and at the bottom** are all kinds and sizes of snakes."

The final examination on how dreamers are falling (*How am I falling?)* revealed that falling can occur as an accident or a consequence of some previous situation. In an example,

"I was **running from a monster** and fell off a cliff and saw myself splat on the rocks below the cliff."

we can see that the falling is causally linked to escape. Not all extracted posts reveal that information, so more dream reports are needed to explore this category further. More examples from the *Flying* and *Falling* networks can be found in the Appendices section.

To summarize: Despite being different when it comes to sensation, the *Flying* and *Falling* dreams have one thing in common: the dream contributors either feel the amplified pull of gravitational force or a floating sensation of zero gravity in the case of flying. In the first case, the visceral feeling of falling is predominantly characterized as a negative experience, whereas the sensations associated with flying are reported as positive and pleasant. We leave a more detailed accounting of the *Flying* and *Falling* dreams and their situational networks to the Supplementary Material (Gutman et al., 2022).

7. Discussion

The quantitative and qualitative results presented here offer strong evidence in favor of the two hypotheses proposed at the outset. A specific focus of the *Oneirograph* on the typical dream situations of *Fighting and/or Fleeing* demonstrated that they have an underlying, computationally mappable semantic structure (H1). The additional application of the *Oneirograph* to the typical dream situations of *Flying/Falling* revealed deep linguistic structures in these dreams, too, showing that other typical situations beyond *Fighting and/or Fleeing* can be computationally mapped by an automated system in terms of a distinctive constellation of characters, settings, and activities (H2).

The results regarding the most frequent content patterns of *Fighting and/or Fleeing* dreams are especially informative. Our analysis treated each part of the two situations separately to extract meaningful contextual groups that we later abstracted into semantic categories. As shown in Tables 1 and 2, we manually extrapolated the automatically generated contextual groups into semantic categories, enabling us to discern the shapes of the most dominant features of the two situations, and thus to identify some fundamental similarities between them.

Our results indicate that both *Flight* and *Fight* dreams share some common semantic categories. In terms of the setting, we observe that the dream contributors frequently mention urban and public spaces in their reports. In dreams of aggressive confrontations (*Fight* response), urban spaces like high schools, restaurants, malls, stores, and streets appear frequently in response to the question *Where am I fighting?*. Similarly, in the *Fleeing (Flight)* dreams, the urban spaces are frequent, both at the start and end of a fleeing dream. In the initial locations of escape (*Where am I fleeing from?*), we found contextual groups such as store, mall, college, Disneyland, temple, hospital, prison, train station. The final place of escape (*Where am I fleeing to?*) revealed another group of public spaces that consist of contextual groups like neighborhood street, theater, school, lab, old neighborhood, park, corner building, library, church, café, and others.

Another *significant overlap can be found between threats and opponents.* We can observe that objects and subjects of a threat in the *Fleeing (Flight)* response (recall the roles *Who or What am I fleeing from?*, and *Where am I fleeing from?*) share the same categories with opponents that the dreamers confront in the *Fight* situation (*recall the role Who/What am I Fighting?*). For example, the character "mother" appears in both *Fight* and *Flight* dreams, playing two different roles. This overlap suggests that the dreamers are either confronting or escaping the same character that poses a danger to them. The dreamers reporting *Fight* responses are confronting their significant others (expressed and first-name persons), family members, fictional characters, animals, and groups of power. Similarly, the dreamers reporting the *Fleeing* response primarily identify their object of threat as family members, fictional characters, animals, and groups of power. Such similarities suggest that the dreamers in *Flight/Fight* situations *are typically confronted by the characters they are escaping from.* Furthermore, urban public spaces seem to be playing an important role *either as a setting where the confrontation is taking place or*, in the case of *Fleeing*, as *the initial or final location of escape*.

Looking beyond the *Fight/Flight* situation, the results of this study have shown that the *Oneirograph* can be applied to other typical

dream situations as well (e.g., *Flying/Falling*), and indeed to any other pre-defined situation that is amenable to a 5W1H question type of method. This opens new possibilities for researchers to use this tool in gaining theoretical and perhaps practical insight into the nature of dreaming. Two examples can be suggested. First, the continuity hypothesis (Hall & Van de Castle, 1966; Domhoff, 2003; Schredl & Hofmann, 2003b) proposes that frequencies of dream content correspond to waking life concerns and preoccupations. Following the method outlined in this paper, the continuity hypothesis could be tested by using the *Oneirograph* to analyze the *Fight/Flight* situation in the dreams of people who in the waking world A) live in a place of extreme physical danger; B) live in a physically safe place but have cultural interests involving extreme physical danger (e.g., in video games), and C) live in a physically safe place and have relatively peaceful cultural interests. The results of the *Oneirograph* analysis would presumably show more semantically rich *Fight/Flight* dreams among the people in groups A and B than in group C. The results might also enable a new and more precise understanding of the differences between *Fight/Flight* dreams with physical continuities (group A) versus dreams with cognitive-emotional continuities (group B). As a second example, the Immersive Spatio-Temporal Hallucination model (ISTH) (Windt, 2010) defines dreaming as the activation during sleep of a reference frame in which the dreamer is oriented as a self who feels present within a (hallucinated) environment of space and time: "dreaming is an inherently variable phenomenon, and according to the ISTH model of dreaming, these variations can be understood as occurring around a common phenomenological core." (Windt, 2010 Pg. 309) To test this claim, the *Oneirograph* could be used to analyze the semantic structure of many different typical dream situations in addition to *Fight/Flight* and *Flying/Falling*, such as swimming, sexual experience, and seeing dead people as if alive, with a special focus on the spatial and temporal qualities that help to constitute each situation. The results of such an analysis, according to the ISTH model, would reveal the range, plasticity, and relative importance of the various elements that together form the common phenomenological core of dreaming.

In addition to prospects for theoretical insight into the nature of dreaming, the *Oneirograph* also has potential applications for practical dream analysis. For example, in the course of therapeutic interventions certain semantic features in dreams, like characters, social interactions, or settings, can change over time; such change may indicate successful conflict resolution or other improvements. Implementing the *Oneirograph* to recurring situations in a client's dreams would give both the therapist and client a statistical observation of the dreaming dynamics, i.e., temporal changes in dream content over time. If the changes are positive, perhaps a conflict-resolution has begun; if the situation in recurring dreams remains without alterations, the conflict may be unresolved or stuck in place. Potential applications to individual treatments and population-based research can thus be listed as follows:

- **Individual-level:**
  1. Extraction of recurring dreams: iterative contents in dreams are a common concern in psychotherapeutic sessions. Computer-implemented tools can extract the essential information from recurring dreams to estimate the current situation in waking life;
  2. Monitoring changes/manipulations in dream content: some elements in dreams, like actants or settings, can change over time. Implementing the *Oneirograph* to recurring situations in dreams will enable statistical observation of dream dynamics;
  3. Remote treatment – the patient can send or upload their dream logs at any time and collaboratively discuss the computation analysis using the *Oneirograph*;
  4. A potential to improve therapy relationship – enhanced understanding of dream dynamics using NLP tools from both parties;
  5. Data-driven decisions about treatments – a therapist can use an empirical tool that helps her/him extract the most important features of dream content and make further decisions about treatment based on individual and/or population statistics.
  6. Improved individual/patient self-awareness – dream reports provide a guided understanding of circumstances, desires, daily concerns, and traumatic events. When designed, developed, and trained on comprehensive data, the computational tools can provide the therapist with applicable statistical models that complement the existing analytic methods of treatment.
- **Population-level and the use of baselines:** The *Oneirograph* could be utilized in future dream research to set up a baseline in typical dream situations, such as *Flight-Fight* and *Flying-Falling*.
  1. The dream science researchers and practicing therapists could use these baselines to analyze the situations in patients' dreams and compare them with findings from our study.
  2. *Oneirograph* can be used in new semantic extractions in the future of dream research; by building a constellation of situations in dreaming, the therapists and researchers would contribute to the global collection of dreams in which more situations would be computationally extracted. In the final instance, the population-based typical situations could represent the new data-driven statistical standards of typically dreamed situations.

The primary purpose of the *Oneirograph* is to identify and extract meaningful semantic structures in dream reports. We have observed that dreams reveal a powerful complexity of our minds – the concerns, desires, fears, and relationships with others – and can thus be considered one of *the most intimate forms of psychological data available*. Thus, an automated semantic extraction can pose ethical concerns if the tool is applied to individual dream series without the prior consent of the dream contributors. Moreover, such analytical treatment of dreams can pose ethical concerns if the research is motivated by other than academic or therapeutic concerns. For this reason, we caution again uses of the *Oneirograph* in which the results are applied without appropriate respect for the individual dreamers and the basic ethical principles of scientific research.

Other researchers can access the *Oneirograph* by following the instructions on the OSF resource.

8. Limitations

Oneirograph marks a first attempt at discovering the latent semantic structure that underlies the situations commonly appearing in

dreams. The resulting graph is a collective of nodes (concepts) and edges (co-occurrences between concepts) that can be sampled in order to produce individual dream reports. The following points list the limitations in our pipeline:

- The TransformerQA model used for inference is deep and while highly effective at returning relevant answer spans from the context in response to a question, sometimes provides incorrect answers or simply no answer to a question in the context. For example, the dream,

  > I can't remember much of the dream, but do remember there was a tall black man standing with his back to me. I see a bunch of white women standing at the table with him. […] I'm hanging around cleaning up the kitchen. The kitchen is a big mess and a woman arrives to help me clean. I'm putting dishes and foods away and making sure we bring home the dishes and foods that belong to us. […].

  queried with the question, "Who was I helping?" returns a larger span than should be admissible: "*cleaning up the kitchen. The kitchen is a big mess and a woman arrives to help me clean.*" Later stages of the pipeline that include a clustering step help to filter these outliers.

- While clustering these answer spans, we use UMAP + HDBSCAN and while the clustering method yields several groups that are thematically similar, there is a propensity to label several spans as noise. The parameters of HDBSCAN can help vary the size of the noise cluster. Based on qualitative examination of the clustering results – the clustering must be granular enough to distinguish between the different contextual groups and yet general enough to aggregate similar mentions – we decided on the values of the parameters (see *Oneirograph* codebase in the Supplementary Material (Gutman et al., 2022) for the exact implementation). But this still leaves room for variability in the quality of the returned groups.
- When we assign a role to each cluster, we use the majority vote among the roles associated to all the answer spans in the cluster. In some cases, when the number of answer spans that constitute a cluster are few, the majority vote (and its confidence measured by the p-value with respect to a random assignment) may be misleading.
- Since situations are prescribed a template of interacting semantic constituents, we require each dream report to contain more than semantic constituent in order to qualify for representation in our network. There are, however, several dream reports (especially the reports that are short) that are relevant but do not contain more than one semantic constituent. For example: "feeling I was falling to my death" or "wHEN YOUNG, DREAMED OF FLOATING AND BEING ABLE TO FLY."
- While *Oneirograph* is designed to extract deeper semantic features from a particular situation or theme, it cannot give detailed information on the prevalence of such a theme in the SDDb database.
- Short dream reports are less likely to be captured by our tool for two reasons: (i) fewer sentences imply lesser semantic constituents and decreases the recall of the TransformerQA model simply because there are not enough answers that would satisfy the roles. For example, single-sentenced dream reports, such as "flying thru the air," "Flying–Hero type dreams," or "yes, I was flying & knew it was not real & told myself to wake up" lack the information about trajectory and thus cannot be detected by the question *Where am I flying?*. In many cases, these shorter dream reports yield isolates: singleton nodes that represent isolated roles without connecting to other components of a situation. Such isolates are filtered out in the automated method: However, these short reports should not be neglected under the false belief that *Oneirograph* is extracting *all* the important information from dream reports – it is targeted toward extracting graph structures that describe situations. Isolates might in fact be very therapeutically valuable when studied on the individual level.

9. Conclusions

The semantic structures of dreams with typical situations like *Fighting*, *Fleeing*, *Flying*, and *Falling* can be identified by the *Oneirograph* analytic system and computationally mapped in terms of their distinctive relations between characters, settings, and actions. The *Oneirograph* builds on significant new advances in NLP to develop a powerful, question-driven method of drawing information from large collections of dream texts. This approach can be applied more generally to the study of dreams in individuals and groups as a way of identifying and precisely measuring the underlying interrelations of various elements of meaningful content.

This project's driving idea has been that integrating conceptual frameworks from dream research with methods in computational linguistics can open new vistas in both the scientific study of dreams *and* technologies of automatic semantic analysis. The interdisciplinary approach demonstrated here consists of close (human) reading in tandem with distant (computer) reading of dream reports. Both methods are different in their nature and thus pose various advances and challenges–close reading is detailed and sensitive to nuance, but slow and dependant on intercoder reliability, while distant reading is automated, fast, but less contextual. Together, however, the limits of each can be overcome while the challenges of each can be enhanced with the strengths of the other. Overall, the combined quantitative and qualitative results of this study offer a vivid illustration of the principle of collaborative synthesis between machine and human methodologies in dream research.

Funding sources

This work was supported by the Fulbright Visiting Scholar program 2020–2021 and the private funding from the Sleep and Dream Database.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. [Dr. Maja Gutman Music, Department of Anthropology, University of California, Los Angeles / Institute for Philosophical Studies, Science and Research Center, Koper. Pavan Holur, Department of Electrical and Computer Engineering, University of California, Los Angeles. Kelly Bulkeley, Ph.D., Sleep and Dream Database.]

Acknowledgements

The authors thank the reviewers for their thorough and valuable comments that helped improve the manuscript. In addition, the authors would like to thank Professor Vwani P. Roychowdhury (University of California, Los Angeles) for outlining the key facets of developing the situational model used in *Oneirograph* and his generous support of ideas and funding. The first author would also like to extend her gratitude to Professor Douglas W. Hollan for hosting her at the UCLA Department of Anthropology during the Fulbright grant period.

Appendix A. Appendices

The networks referred to in this section are provided as a joint interactive cys file attached in the Supplementary Material (Gutman et al., 2022).

A.1. A deeper analysis of the Fighting Situation Network

Before diving into the network, the reader should first be introduced to the color-coded roles, represented as a circular group of nodes. The *Fighting* situation consists of four circularly outlined groups: (i) the group of red nodes represents the fighting strategy and thus answers the question *How am I fighting?*(ii) Fighting strategies naturally pose a question about the opponent. We formulated this query as *Who or What am I fighting against?* as represented in the light pink colored group of nodes. (iii) The group of blue nodes represents the fight's location, i.e., answers to the question *Where am I fighting?*; (iv) the turquoise group of nodes represents reasoning behind fighting by answering the question *Why am I fighting?*

Nodes are defined as contextual groups that consist of semantically similar answers (n-grams). To better demonstrate the clustering functionality, we observe all four circular groups of nodes, examine the edges and identify examples of phrases from the posts that validate the accuracy of the roles:

- **How am I fighting?**: The strategy of fighting consists of various approaches and techniques that can be abstracted into four categories:
  - **Tools and weapons** ('knife,' 'shooting,' 'rifle,' 'gun,' 'pulled revolver, rifle shot, revolver pocket shot, gun shot,'' sword'). An example: "I have a sword and the means to win the fight."
  - **Physical confrontations** that rank from moderate to severe ('grabbing,' 'pushing,' 'hitting,' 'slapping,' 'beating,' 'stabbing,' 'repeatedly kicked,' violently kick, 'attack fists,' 'stomp, started throwing, use weird power,' 'snapping, punched,' 'killing.'. Examples like "I **stabbed** him in the heart with a piece of wood" and "I started working my **gijistu moves** on them, and then I grabbed some bottles that were next to the stairs and started throwing them at them" illustrate various uses of physical aggression.
  - **Voiced confrontations – verbal** ('arguing,' 'accusing'), for example: "An unknown female was **arguing** with me about the length of my hair."
  - **Voiced confrontation – Non-verbal** and Emotional distress (represented in clusters like 'anger,' 'crying,' 'screaming, 'yelling'). An illustrative example of emotional distress can be found in the following situation: "I kept hitting him and started to **scream** when he stabbed me."

- **Where am I fighting?**: The fighting location varies from private to public places, emphasizing battle sights and educational spaces that metaphorically serve as the battleground. A close observation of the aggregated clusters prompted us to categorize 6 locations where fights are mainly taking place:
  - **Urban spaces** (consists of clusters such as: 'restaurant,' 'mall,' 'store,' 'street,' 'top floor, floor building,' 'pub,' 'medieval castle'), for example: "/a/t an unknown **pub** in town. with my dad mat and Danny and unknown guy who I started fighting with and Danny broke it up" or "I AM **in a building** with lots of odd monsters I have a sword and the means to win the fight."
  - **Educational environments**('school,' 'high school'), for example: "I'm in the big gym at JR Robson high school. I'm in **junior high…**"
  - **Battle sites** (consists of clusters such as 'big war,' 'military people,' 'vietnam,' 'army,' 'military,' 'battle,' 'combat,' 'jet fighter,' 'iraq,' 'combat area,' 'marinespecial forces'), for example: "I was in a **combat area**, fighting the enemy. There were many enemy soldiers attacking our zone, and I kept shooting them down with my rifle."
  - **Domestic environment** (consists of clusters such as 'mum's house, parents house,' 'home,' 'bed,' 'bathroom,' 'kitchen,' 'outside' 'house,' 'house road,' 'barn,' 'farm'), for example: "At my **parents' house** and I'm getting mad at my mom for making me go out with her in the wind and snow, even though her hip is broken."

- **Natural environments** (clusters like 'jungle,' 'water,' 'rainforest,' 'ship, beachfront,' 'wooden dock,' 'lake'), for example: "I am on a wooden raft **in a lake**. A huge black man and I are fighting."
- **In transport** ('train,' 'bus,' 'car,' 'ship, beachfront'), for example: "Later in the dream I was running and fighting **on a moving train** with the group of women."
- **Whom or What am I fighting (against)?**: These results suggest that dreamers prioritize the opponent's identity more than any other semantic information in the text. Many clusters with personal names were detected in this group, suggesting that we often confront significant others in dreams. The contextual groups that yielded answers on the identity of the opponent can be abstracted down to 9 categories:
  * **First name persons** (clusters like 'marc,' 'tommy,' 'jess,' 'emily,' 'hannah,' 'alec'), for example: "**Marc** is being a prick, and he throws orange juice on kerry hair so I pour it all over him. We start to fight. He jumps on my back so I slam him to the floor…"
  * **Family** numerous clusters such as 'family,' 'cousin,' 'son,' 'parents,' 'sister,' 'daughter,' 'nephew,' 'dad,' 'father,' 'stepfather,' 'brother,' 'monther-in-law,' 'mother' suggest that close family circles can often pose a site of friction. The following two examples demonstrate the threatening role of a mother: "**My mom** came out and the more I faded and tried to talk the stronger her energy was towards me. She was beating me down; I could feel it on my legs. She was yelling at me. I tried to get away, I went across the street, I was hiding in the grass but I could NOT open my eyes! I couldn't see where I was going! I was afraid. My eyelids were glued. She was hitting me on the leg. The more I struggled the worse it got." And: "I was really mad at my **Mother**, she was making me feel guilty and I had had it. I began banging on her bedroom door and yelling "f*&ŷou" again and again. I was very angry, then I woke up".
  * **Close Relations (Present and Previous Spouses and Partners)** clusters like ('wife,' 'husband,' 'ex girlfriend,' 'ex husband,' 'ex,' 'girlfriend,' 'boyfriend,' 'lovers') indicate that dreamers often confront their present or past partners, for example: "Arguing with **my wife** for trivial reasons; felt angry, woke up in a discordant mood."
  * **Fictional characters** ('vampires,' 'creatures,' 'witch,' 'monsters,' 'ghosts,' 'demons,' 'spirit'), for example: "A **ghost or a evil spirit** had me up off the ground and pinned against my bedroom window. I was pushing against the window with my hands and trying to push my head away from the window, but niether one would move."
  * **Animals** ('bear,' 'panther,' 'dog,' 'wild animals,' 'cat,' 'horses,' 'sharks,' 'wolf,' 'snakes,' 'spider,' 'dinosaurs,' 'alligators,' 'bees,' 'rats'), for example: "Flying in an open cockpit WW1 era biplane, engaged in a **dogfight**."
  * **Groups of Power** ('security officers,' 'police,' 'cop,' 'boss,' 'gangs,' 'savages,' 'pirates,' 'ninjas,' 'rebels,' 'enemy,' 'left wing utopians'), for example: "Then a **group of pirates** showed up and tried to steal the animals which were my responsibility. I tried to stop them but the captain (Who oddly resembled Captain Hook from the cartoon Peter Pan) picked me up and was alughing at me. I kicked him he dropped me and the pirates ran away."
  * **State leaders and personalities** ('clinton,' 'obama,' ' president, trump,' 'jim carrey,' 'brittany spears'), for example: "**Barack** and another politican who was a democratic nominee – but I don't think exists, or maybe it was John Edwards – were having a gun fight in my basement."
  * **Undefined characters** ('somebody,' 'someone,' 'man,' 'lady,' 'woman,' 'girl,' 'little boy,' 'boy,' 'people,' 'guy,' 'white woman'), for example: "/t/he **man** was fat and he grabbed me and I slipped away from him and then I grabbed him by the hair and pushed his head under the water in my washing machine."
- **Rationale**: Finally, we examined the *rationale* for fighting. Based on our observation of the clusters, we identified eight semantic categories:
  * **Accusations and resentments** ('angry still angry, got mad, angry still, mad girl (-) hit bow arrow, set arms, bow arrow set, hit the bow,' 'confronted someone, accuse, accuses shallow, confronted someone thought'), with the following example: "I remember being in a car with someone who was on my side and with someone who was not. The one who was not on my side was an Asian woman and she **said something to make me mad** so I hit her with a bow and arrow set that I had in my arms. I hit her in the eye and it almost cut her eye."
  * **Direct threat** ('man wants harm, wants harm family,' 'egor, fucking dudes,' 'giant zombie deer, deer bloody mouth,' 'chasing, monsters chasing, chasing monster'), for example: "I was being chased **by a huge blob monster** that looked like purple jello. I shot it with a rifle…"
  * **Mistreatment** – numerous examples of wrongdoings as motivators for the fight can be found in different nodes such as ('big fight ava,' 'got mad beat, yell dad,' 'egor, fucking dudes egor…'). The following examples illustrate various types of mistreatment: "I got in a big fight with them and Ava **cause I wasn't treating AVA well**."/ "Fighting with Josh **about stupid stuff on his Facebook profile…**"
  * **Discomfort** ('woman emerald'). Example: "At my parents' house and I'm getting mad at my mom for making me go out with her in the wind and snow, **even though her hip is broken**."
  * **Disagreements** ('coat trying fix, expensive coat,' 'spot wanted, tried take nice'), for example: "…arguing how they were going to decorate the house."
  * **Justice** ('onto floor saying, floor saying disrespectful'), for example: "I was fighting with a very large man, medieval style, **over a series of murder charges against him**."
  * **Personal traits** ('scolding father scolding'), for example: "I am irritated by my father's **looseness…**"

    It is worth pointing out that some clusters are not immediately clear in a strict semantic sense, which can be attributed to the poorly defined or indirect reasons stated in the posts. However, an increasingly high level of accuracy in answers can be found in posts with high-order syntax, where the reasoning behind the fight has solid psychological components. Such examples can be found in the *Mistreatment* category.

A.2. A deeper analysis of the Fleeing (Escaping) Situation Network

The *Fleeing* situation, which represents the largest network in our study, consists of five groups: (i) The group of red nodes represents the initial place of escape and thus answers the question *Where am I fleeing from?*; (ii) The group of blue nodes, which makes the largest group in the network, represents the final location of escape, formulated as *Where an I fleeing to?*; (iii) The group of pink nodes represents the subject or object of a threat, and is formulated as *Who or What am I fleeing from?*; (iv) The green group of nodes represents the final place of comfort which can be a person or object. The answers in this group correspond to *Who or What am I fleeing to?* question; (v) The fifth group, color-coded in yellow, represents the strategies of fleeing, which correspond to the question *How am I fleeing?*.

- **Where am I fleeing from?**: We abstracted the results into 8 semantic categories:
  * **Animals** ('snake,' 'bear,' 'great white shark,' 'chasing squirrel,' 'dog,' 'spider giant,' 'gorilla,' 'buffalo,' 'black horse'). Example: "At that time, I saw a **big scary snake** that was coming towards me. I started running away from there and the snake was chasing me." Example: "I was wondering through the woods when**a bear**came upon me. I began running away from the bear for a long time."
  * **Public spaces** ('store convenience, mall,' 'college,' 'disneyland,' 'temple,' 'hospital,' 'prison,' 'station train station,' 'police station,' 'class linguistics class, classroom'). Example: "I was running away from bad guys **at a hospital**." Example: "I killed another woman in this dream, and I am escaping from the police. I am hiding in the **underground**, and I am living there. Now I live with a terrified feeling all the time. I am very scared that the police will find me."
  * **Battle sites** ('attack building,' 'air force base, people military hassling, people military, force base air'). Example: "With friends as part of a rescue party escaping some **building that was under attack**. I had to find certain items to bring along and had to keep doubling back to find some straggler."
  * **Natural places** ('night owls, woods night bunch, outside woods night'). Example: "A bunch of people were outside some **woods** at night. somthing scary was going on, and annette gave me her kid to get it away from there…"
  * **Fictional characters and Creatures** ('monster,' 'vampire,' 'werewolf,' 'dracula justice,' 'running vampires,' 'demon,' 'dragon,' 'land sharks,' 'alien outer space,' 'awaken, zombie awakened zombie'). Example: "I dreamed that I was in church when a **scary monster** came in and started attacking people. Most people ran out of the church, but I couldn't run because I had a broken leg, so I hid underneath one of the pews." Example: "I was with a friend at the mall. Suddenly, I was being chased by **Dracula** through a mall and was hiding in a bathroom trying to escape from him."
  * **Natural phenomena** ('tornado,' 'twister, inevitable end world,' 'flowing lava,' 'streak lightning coming, thunder cloud, cloud perpetual streak, coming also tornado'). Example: "I was somewhere in the tropics, and than all of a sudden this huge **tsunami wave** came." Example: "I was with a man and we were running from the **inevitable end of the world** in sweltering LA heat. I wasn't quite sure what it was that was coming, an **asteroid or tidal wave** or something. The feeling was terrifying but I was somehow relaxed even while I was running. Example: "I was running away from a **thunder cloud** with a perpetual streak of lightning coming out of it and also a tornado. I was all by myself and scared to death."
  * **Distant Non fictional characters** ('some bad men,' 'some guy,' 'agent corrupt,' 'deranged serial killer, pedophiles deranged,' 'began chasing, man began chasing, man began, unknown man began,' 'away bad guys, bad guy,' 'black coat man,' 'man,' 'chased somebody something, chased someone know, somebody something chased, chased somebody,' 'something murderer,' 'chasing someone chasing'). Example: "All the time, I try to run away from **a man**. I run very fast, almost flying." Example: "I was being chased **by somebody or something** and running from building to building…"
  * **Non fictional characters – defined** ('Morgan's house,' 'jason bateman')
  * **Relatives, family members** ('sister,' 'stepfather'). Example: "I was hiding Fromm **sister**. I think she was going to hurt or kill me."
- **Where am I fleeing to?**: We identified 6 semantic categories, as listed below:
  * **Public urban places** ('ballroom electric park,' 'abandoned shed,' 'neighborhood street,' 'underground motel bikers,' 'theater anymore, working theater, working theater,' 'school lab' 'school,' 'old neighborhood,' 'park park park ivy hill park, park park, park country park,' 'corner building,' 'library,' 'vermont church, church church,' 'place café place, café place café, allies cafe, cafe allies café,' 'high school,' 'stadium dark,' 'store back,' 'graduation graduation university, fredonia state, university homecoming,' 'village small english,' 'mall'). Example: "I'm in **a school**, in a classroom, hiding from the Devil amongst the students." Example: "/i/n one part, running around, hiding **in the library**. lady up front asks me not to run…." Example: "The I darted over **to church** ran inside and asked to use the phone because someone was outside trying to hurt someone. I felt a little freaked. I was thinking "I've got to get out of here."
  * **Places of safety** ('safety distant,' 'parents parents parents, parents parents (-) subterranean room,room subterranean,subterranean room subterranean,room subterranean room'). Example: "I finally killed the enemy but was afraid that his companions would kill me. Suddenly a young man dressed in a Tyrolian alpine costume appeared from the room and led me **away to safety**." Example: "Running through a destroyed city with my family running behind me. We get **to shelter** and my parents are there ahead of time…"
  * **Places in nature** ('nitt mountains,' 'cave,' 'shallow lake beautiful, lake beautiful blue,' 'large forset,' 'sankey valley woods, wood, place running, running straight line,' 'mountains mountains mountains, top mountain,' 'field across'). Example: "While I was roaming in the jungle I was observing the natural elements. At that time, I saw a big scary snake that was coming towards me. I started running away from there and the snake was chasing me. I reached a place that looked like a **cave** and I went down into it

and felt safe." Example: "We were discovered when I filled a tax form. We escaped out of a window and ran **into the woods** with the police in pursuit…." Example: "I was being chased by a bear in the forest going **down a hill**." Example: "Someone is chasing me. I run away, **across a field**."

* **Cities and countries** ('chicago,' 'grenada,' 'new York,' 'england,' 'unseen force, mexico,' 'italy rome,' 'tokyo, dinosauers,' 'friend Jamaica,' 'australia Victoria,' 'Ireland,' 'europe,' 'new zealand,' 'california,' 'city,' 'sao paulo'). Example: "I'm running away from somebody (the law? trying to kill me?). I have everything packed so I can leave at a moment's notice. in **Grenada**, I'm hiding out in the "projects." Example: "we?re secretly fleeing the city in a Jeep type car after being chased by some hostile people? I don?t know them? and we?re making for a very remote part of **New Zealand**."
* **Private places** ('home,' 'friends apartment,' 'mothers house,' 'yard back yard,' 'bedroom window,' 'bedroom,' 'grandparent home,' 'loft attic,' 'house grandma, grandmother house,' 'bathroom,' 'dad room mom, parents room, room mom dad,' 'friend house,' 'treehouse finn,' 'jamba juice jamba, mom room, mom room mom, room mom room (-)', 'country empty house,' 'weli house house, house house house, floor hom,' 'room four, see stars room, room four corners, ceiling could (-', 'mom boss mom, office mom office, office mom mom, mom,' 'beachy looking house,' 'kitchen get money, kitchen kitchen kitchen, kitchen told mom, kitchen told,' 'room room room, unit unit, room room, unit unit unit,' 'house parents,' 'kitchen get money,' 'house friends, house friends house, friends house,' 'swig front dorms, dorms,' 'room brother,' 'parents home,' 'room used live, live old room, abandoned cabin abandoned, old room used,' 'bathroom airport,' 'aunt house,' 'moms room,' 'hotel hotel, hotel room hotel, room hotel room,' 'farm house,' 'someone house, house person'). Example: "I was walking home down a country road at dusk when an unknown man began chasing me. I was terrified and ran furiously **toward my home**. I managed to get in the back door and lock it." Example: "I think there was a bomb threat. So, we ran too and broke into a **nearby apartment** to hide out." Example: "Being chased by a wolf, I ran into **my grandparent's home**. But the wolf was able to get into house where he began the chase anew." Example: "I was being chased by a little doll with a spear and when I went **into a kitchen** there were a bunch of people having a seance so I sat with them to hide and when I looked in back of me there was a hand coming out of a toaster to get me." Example: "I went **from friend to friend's house** to try and hide from my dad." Example: "I was trying to run **to my parents house** because there were tornados all around me. My parents house was the only thing that could save me." Example: "I think I see two large black roaches in the foyer. There is still a lot of clutter. Mom has a vacuum cleaner there. I run **to the kitchen**." Example: "I leave the house and run down the street **towards my friends house** and I can hear it's feet pounding behind me and it's antennae hitting my back." Example: "I was being chased by Dracula through a mall and was hiding **in a bathroom** trying to escape from him."
* **Vehicles** ('truck behind,' 'armored car,' 'camp trailer place, summer camp trailer,' 'begin drive children, got car, car drive, get car (-) 'house fraternity, maui maui, fraternity house fraternity, maui,' 'bus bus alviso, alviso apartments bus,' 'white car,' 'car'). Example: "Then a guy got beheaded and everybody was screaming. I hid **behind a truck**, but he saw me and started swinging his machete at my head. I woke up just before it hit me." Example: "A serial killer lured me into his house, acting like he liked me. But then he tried to kill me but I escaped **to my car** at the bottom of the hill of my house."

- **Who/ What am I fleeing from?** The initial place of escape can we manually abstracted into the following 8 semantic categories:
  * **Family and Close relations** ('moe dad, dad dad mo,' 'nana mother,' ('quynh hoa grandmother, aunt quynh, aunt,' 'son son, youngest brother friend,' 'husband kevin left,' 'partner, partner my, bear polar bear,' 'college ex boyfriend,' 'husband,' 'girl friend,' 'old mother, year old mother, 95 year, mothers mother'). Example: "I dreamed that **my dad** was trying to find me. I was hiding from him because I had done something wrong but I can't remember what it was." Example: "I am in a locked room an I realize my **husband** wants to kill me, he is planning how to do it, he is going to shoot me, I am looking around how to get out of the room…"
  * **Distant Non-fictional characters** ('guy random homeless,' 'land owners oriental,' 'dorm girl, girl katie,' 'bad guys group, guys group,' 'burglar,' ' evil men,' 'catch someone, catch someone trying, trying,' 'skinned woman following, dark skinned,' 'slender man,' 'someone going kill,' 'intoxicated older,' 'man long knife, giant black widow, long knife hand, long knife,' 'Asian man,' 'guys bad guys, bad guys,' 'girls two black, black girls two,' 'man broken house, broken,' ' person chasing mean, old person chasing,' 'guys bad guys, bad guys, bad,' 'man kill, kill man kill,' 'bangers gang bangers,' 'guns people guns,' 'third one go, pursued someone vehicle, pursued someone,' 'guy kidnapping women, prostitution,' 'chased somebody something,' 'bad guys wild, guys wild animals,' 'done something mob,' 'chasing people chasing, people chasing,' 'woman ugly old, old woman old,' 'guy random homeless). Example: "I was trying to escape **some evil men**. I hid behind a dog." Example: "Came out of mall and walked a few blocks when **three young men** stopped me and started hitting me. I jumpt to go down a hill." Example: "**A man** was trying to catch me and i was so scared i would sleep with my mother and she made me feel better." Example: "I don't remember very much of it. We were running away **from someone** who was going to kill us." Example: "I'm at a house, **some bad guys** take it over, people are held hostages….I flee to the back yard."
  * **Fictional characters** ('monster would, would get,' 'darth vaders darth, darth vaders,' 'hiding devil amongst, evil monster evil, devil amongst students, monster evil monster,' 'santa santa dad,dad dad father' non related answers, 'monsters would chase, mermaids chasing,' 'aliens scary,' 'monsters,' 'witch, maya, witch witch,' 'teddy bear, giant teddy,' 'giant bug, millions, big monster creature, bug,' 'monster godzilla, king kong type,' 'mutant zombie terrorist, zombie apocalypse,' 'castle chase defenseless, chase defenseless daughter, vampire'). Example: "**A monster** is chasing me. It has the shape of a human, but in reality it is a monster." Example: "I am being chased by many **Darth Vaders**." Example: "I'm in a school, in a classroom, hiding from the **Devil** amongst the students." Example: "I was running from **a monster**and fell off a cliff and saw myself splat on the rocks below the cliff." Example: "**Aliens** secretly harvesting humans under the guise of bring in utopia but I get suspicious and run away with the girls." Example: "A **witch** living either in my bedroom closet or under my bed. A scary hand coming out of the

closet to slide the door open…eek!" Example: "There was a **giant teddy bear** chasing after me." Example: "Me and my fellow soldiers fought of **mutant zombie** terrorist hordes and I had to street style race hellish monsters to get home to save my girlfriend and her family before the monsters made it there." Example: "My family and I were trying to escape from **vampires**, and I felt like I didn't have to run anymore and not be so afraid, because it was just a dream."

* **Animals** ('black bear big, black bear,' 'snake chasing,' 'big ostrcihes big, ostrcihes,' 'pig, guenia pig guenia, pig guenia, guenia pig,' 'bees swarm bees, bees swarm,' 'fox,' 'leopard snow, leopard, snow leopard, guys trying molest,' 'tiger big tiger, tiger,' 'shark big shark,' 'wolf,' 'panther child, black panther,' 'spiders lots spider,' 'wild boar, rabbits rabbits,' 'chasing gorilla chasing,' 'mean grizzly, grizzly bear mean,' 'dinosaur game,' 'beast chasing everyone, panther jumped rows, beast chasing, okra started chasing,' 'bat bat bat, death death death,' 'alligator large alligator'). Example: "A huge, **black bear** was sleeping right in the front of our house. I woke it up by coming downstairs and it started to head back into the woods." Example: "I am being pursued by **a snake**. It moves as if flying on the air. I feel anxious. I escape." Example: "Grandma and me and **2 big ostrcihes**.They were trying to eat me and grandma got in her limo and drove off." Example: "About 20 years ago, I would have a reoccuring dream I was outside being chased by **a red fox**." Example: "Being chased by a **wolf**, I ran into my grandparent's home." Example: "Then I see shadows, and a great **black panther** appears in the room. I get very frightened and run across the room. I hide behind the sofa." Example: "So I started to run down the street and **a gorilla** started to chase me but I made it back in time."
* **Groups of power** ('police,' 'boss,' 'mayor county,' 'chasing sheriff,' 'officers one security, policeman policeman,' 'teacher really, school teacher,' 'soldiers german soldiers'). Example: "I was Running from the **police man** when I did nothing don't know what I was running dinner with my grandma's house he caught me …" Example: "It was something about me hiding from my **boss** at work." Example: "The mayor our county, province, came to our house, and I ran away, because I don't like people like him, **the authorities.**" Example: "I was Running from the **police man** when I did nothing don't know what I was running dinner with my grandma's house he caught me I said I was running to get to my girlfriend…" Example: "I remember running through my family farm being chased by **German soldiers**. I jumped from the top of a gravel pit that's by the edge of a hayfield."
* **Known figures** ('mel gibson trying, chasing lava chasing,' 'hugh jackman hugh, captain jack sparrow, jackman hugh'). Example: "My boyfriend and I were trying to escape out of a bathroom window because **Mel Gibson** was trying to kill us." Example: "She asked me why I tweeted about her and put instagram pics of her up. But then **Hugh Jackman** was chasing us."
* **Violent environments** ('holocaust,' 'nazi checkpoint nazi, nazi checkpoint,' 'streets war, war streets, war streets war, streets war streets (-) europe europe'). Example: "I dream that I wrote a script about a Jewish man trying to escape the **holocaust** by hiding in a book." Example: "I am a Jew and I've run to a **nazi checkpoint**. There is a stone opening, stairs and rails, a descent. I jump into it and find my way down rails to a secret passageway…" Example: "I am in an industrial city in Eastern Europe. There is a **war in the streets**. To escape the violence I am jumping from rooftop to rooftop with fifteen other people."
* **Opressors** ('russians enslaving, russians enslaving'). Example: "The **Russians** were invading our town and I was trying to find shelter around the high school. I could hear gunfire and knew they were close."

- **Who/ What am I fleeing to?** The final point of safety can be summarized in one semantic category:
  * **Friend's or Family's homes**: Someone's home ('fathers house, house dad house,' ' neighbor house, house stranger house,' 'parents children, parents parents,' 'mom grandparents, grandparents mom, mom grandparents,' 'mom house, mom house mom, mother,' 'moms moms mom,' 'parent, parents,' 'mom house,mom house mom,' friend house,house friend'). Example: "I have divorced parents Mother is mean in the first day if school so I run to my **fathers house.**" Example: "My sister and I got off the bus and ran into a **stranger's house** to call home." Example: "I tried to scream but couldn't get my voice loud enough and I knew no one would hear me. I ran to the other side of the house and woke my **mom and grandma** up. I was frantic and trying to explain to them what happened." Example: "I backed up hurriedly, turned around on the front lawn of the closest house knocking down several little fences bordering the walk, and drove away from the scene. I later found myself in the garage at my **mother's home**." Example: "A scary character from cartoons was chasing me around a skating rink and I had to skate very hard to keep from getting caught. **My mom was there and came and helped me**." Example: "I have divorced parents Mother is mean in the first day if school so I run to **my fathers house.**" Example: "My sister and I got off the bus and ran into a **stranger's house** to call home."

- **How am I fleeing?** The strategy of escape showed 11 distinct semantic categories:
  * **Primary movement**:

    Jumping and flying ('escapes jumps pool,' 'could fly, endless rolling, jumped jumped, could fly halls, 'jumped window, window jumped,' 'jump bike, bike get away, trying catch vineeta, bike trying,' 'fly, overhead, overhead heading, overhead heading towards,' 'away tried fly, start fly away, fly away,tried fly away,' 'wall flying, wall flying outside'). Example: "**Jumping** into giant pools and hiding deep in the water from him but he would find us." Example: "I was running from a monster and i **ran straight off a cliff** and i watched myself falling and splat on the rocks below." Example: "We are on the third floor and I finally decide to **jump from** the window, Somehow I manage to land gracefully and take off running, I remember being impressed by my own getaway skills." Example: "I went into my bathroom threw a towel over my back and used one of my birthday candles to make a wish that I could **fly**. I **blasted off into the sky** catching up with the space ship."
    Escaping ('escaped lab kept, lab kept made, lab,' 'escaped could, away escaped, kind ran away, ran away'). Example: "Then the baby and I must **escape**. We hide in the movie crowd." Example: "I was being chased by a big baseball. It was going to hit me but I was able to hit a homerun and **get away.**" Example: "With friends as part of a rescue party **escaping** some building that was under attack."
    Running ('ran another woman, us someone chasing,' 'way home ran, ran way home, home ran way,' 'sneak, escaped running top, bolted back store, bolt,' 'door ran, door ran door, ran away house, ran door ran,' 'nephew ran behind, nephew ran, tree

nephew, ran behind huge (-),' 'running away creature, away creature running,' 'street running, street toward, street toward house, run neighbors,' 'ran house hila, took ran house, back door,' 'run home, ran home, home ran home, ran home ran,' fast could ran, could ran fast'). Example: "There's a part of me that wonders if I should have stayed and waited to be questioned with the possibility of being freed. But this is the choice I've made - **to run. So I continue running**." Example: "A lion is chasing me, and **I rush into** my house, and slam the door. I feel secure." Example: "I was walking home down a country road at dusk when an unknown man began chasing me. I was terrified and **ran furiously** toward my home. I managed to get in the back door and lock it." Example: "War going on, big fancy dinner/banquet, woman with a pearl in her hair, **running through** a gravel parking lot."

Climbing ('climbed tree, tree climbed tree'). Example: "We stayed out too late and like the only way you could escape was by **climbing up** a tree, but like not everyone could climb it."

Swimming ('bridge river fall, fall bridge, fall bridge river, river swam buoy'). Example: "Jacob and I continued **swimming** furiously towards the boat. We were almost there when I couldn't shake the feeling the man had just hopped on a speedboat and was now in the water."

Hiding ('hiding fromm, hiding behind, hid behind, hide'). Example: "I dreamed that I was in church when a scary monster came in and started attacking people. Most people ran out of the church, but I couldn't run because I had a broken leg, so I **hid** underneath one of the pews." Example: "I ran through the house and found a secret attic **hideaway** from a closet." Example: "I killed another woman in this dream, and I am escaping from the police. I am **hiding** in the underground, and I am living there."

* **Secondary modes of transport**:

Driving ('got drove away, drove away, away walking back, ran car escaped,' 'jeep type, car jeep, jeep like car,' 'begin drive children, got car, car drive,' ' riding car sidewalk, car sidewalk'). Example: "I took someone who was in the house with me, packed some pie crusts in a backpack and tried to **force a car** on the freeway to pick us up before the zombie got to us." Example: "I don?t see his face, we?re secretly fleeing the city in a **Jeep type car** after being chased by some hostile people." Example: "Russian theme park turna into construction complex that we get caught at. Cops detain us and ontals to bonobfor a while he offers an OPPERTUNITY. We escape in **my car**."

By plane ('disaster plane dad, plane'). Example: "…mother evil, trying to kill my friends by cursing me trying to escape natural disaster on a **plane** with dad." Example: "Then I got even angrier and decided that there is no such thing as fate, and that I could save them. All of a sudden I was on a **plane** and I was the copilot, my friend Dan was the pilot."

By boat ('leave one boats, boats leave one, ship ship'). Example: "Being chased by big yellow blobs (I could smell them). **Floating** down a river to get away and when that didn't work flying. But flying low to the ground and unable to gain altitude so the possibility of being caught was still there." Example: "I crawled on the ground until I passed them then left in a **boat**."

By stairs ('stairs ran, ran flights stairs, flights stairs go, stairs go home'). Example: "I tried to escape him and fell down on the **stairs**." Example: "Apparently this was not what he had come for. He started to come toward us and while screaming, we ran out the back way and up the **stairs** where I used to live."

By train ('train catch train, train'). Example: "Last thing I remember we were hanging on to **railroad tracks** running away from Hugh Jackman."

A.3. A deeper analysis of the Flying Situation Network

In the *Flying* situation, we were primarily concerned with the location, the presence or absence of other characters, and how the dreamer was flying. The network is broken-down into three color-coded node groups, as follows: (i) The group of green nodes corresponds to the question *How am I flying?*; (ii) The group of purple nodes attempts to answer if the flying is a solitary or group activity and is thus formulated as *Whom am I flying with?*; (iii) The group of turquoise nodes represents answers about location of flying and is queried as *Where am I flying?*; In terms of location, we identified two categories:

- **Where am I flying?** The settings can be grouped into 2 semantic categories:
  - **Urban spaces** ('restaurant,' 'buildings mountains, pyramids Egypt,' 'old neighborhood,' 'giants park, chicago, city field,' 'grade school,' 'building next, cathedral,' 'road right, avenue,' 'house street district, house neighbors'). Example: "I notice mosquitoes in the air. I begin to fly with them, noticing **buildings and streets below**." Example: "Flying **over my old neighborhood** and noticing changes in landscape, land cleared/grown over." Example: "Flying **over my school** on my own power." Example: "I flew and was climbing **inside a building** above where Levi was standing trying not to be seen by him so he wouldn't realize I'd been flying." Example: "There was a big storm in the city and I started to fly**over one of its larger avenues** which became a river with two different flows."
  - **(ii) Natural environments** ('mountain tall,' 'trees,' 'blue sky,' 'ocean,' 'space, dead space', 'hawaii,' 'clouds,' 'mountains'). Example: "I was flying **over a tall mountain** and was looking from above on my classmates sports activities." Example: "I would start flying (floating upright and without effort) until I had cleared **the trees** (maybe 100–150 feet up)…" Example: "Flying through my home and **out into the blue sky**." Example: "I am flying as high as I can get, to the sun. At first I go slowly, but then, like a rocket, faster and faster. I penetrate **the layer of clouds** and forget about everything that has happened…" Example: "I use a piece of plastic to fly **over the ocean…**" Example: "I am able to fly, and I fly all over the globe, especially to **exotic islands**, where I stay most of the time." Example: "In mountains I was flying (with my hands streched out) **from one hill to next** to avoid from being captured by some one…"

- **How am I flying?**
  - **Bird-like imitations** ('bird flew, like bird,' 'fly, flying trapeze, air flying,' 'fly halls away, high enough,' 'balloon basket balloon,' 'float, floating near, floating near top,' air suspended air, air flying air,' 'apollo like rocket '). Example: "I had a dream last night. I was **turning to a bird** and fly in sky it felt so good that I could flying like that even though it was just a dream." Example: "I am running so **silently** so as not to alert anyone my feet barely touch the ground and it is like I am flying, **skimming the ground**."

- **Whom am I flying with?**: The posts answering this question were limited to *air traveling* that occurred in the company of others. Example: "**Ava, me and my mom** were on a flight back from Russia. Plane crashed, and my mom didn't make it. Then we called into some therapist radio show and told them the sob story. Then I worked at camp." Example: "Last night I had a dream that I had to go to Japan **with my mother** for some reason. We drove to the airport, which was at my high school, and I was able to pack everything into this small DVD case." Example: "I was traveling in an airplane to the state of Rio Grande de Sulin southern Brazil. With me were **my aunt and two female cousins**." Example: "I dreamt that **Caitlin and I** landed in Toronto and that we were going to meet up with Chris Taylor…"

  It is worth mentioning that flying with one's body was manually detected in close readings, where we found 541 posts related to flying/levitating/hovering. Most of the posts describe flying as a *solitary activity*.

A.4. A deeper analysis of the Falling Situation Network

In the *Falling* situation, we were primarily concerned with the trajectory of falling (*from-to*) and how the dreamer was falling. The network is broken-down into three color-coded node groups, as follows: (i) The group of yellow nodes corresponds to the question *How am I falling?*; (ii) The group of turquoise nodes represents the initial location of the fall and it is thus formulated as *Whom am I falling from?*; (iii) The group of red nodes represents the final location of the fall and it is thus formulated as *Where am I falling to?*; In terms of initial location (*Where am I falling from?*), we identified 2 semantic categories:

- **Where am I falling from?**
  - **High man-made structures and Aeroplanes** ('fall, fell ground (-) airplane airplane,' 'bridge, bridge, bridge,' 'high place,' building tall'). Example: "We were flying through some clouds and Richie Rich (the comic book character) **pushed me out of the airplane**. I fell through the air for a long time." Example: "/w/as **falling down a really long mountain** on my white horse unicorn and landing in water. I was falling forever and it was a freeing but terrifying feeling…" Example: "…falling **from a ferris** wheel into the river at our local amusement park." Example: "…falling **through a wooden bridge** into a gorge." Example: "… falling **off a ledge** into deep, dark, water that has no bottom." Example: "A small animal is running with me. I fall **off a bridge** into the river. There are some high cliffs nearby and someone is watching me as I fall…" Example: "Falling **off a high building**, hitting the ground. I woke up when I hit the ground." Example: "I was falling **off a skyscraper** and Robin caught me (I had a crush on him at the time). Batman was standing beside him…"
  - **Highlands** (cliff steep cliff, cliff high,' 'mountain climbing mountain,' 'mountain tall, tall mountain,' 'hill hill, steep hill holiday'). Example: "Repeatedly dreamt I was on top **a high cliff** and jumped off – beautiful feeling and enjoyed the falling – at the bottom was a Kroger grocery store." Example: "I was running from a monster and **fell off a cliff** and saw myself splat on the rocks below the cliff." Example: "Being chased by a lion, I think, and **falling off a cliff**, but waking up just as I wondering what it was going to be like when I hit the bottom." Example: "Falling **off a cliff** on a mountain near where I lived."
- **Where am I falling to?**: For this query, we identified 2 semantic categories:
  - **Water** ('tank water, water water, water tank water,' 'water start, go water start, water start sink,' 'ocean,' 'patch ice patch, water cold water,' 'deep ocean falling, ocean falling deep,' 'white water river'). Example: "I've actually fallen from very high distances, hit the ground, gone through and **ended up in water**. But I can still breath and finally end up in air again. Then I start flying. It's a very cool dream …" Example: "After flying for some time I got tired and **fell in water**." Example: "I fell **into the ocean** off of the porch and the ocean was really choppy." Example: "I suddenly felt myself falling into space without even coming to a halt. Suddenly I plunged into the **icy cold water**. I could not swim. I cried for help…" "I'm dropped down **at the bottom of the sea**."
  - **Floor or under the floor** ('tunnel endless spiral, underground tunnel,' 'oncoming traffic, train trestle,' 'ground floor, floor hatch,' 'hole'). Example: "Falling **in a hole**." Example: "There is **a hole** and I descend into it." Example: "I am running along a path when I suddenly fall **into a hole** and at the bottom are all kinds and sizes of snakes." Example: "I looked at him, and with the power of mind, he **pushed me to the hole** from which he had originally come…" Example: "(I frequently dream that) I am falling straight down **into a deep shaft or well**. I start out walking along a street that seems very safe. Then, suddenly, the earth opens up and I start to all. I am very scared but there is nothing I can do to stop falling. And I can not see the bottom of the shaft." Example: "The only part of it that I remember is me in midair falling from the second floor (inside) of my childhood home **to the ground floor**, with me waking up as I hit the ground."

    Some contextual groups ('slumber,' 'deep sleep,' 'fall sleep, fell asleep') suggest the **metaphorical fall**, for example: "When I first **fell asleep**, I fell into a really deep sleep" or "I had a dream and it was set in school. **I fell asleep** in class and woke up in a prison for mutated monsters."
- **How am I falling?**: There is one major group that emerges from this query, which can be semantically categorized as follows:

  Accidents and Escapes ('sandy slope, giant hill sand,' 'bridge covered bridge,' 'fell cold water, somehow end water, broke fell cold,' 'throw fall, end bed fall,' 'stairs falling,' 'falling ferris, falling overboard,' 'floated roof house, falling skyscraper, skyscraper, falling high building,' 'cliff fell, cliff falling,' 'deep shaft, fell well,' 'water start, go water start, water start sink,' 'hole never ended,

hole endless hole'). Example: "Falling from a covered bridge that I was **trying to cross** because all my friends were on the other side." Example: "She took me across a **bridge which was decrepit and falling apart. I knew it wouldn't hold out** and it broke under me and I fell into cold water." Example: "My dad and I went to someplace where you got to be on top of a roof but it was very slanted and steep. I was **wearing slippery pants** also, so I kept on falling and almost dropped my phone and camera. But this place was so high up you could see almost all of America." Example: "I was running through a forest and **I tripped**. I fell into this deep hole and I was alone for a long time…" Example: "...i had to jump out of his top window **with a blanket as a parachute** cause green, yellow and white ninjas werre climbing up the side of a cliff to attack. i ended up sailing over the ocean in greece and getting picked up by a fisherman." Example: "I was **running from a monster** and fell off a cliff and saw myself splat on the rocks below the cliff." Example: "I was **running from a monster** and fell off a cliff and saw myself splat on the rocks below the cliff."

References

Azunre, P. (2021). *Transfer learning for natural language processing*. Simon: Schuster.

Bostock, M. (2012). *D3.js - data-driven documents*. http://d3js.org/.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners.

Buck, C., Bulian, J., Ciaramita, M., Gajewski, W., Gesmundo, A., Houlsby, N., & Wang, W. (2017). Ask the right questions: Active question reformulation with reinforcement learning. *arXivpreprint arXiv:1705.07830*.

Bulkeley, K. (2016). *Big dreams: The science of dreaming and the origins of religion*. Oxford University Press.

Calkins, M. W. (1893). Statistics of dreams. *The American Journal of Psychology, 5*(3), 311–343. Retrieved June 20, 2022, from http://www.jstor.org/stable/1410996.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding.

Domhoff, G. W. (2003). *The scientific study of dreams: Neural networks, cognitive development, and content analysis*. American Psychological Association.

Domhoff, G. W., & Schneider, A. (2008). Studying dream content using the archive and search engine on dreambank. net. *Consciousness and Cognition, 17*(4), 1238–1247.

Eggan, D. (1952). The manifest content of dreams: A challenge to social science. *American Anthropologist, 54*(4), 469–485.

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

Fogli, A., Maria Aiello, L., & Quercia, D. (2020). Our dreams, our selves: Automatic analysis of dream reports. *Royal Society Open Science, 7*(8), 192080.

Griffith, R. M., Miyagi, O., & Tago, A. (1958). The universality of typical dreams: Japanese vs. americans. *American Anthropologist, 60*(6), 1173–1179.

Grootendorst, M. (2020). Bertopic: Leveraging bert and c-tf-idf to create easily interpretable topics. https://doi.org/10.5281/zenodo.4381785.

Gustavsen, A., J., Pai, Shraddha, Isserlin, Ruth, Demchak, Barry, Pico, & R., A. (2019). Rcy3: Network biology using cytoscape from within r. F1000Research. https://doi.org/10.12688/f1000research.20887.3.

Gutman, M., HOLUR, P., & Bulkeley, K. (2022). Dreams and archetypes. https://osf.io/e9jnv/?viewonly=3bfdb80f287e44f2a67c0a8d046995ca.

Hall, C., & Van de Castle, R. (1966). The content analysis of dreams.

He, L., Lewis, M., & Zettlemoyer, L. (2015). Question-answer driven semantic role labeling: Using natural language to annotate natural language. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 643–653).

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach.

Maggiolini, A., Di Lorenzo, M., Falotico, E., & Morelli, M. (2020). The typical dreams in the life cycle. *International Journal of Dream Research*, 17–28.

Martin, J. M., Andriano, D. W., Mota, N. B., Mota-Rolim, S. A., Araújo, J. F., Solms, M., & Ribeiro, S. (2020). Structural differences between rem and non-rem dream reports assessed by graph analysis. *PloS One, 15*(7), e0228903.

Mathes, J., Schredl, M., & Göritz, A. S. (2014). Frequency of typical dream themes in most recent dreams: An online study. *Dreaming, 24*(1).

Mausam, M. (2016). Open information extraction systems and downstream applications. In *Proceedings of the twenty-fifth international joint conference on artificial intelligence* (pp. 4074–4077).

McInnes, L., & Healy, J. (2017). Accelerated hierarchical density based clustering. Data Mining Workshops (ICDMW). In *2017 IEEE International Conference on* (pp. 33–42).

McInnes, L., Healy, J., & Melville, J. (2020). Umap: Uniform manifold approximation and projection for dimension reduction.

Mota, N. B., Furtado, R., Maia, P. P., Copelli, M., & Ribeiro, S. (2014). Graph analysis of dream reports is especially informative about psychosis. *Scientific Reports, 4*(1), 1–7.

Nielsen, T. A., Zadra, A. L., Simard, V., Saucier, S., Stenstrom, P., Smith, C., & Kuiken, D. (2003). The typical dreams of canadian university students. *Dreaming, 13*(4), 211–235. https://doi.org/10.1023/B:DREM.0000003144.40929.0b

Palmer, M., Gildea, D., & Xue, N. (2010). *Semantic role labeling (1st)*. Morgan: Claypool Publishers.

Pesonen, A.-K., Lipsanen, J., Halonen, R., Elovainio, M., Sandman, N., Mäkelä, J.-M., Antila, M., Béchard, D., Ollila, H. M., & Kuula, L. (2020). Pandemic dreams: Network analysis of dream content during the covid-19 lockdown. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.573961

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer.

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). Squad: 100,000+ questions for machine comprehension of text.

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bertnetworks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. https://arxiv.org/abs/1908.10084.

Ruppenhofer, J., Ellsworth, M., Schwarzer-Petruck, M., Johnson, C. R., & Scheffczyk, J. (2016). *Framenet ii: Extended theory and practice*. International Computer Science Institute (tech. rep.).

Schredl, M. (2010). Characteristics and contents of dreams. *International Review of Neurobiology, 92*, 135–154.

Schredl, M., & Hofmann, F. (2003a). Continuity between waking activities and dream activities. *Consciousness and Cognition, 12*(2), 298–308.

Schredl, M., & Hofmann, F. (2003b). Continuity between waking activities and dream activities. *Consciousness and Cognition, 12*(2), 298–308.

Schwartz, S. (2004). What dreaming can reveal about cognitive and brain functions during sleep? a lexico-statistical analysis of dream reports. *Psychologica Belgica, 44*, 5–42.

Seligman, C. G. (1923). Type dreams: A request. *Folklore, 34*(4), 376–378.

Suchanek, F. M., Kasneci, G., & Weikum, G. (2007). Yago: A core of semantic knowledge. In *Proceedings of the 16th International Conference on World Wide Web* (pp 697–706). https://doi.org/10.1145/1242572.1242667.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need.

Windt, J. M. (2010). The immersive spatiotemporal hallucination model of dreaming. *Phenomenology and the Cognitive Sciences, 9*(2), 295–316.

Windt, J. M. (2010). The immersive spatiotemporal hallucination model of dreaming. *Phenomenology and the Cognitive Sciences, 9*(2), 295–316.
Yin, F., Shen, H., He, Y., Wei, Y., & Cao, W. (2013). Typical dreams of "being chased": A cross-cultural comparison between tibetan and han chinese dreamers. *Dreaming, 23*(1), 64.
Yu, C. K.-C. (2008). Typical dreams experienced by chinese people. *Dreaming, 18*(1), 1.