A Behavioral Data Approach to Understanding Time-Blindness in ADHD

Name: Amy Yu
USC Email: yuamy@usc.edu
USC ID: 5974518509

1. Description

Time-blindness is a commonly reported but under-researched cognitive difficulty associated with Attention-Deficit/Hyperactivity Disorder (ADHD), characterized by distorted time perception, chronic lateness, and difficulty planning or transitioning between tasks. This project investigates how time-blindness, commonly associated with ADHD, is framed and discussed in psychoeducational online articles. Time-blindness refers to persistent difficulties in perceiving, estimating, and managing time, which can affect punctuality, task initiation, and transitions between activities. By applying text-based data analysis techniques to expert-written content, this project aims to identify recurring themes, emotional tone, and dominant vocabulary used to describe time-blindness. The findings offer insight into how this phenomenon is communicated to the public and how challenges related to time perception are contextualized within broader discussions of ADHD.

2. Data

2.1 Data Sources

Data for this project were collected through static HTML scraping of publicly available psychoeducational articles related to ADHD and time management. All sources are reputable health or ADHD-focused platforms commonly referenced for educational purposes. No user-generated or personally identifiable data were collected.

The six articles included in the final dataset were sourced from:

- **ADDitude Magazine**
  - https://www.additudemag.com/how-to-plan-ahead-when-you-have-adhd-understand-time/
  - https://www.additudemag.com/slideshows/stop-wasting-time/
  - https://www.additudemag.com/time-on-your-side/
- **PsychCentral**
  - https://psychcentral.com/adhd/why-are-people-with-adhd-always-late
  - https://psychcentral.com/adhd/time-management-tips-for-people-with-adhd
  - https://psychcentral.com/adhd/the-adhd-iceberg

These sources were selected because they explicitly address ADHD-related time perception, lateness, and executive functioning challenges, making them well-suited for a focused analysis of time-blindness.

2.2 Number of Data Samples

The final dataset consists of six long-form articles, each containing multiple paragraphs of explanatory and behavioral content. After cleaning and segmentation, the dataset yielded several hundred sentences and thousands of word tokens, providing sufficient textual material for frequency-based analysis, sentiment evaluation, and thematic clustering.

3. Data Cleaning, Analysis, and Visualization

3.1 Data Cleaning

The raw HTML content was first parsed using BeautifulSoup to extract article titles and main body text. The data cleaning process involved removing HTML tags, navigation elements, and unrelated page content, converting all text to lowercase for consistency, removing punctuation, numbers, and common English stopwords, normalizing whitespace and encoding issues, and tokenizing text into individual words.

The cleaned text was stored in structured CSV format in the data/processed directory to ensure reproducibility and separation from raw data.

3.2 Analysis Methods

This project employed frequency-based and unsupervised text analysis techniques commonly covered in class.

To examine vocabulary patterns, term frequency (TF) and TF-IDF (Term Frequency–Inverse Document Frequency) were computed to identify words that are both frequent and distinctive across articles. This allowed the analysis to surface concepts most strongly associated with time-blindness, such as lateness, planning, deadlines, and executive function. These results were later visualized in the top_words bar chart and the keyword_heatmap, which highlight dominant terms and their relative emphasis across articles.

To explore higher-level thematic structure, K-means clustering was applied to TF-IDF vectors. The number of clusters was selected based on interpretability and distribution of terms rather than formal optimization, given the small dataset size. Clustering enabled the grouping of articles and text segments into conceptual themes, such as emotional frustration, behavioral coping strategies, and structural time-management advice.

For emotional tone, VADER sentiment analysis was used to calculate compound sentiment scores. VADER was chosen due to its effectiveness on explanatory and emotionally descriptive text, making it appropriate for psychoeducational content that blends emotional validation with instructional guidance. The resulting sentiment values were summarized and visualized using the sentiment_scores distribution chart.

3.3 Visualizations

Several visualizations were generated using Matplotlib to support interpretation of the analytical results. The top_words visualization presents a bar chart of the most frequently occurring terms across all articles, revealing a strong emphasis on time awareness, lateness, planning, and task management.

The keyword_heatmap illustrates how selected time-related concepts, such as procrastination, routines, focus, and scheduling, vary in prominence across individual articles, highlighting differences in thematic emphasis between sources.

The sentiment_scores plot shows the distribution of compound sentiment values across articles, indicating that the overall tone of psychoeducational content is largely neutral with mild positive framing.

Together, these visualizations provide a clear, interpretable representation of how time-blindness is discussed across expert-written sources.

3.4 Hypothesis and Conclusions

Initial Hypothesis:
Psychoeducational articles on ADHD frame time-blindness primarily as a behavioral and executive functioning challenge, combining emotional validation with practical coping strategies rather than emphasizing clinical pathology.

Conclusions:
The analysis supports this hypothesis. The top_words and TF-IDF results show that discussions of time-blindness consistently center on planning, lateness, routines, and time awareness rather than diagnostic terminology. The sentiment_scores visualization indicates a predominantly neutral-to-positive emotional tone, reflecting an educational and supportive framing. The thematic patterns observed through clustering and the keyword_heatmap further demonstrate that expert content balances acknowledgment of emotional frustration with actionable strategies, reinforcing the idea that time-blindness is treated as a manageable behavioral pattern rather than a fixed deficit.

4. Changes from Original Proposal

The original proposal planned to include Reddit-based community discussions as a primary data source. However, due to significant changes in Reddit API accessibility and restrictions on third-party services such as Pushshift, Reddit data collection was ultimately removed from the project. To address this challenge, the project scope was refined to focus exclusively on expert-written psychoeducational content. This adjustment ensured reliable data access, ethical data usage, and full project feasibility while still addressing the core research question.

5. Future Work

With additional time and resources, this project could be expanded in several directions. Incorporating community-generated content from accessible forums or archived datasets would enable direct comparison between lived experiences and expert framing. More advanced topic modeling techniques such as Latent Dirichlet Allocation (LDA) could provide deeper thematic insight. Finally, expanding the dataset to include more articles or longitudinal content could support temporal analysis of how discussions around time-blindness evolve over time.

6. References

ADDitude Magazine. (n.d.). *Time management and ADHD*. https://www.additudemag.com/
PsychCentral. (n.d.). *ADHD and time management*. https://psychcentral.com/