Journal of Data Mining and Digital Humanities

Interpretable Socioeconomic Profiling: A Deep Dive Beyond Narrative Classification

Nahed Abdelgaber*[1], Labiba Jahan[1], Joshua R. Oltmanns[2], Mehak Gupta[1], Jia Zhang[1]

[1]Southern Methodist University, Dallas, TX 75205
[2]Washington University of St. Louis, MO 63130
nabdelgaber@smu.edu

Abstract

This paper develops a narrative-driven SES profiling approach that identifies underlying themes within life stories. Building upon our prior work published at NLP4DH 2025, we extend the SES understanding pipeline by integrating interpretability mechanisms such as theme-topic association heatmaps and SES-separated semantic visualizations. The profiling system operates by segmenting narratives into predefined themes, extracting representative topics using BERT-based embeddings, and assigning SES relevance scores. Although not designed for classification, the system demonstrates robust performance in out-of-distribution (OOD) evaluation: when used in a binary classification task, the profiling-based method achieved an accuracy of 59.4%, surpassing Random Forest (56%), the strongest classical ML baseline reported in our earlier study. This highlights the profiling system's potential to generalize across domains while offering interpretable insights into SES indicators embedded in personal stories. We release our code and visual analysis tools to support further research in narrative understanding and computational social science.

Keywords

Socioeconomic Status; Narrative Analysis; Interpretability

I INTRODUCTION

Socioeconomic status (SES) continues to be a significant determinant of life outcomes, influencing individuals' access to education, healthcare, employment, and overall well-being. While SES is traditionally measured using structured indicators such as income, occupation, and education, recent research has explored the use of language as a complementary source for SES inference. In particular, personal narratives provide a rich and nuanced window into individuals' lived experiences, revealing subtle indicators of their socioeconomic context.

Building on our earlier work Abdelgaber et al. [2025], which demonstrated the feasibility of classifying SES from life narrative interviews using transformer-based models, this paper introduces an expanded profiling system that moves beyond classification to provide interpretable, theme-aware representations of socioeconomic patterns. This system leverages the thematic structure of narratives to extract high-salience topics for each SES group and computes semantic similarity between unseen narratives and SES-specific topic sets.

The profiling pipeline begins by segmenting narratives according to predefined life themes—such as education, health, and personal satisfaction—identified in our previous study. Within each theme, we extract representative topics associated with each SES class using SBERT embeddings and a similarity-based selection mechanism. For any new narrative, the system computes narrative-level semantic similarity to the SES-specific topics, enabling an interpretable SES prediction based on top topic voting.

To enhance interpretability, we introduce several visualization tools that help explain the model's reasoning. These include heatmaps that align extracted topics with life themes, as well as low-dimensional scatter plots that reveal the spatial distribution of narrative and topic embeddings across SES groups. We also apply OOD evaluation to test the generalization of our topic-based inference pipeline.

By integrating thematic segmentation, semantic similarity, and visual interpretability, the proposed system offers a transparent and extensible approach to SES profiling. This work contributes to the growing body of research on the intersection of computational social science and NLP, highlighting the role of language as both a signal of social inequality and a tool for understanding it.

The remainder of this paper is structured as follows. Section II reviews prior research on identifying SES. Section III describes the dataset and preparation steps, and Section IV details our methodology, including the thematic filtering, SES-aligned topic modeling, visual interpretability tools, and out-of-distribution evaluation. Section VI provides a broader discussion of the findings, their implications, and outlines directions for future research. Section VII summarizes the study's key contributions. Section VIII discusses the ethical and societal impact of the work, and Section IX addresses the limitations of the current study.

II RELATED WORK

Understanding SES through language has garnered increasing attention in both computational social science and natural language processing (NLP). Prior research has leveraged diverse linguistic sources—such as social media posts Preoţiuc-Pietro et al. [2015], Field and Tsvetkov [2017], survey responses Volkova et al. [2016], and spoken transcripts from interviews or counseling sessions Buechel et al. [2019]—to infer SES or related demographic indicators. These approaches often rely on lexical, syntactic, or affective features, and typically use supervised classification frameworks to map these features to SES labels.

More recent efforts have incorporated pretrained language models (PLMs) such as BERT, RoBERTa, and DeBERTa to better capture semantic nuances in text. These models provide richer contextual embeddings that have shown significant improvements in tasks like social attribute inference, including age, gender, occupation, and income prediction Liu et al. [2019], Zhang et al. [2022]. Our previous work Abdelgaber et al. [2025] applied transformer-based models to life narratives and demonstrated that SES classification benefits from the depth of contextual understanding these models provide—particularly in free-form, subjective narratives.

Interpretability remains a core concern in SES modeling. While black-box models may perform well, they often lack transparency, which is crucial when dealing with socially sensitive attributes like SES. Topic modeling has served as one interpretability bridge, enabling the discovery of latent themes in text. Traditional methods such as Latent Dirichlet Allocation (LDA) Blei et al. [2003] have been widely used, but they often suffer from poor coherence and interpretability in small or noisy datasets. More recent methods, such as BERTopic Grootendorst [2022], leverage transformer embeddings and clustering to yield more semantically coherent and meaningful topics.

Beyond topic modeling, advances in model visualization and representation analysis—such as t-SNE, PCA, and attention heatmaps—have played a crucial role in making model decisions more transparent. These tools have been used in clinical NLP Zhang et al. [2020], psychological profiling Pennebaker and Chung [2013], and sociolinguistic research Hoyle et al. [2022], among others.

Despite these advances, limitations remain in combining interpretability, generalizability, and SES relevance. Most SES classification systems prioritize performance on in-distribution tasks and fail to offer insight into why a model makes a particular prediction. Furthermore, theme discovery in prior work is often driven by topic frequency rather than alignment with psychological or socioeconomic theory. The resulting models may not generalize well across contexts or populations.

While prior Digital Humanities (DH) research has employed interpretable models such as topic modeling and lexicon-based analysis to explore cultural or historical corpora, these methods often lack socio-psychological grounding. Our system extends this line of work by aligning thematic markers with SES-relevant theories, positioning it within current debates on explainable AI for socially sensitive domains.

Our proposed profiling system addresses these limitations by integrating PLM-based sentence representations with semantically grounded thematic segmentation, SES-specific topic filtering, and structured visualization tools. Unlike traditional SES classifiers, our system is designed to highlight interpretable SES-aligned themes across narratives and demonstrates robust OOD performance. This positions our work at the intersection of narrative understanding, SES modeling, and explainable AI.

III DATA

3.1 Data Overview:

The dataset employed in this study originates from the St. Louis Personality and Aging Network (SPAN) Oltmanns et al. [2014]. Over a span of 3.5 years, 1,630 older adults were recruited from households within a 100-square-mile radius of the St. Louis metropolitan area. A combination of listed phone number directories and the Kish selection method Kish [1949] was used to ensure a representative sampling process within households. Participants completed in-lab interviews focused on life history, mental health, and medical background. For this research, we utilized narrative transcripts from 1,408 participants who completed the life story component.

STD	**Low**	**Mid**	**High**	**Total**
1.0	242	891	275	1408
0.5	474	541	393	1408
SES Class	**Total Texts**	**Avg. Sen.**	**Avg. Words**	**Total Words**
Low	474	113	1669	791145
Mid	541	105	1622	877750
High	393	101	1560	612951

Table 1: SES category distributions and corresponding text statistics (STD = 0.5). Sen. = Sentences

As shown in Table 1, we examined the textual structure across SES groups, focusing on sentence count, average word count per narrative, and total word count. While sentence and word length remained relatively stable across groups, total word volume differed, which may indicate variation in narrative elaboration styles across SES levels.

3.2 Data Preparation

We retained only participant speech in each transcript and excluded interviewer utterances to minimize irrelevant content. The text data was then preprocessed by converting to lowercase, tokenizing words, and removing stop words using the Natural Language Toolkit (NLTK) Bird and Loper [2004].

For SES labeling, we adopted a composite score based on the average of parental education, participant education, and annual household income Iacovino et al. [2014]. This composite was used to categorize narratives into three SES tiers—low, mid, and high—by applying both 1.0 and 0.5 standard deviation cutoffs from the mean (see Table 1). This tri-level stratification aligns with established sociological methods for SES classification Lampos et al. [2016]. To normalize SES scores, we used the StandardScaler from scikit-learn Buitinck et al. [2013] prior to training our models.

IV METHODOLOGY

This study extends the SES classification framework introduced in our previous work, *Understanding Socioeconomic Status from Life Narrative Interviews* (NLP4DH 2025), by integrating a post-classification profiling system designed to enhance model interpretability. While the original system focused on using transformer-based models for segmenting narratives and predicting SES labels, it offered limited insight into the underlying reasons behind those predictions.

The newly introduced profiling module serves as an explanatory layer: after a narrative is classified (e.g., as ”Low SES”), the profiling system analyzes the narrative content to identify the dominant topics that contributed to that decision. For instance, if the classification outcome is ”Low SES,” the profiler may reveal that the narrative centers on themes such as financial hardship or chronic debt. This integration transforms the pipeline from a black-box predictor into an interpretable system capable of revealing the semantic basis behind each SES label.

4.1 Thematic Segmentation Using Semantic Markers

In our previous work, thematic analysis of narratives was hindered by the reliance on global topic modeling techniques that lacked fine-grained semantic alignment. To overcome this, the current profiling system introduces a marker-driven segmentation mechanism, enabling us to extract and align narrative content with predefined social and psychological themes.

We developed a taxonomy of nine core themes in collaboration with a domain expert, drawing from established psychological and socioeconomic literature. These themes include: *physical health literacy*, *mental health*, *psychological traits*, *life satisfaction*, *educational background*, *financial status*, *relationship satisfaction*, *cultural identity*, and *age and generational indicators*. Each theme was accompanied by a carefully curated list of representative keywords intended to capture relevant linguistic patterns within free-form life narratives.

To assign sentences to these themes, we employed a sentence-level semantic matching technique. Narratives were decomposed into individual sentences, and content-rich terms were extracted through preprocessing. Each term was encoded using Sentence-BERT (SBERT), and its cosine similarity was computed against pre-encoded embeddings of theme-specific keyword sets. A sentence was assigned to a theme if at least one of its terms achieved a similarity score above a threshold of 0.5. This value was chosen empirically after experimentation: higher thresholds (e.g., 0.6 or 0.7) yielded very few sentence assignments, limiting coverage, while lower thresholds (e.g., 0.3 or 0.4) led to over-assignment, where sentences were mapped to multiple or all themes. Thus, 0.5 provided an effective balance between thematic precision and representational coverage.

This semantic filtering yielded a robust and interpretable representation of thematic presence in each SES group. For example, the number of sentences associated with the *relationship satisfaction* theme reached 11,509 in the low SES group, 13,049 in the medium SES group, and 8,691 in the high SES group. Similarly, *educational background* was prominently discussed across all groups (7,706 in low, 9,658 in medium, and 6,834 in high), highlighting the theme's consistent relevance to perceived SES. These sentence-level corpora per theme formed the basis for subsequent topic modeling and profiling analysis.

All outputs were saved in JSON format per SES group, forming a structured thematic map that informs both topic extraction and interpretability stages in the profiling pipeline. Table 2 presents the number of sentences extracted per theme across SES groups. Notably, themes such as *relationship satisfaction*, *educational background*, and *mental health* appear frequently across all SES levels, while more nuanced differences emerge in themes like *financial status* and *life satisfaction*.

Theme	**Low SES**	**Medium SES**	**High SES**
Physical Health Literacy	4,163	4,383	3,118
Mental Health	9,083	10,066	7,221
Psychological Traits	1,229	1,446	984
Life Satisfaction	2,247	2,579	1,865
Educational Background	7,706	9,658	6,834
Financial Status	4,305	5,073	3,440
Relationship Satisfaction	11,509	13,049	8,691
Cultural Identity	1,880	2,380	1,796
Age and Generational Indicators	7,970	9,245	6,145

Table 2: Number of extracted sentences per theme across SES groups.

4.2 Topic Extraction and Filtering

A key limitation of the original system was the lack of clear, SES-distinctive topics—largely due to extracting topics from entire life narratives, which often contain overlapping and diverse content. This approach diluted the specificity of topic representations, leading to noisy outputs with limited discriminative power. In the updated system, we address this by introducing a multi-stage filtering pipeline that segments narratives into meaningful units and generates semantically distinct, SES-aligned topic sets.

We began by segmenting narrative sentences according to predefined themes (as described in the previous subsection), and applied BERTopic Grootendorst [2022] to the theme-specific sentence sets for each SES group. To improve topic coherence and relevance, we used a custom stop word list developed in collaboration with a domain expert. This list filtered out filler words and informal conversational markers commonly found in spoken narratives (e.g., *"um," "yeah," "gonna"*). BERTopic was configured with UMAP McInnes et al. [2018] for dimensionality reduction and HDBSCAN Campello et al. [2013] for clustering, with the number of topics constrained to 10 per theme to ensure consistent granularity.

From this step, we extracted an initial pool of 45 candidate topics per SES group. We then computed pairwise semantic similarity using Sentence-BERT (SBERT) embeddings of topic descriptors (i.e., the top words per topic). Topics with cosine similarity greater than 0.55 were treated as semantically redundant and removed, yielding a more diverse topic set.

To further evaluate quality, we computed the c_v coherence score Röder et al. [2015] for each topic based on its constituent words. Finally, we assessed topic discriminability by calculating the average SBERT-based cosine similarity between each SES group's narrative embeddings and each topic embedding. Topics were retained only if they showed strong alignment with one SES group and minimal similarity to the others. An additional cross-group filtering step ensured that SES-specific topics were semantically distinct from those assigned to other groups (threshold: cosine similarity < 0.6).

This pipeline resulted in 29 filtered topics for the Low SES group, 25 for the High SES group, and 24 for the Medium SES group. These topics, associated with specific SES identities through semantic and statistical separation, are used in subsequent profiling and visualization tasks. The complete list of SES-aligned topics and their scores is provided in the supplementary material. A detailed breakdown of these SES-specific topics for each SES group, is presented in Appendix A, with separate tables provided for Low, Medium, and High SES groups. All code and scripts supporting the profiling pipeline are publicly available at https://github.com/NahedAbdelgaber/JDMDH-2025.

V RESULTS

5.1 Topic Distribution and Visualization

To gain interpretability into how extracted topics align with SES group narratives, we visualized their distribution using two complementary techniques: heatmap-based comparison and t-SNE projection.

Heatmap of SES Topic Similarities To better understand how SES-specific profiles diverge semantically, we visualized the 25 most distinctive topics based on the highest pairwise differences in cosine similarity scores across the three SES categories. These similarity scores were calculated between the SBERT-based embedding of each topic and the aggregated narrative embedding per SES group. Figure 1 presents a heatmap where each row corresponds to a unique theme-topic pair, and each column reflects similarity to one of the SES categories (Low, Medium, High). Color intensity captures the magnitude of alignment, with deeper hues indicating higher similarity. The topics are grouped and color-coded by their associated themes to enhance interpretability.

This visualization reveals several important patterns. Topics related to psychological traits (colored in red) show the strongest alignment with the Medium SES group. For example, topics such as "law, lawyer, firm, practice, school, decide, belief..." and "chance, risk, time, school, college..." demonstrate substantially higher similarity scores for Medium and High SES narratives, suggesting that psychological complexity and self-conceptualization may be more salient in these groups. Although these terms also overlap with educational and occupational contexts, within the narratives they often occur in discussions of self-conceptualization (e.g., deciding on career paths, weighing risks, or framing one's beliefs). Thus, their placement within the psychological traits theme reflects not just domain-specific language but also deeper cognitive processes of identity and decision-making.

Educational background themes (in green) also show high similarity scores in the Medium and High SES groups. Topics such as "college, graduate, school, job..." and "music, hearing, sing, implant..." indicate a stronger emphasis on formal education and artistic or technical training among individuals with higher socioeconomic standing.

Life satisfaction topics (light blue) exhibit peak similarity for the High SES group. Phrases like "function, vet, lofty, maybe…" suggest deeper self-reflection, stability, and goal-setting language that may characterize narratives from individuals in this SES bracket.

Financial status topics (in purple) show mixed patterns. Some themes associated with financial hardship and instability are more prominent in the Low SES group, such as those involving tension, jeopardy, or medical-related expenses. Meanwhile, other financial topics linked to stability or abstract financial discussion align more closely with the High SES group.

Mental health themes (yellow) appear moderately across all SES groups, but slightly more in the Low and Medium SES groups. This pattern might reflect greater narrative content related to trauma, disorder management, or emotional hardship.

Themes of cultural identity and generational indicators (orange) are especially visible in Medium and High SES narratives. Topics such as "machine, technician, product…" or "girl, rid, spider…" reflect a broader range of generational or socio-technical reflections, possibly influenced by education, mobility, or broader cultural exposure.

Physical health literacy (cyan) shows slightly higher similarity in the Low SES group, suggesting that issues such as injury, illness, or chronic conditions are more frequently narrated or emphasized in these life stories.

Finally, relationship satisfaction topics (pink) tend to align more with the High SES group. This may indicate that individuals from this group narrate more complex or emotionally nuanced relationship dynamics, possibly due to different relationship expectations or available support systems.

Overall, this visualization confirms that the profiling system is capable of uncovering meaningful, theme-based distinctions in how different SES groups construct their narratives. The use of color-coded themes enhances interpretability, and the patterns observed here reflect a nuanced understanding of the interplay between socioeconomic identity and linguistic expression.

t-SNE Visualization of Narratives and Topics We further applied t-SNE to project both narrative and topic embeddings into two-dimensional space for each SES group. As shown in Figure 2, panel (a) presents the Low SES group, panel (b) the Medium SES group, and panel (c) the High SES group, with narrative points (dots) and topics (stars) displayed in each. Outliers were removed using Isolation Forest to reduce noise in the projections, ensuring clearer visualization of the group-level alignments.

The spatial proximity between narrative clusters and their corresponding topics visually demonstrates coherence within groups. As shown in Figure 2, panel (a) illustrates that topic stars for the Low SES group are embedded near dense narrative regions, indicating tight semantic alignment. Panel (b) displays the Medium SES group, where narrative-topic correspondence highlights reliable clustering. Panel (c) shows the High SES group, again confirming that the selected topics serve as meaningful representatives of the respective SES groups' experiences. Collectively, these visualizations validate the coherence of the extracted themes and their interpretability across SES groups.

Together, these visualizations provide a multi-faceted validation of the improved profiling system, offering interpretability, coherence, and group-level semantic differentiation that were previously lacking.

5.2 Evaluation on Unused 20% Data

To assess the generalizability of our profiling system, we evaluated the classifier on the 20% of data that was held out from training and topic similarity analysis. The classification yielded an overall accuracy of 64%. For the Low SES class, the precision was 0.43, recall was 0.23, and F1-score was 0.30. In contrast, the Not Low SES class achieved a precision of 0.68, recall of 0.84, and F1-score of 0.76.

The macro-average precision, recall, and F1-score were 0.56, 0.54, and 0.53 respectively, while the weighted averages were 0.60, 0.64, and 0.60. These results indicate a modest but meaningful improvement in detecting Low SES narratives compared to earlier iterations. In particular, the increase in recall and F1-score for the Low SES class reflects better sensitivity to underrepresented group characteristics, showcasing the profiling system's potential for equitable classification across socioeconomic groups.

5.3 Example Narrative Analysis

To illustrate the interpretability of our profiling system, we present an example narrative and its corresponding detected topics. The input narrative is as follows:

> *Growing up, I faced hardship as my father suffered from diabetes and a stroke, leaving him partially paralyzed. Lacking insurance, my mother struggled to afford his medication, relying on church visits and faith for support. The pastor often brought us both spiritual and physical aid. Amid these challenges, I remained committed to school, aspiring to become a nurse to help others like my father—driven by faith, resilience, and the will to rise above our situation.*

The narrative was semantically analyzed using SBERT embeddings, revealing the top three most similar topic clusters. The most similar topic was: "chance, risk, time, school, church, college, sit, learn, god, home" with a similarity score of 0.4692, which corresponds to the Low SES group. The second most similar topic was: "walk, minister, movement, church, read, chaplain, black, faith, ordained, lottery" with a similarity score of 0.4505, also aligned with the Low SES group. The third closest topic was: "church, catholic, bible, religious, faith, seminary, priest, spiritual, lutheran, sunday," which had a similarity score of 0.3452 and is categorized under the Medium SES group.

This narrative clearly aligns with low SES characteristics, reflected by two of the top three topics being labeled as Low SES. The first topic captures themes of adversity and striving through education and faith. The second emphasizes the religious and community support received, while the third, although categorized as Medium SES, remains semantically close to the narrative's religious aspects. This illustrates the system's ability to identify thematically and socioeconomically meaningful patterns from narrative input.

5.4 Out-of-Distribution Evaluation

Although the profiling system is not explicitly designed for classification, we assess its potential in supporting binary SES prediction (Low SES vs. Not Low SES) in OOD settings. This evaluation follows the same task setup used in our previously published work at NLP4DH 2025 Abdelgaber et al. [2025], allowing direct comparison with classical supervised models.

To evaluate the generalizability of our models, we performed an OOD test using 74 low SES student narratives from Kelbessa et al. [2024] and 74 manually curated non-low SES narratives from Reddit communities such as 'college' and 'ApplyingToCollege'.

The current profiling approach estimates SES by aggregating similarities between narrative sentences and a refined set of SES-distinctive topics (as described in Section 4.2). A voting scheme based on the most similar topics determines the predicted SES class. Applied to unseen narratives, this approach yields an accuracy of **59.5%**, which surpasses the best-performing classical model—Random Forest—from the original study, which achieved only **56.0%** accuracy on the same binary OOD classification task.

Metric	Low SES	Not Low SES	Macro Avg.	Accuracy
Precision	0.65	0.57	0.61	**0.595**
Recall	0.41	0.78	0.59	
F1-score	0.50	0.66	0.58	
Support	74	74	148	

Table 3: Performance of the profiling system on OOD binary SES classification. Despite not being optimized for classification, it outperforms Random Forest from the published work (56.0%).

This result highlights the strength of the profiling representation in generalizing across SES groups. It demonstrates that a linguistically grounded, theme-guided approach can provide competitive accuracy while offering interpretability—something lacking in conventional black-box classifiers.

Qualitative Comparison In addition to performance gains, the profiling system offers semantically rich and interpretable outputs. While Random Forest achieves SES classification via opaque feature weights, the profiling model provides concrete topic-level explanations. For example, narratives associated with Low SES often surface themes like *family relationships*, *religious coping*, and *mental health*, which align with psychological and sociological theories such as the Family Stress Model Conger et al. [1994] and Religious Coping Theory Pargament [1997].

By contrast, narratives classified as Medium SES highlight themes of *routine, fitness,* and *self-identity*, suggesting a developmental shift in psychosocial focus (e.g., Maslow's hierarchy, Resilience Theory Luthar et al. [2000]). High SES narratives are dominated by topics like *legal and financial literacy*, and *institutional success*, reflecting concepts from Cultural Capital Theory Bourdieu [1986] and Control Theory Kraus et al. [2012].

SES Group	Dominant Themes	Theoretical Alignment
Low SES	Family, Religion, Mental Health	Family Stress Model (Conger), Religious Coping (Pargament)
Medium SES	Wellness, Routine, Identity	Maslow's Hierarchy of Needs, Resilience Theory (Luthar & Cicchetti)
High SES	Law, Finance, Success Framing	Cultural Capital (Bourdieu), Self-agency Theory (Kraus)

Table 4: SES-aligned topic themes and their alignment with social science theories.

These qualitative insights expose a critical trade-off: while conventional models may yield accurate predictions, they fail to offer explanations grounded in lived experience. Profiling, by contrast, allows researchers and practitioners to understand *why* an SES label may emerge based on specific narrative cues.

VI DISCUSSION AND FUTURE WORK

This study introduces a refined SES profiling system that builds on our previously published framework, enhancing its interpretability, thematic alignment, and generalization capacity. Compared to traditional classification models, the proposed system demonstrates superior OOD performance, with an accuracy of 59.5% on binary SES prediction—outperforming the Random Forest baseline from our prior work. More importantly, the system offers rich, human-interpretable insights grounded in social and psychological theories.

Through thematic segmentation and topic filtering based on SES-specific semantic variance, we uncovered meaningful distinctions in how different socioeconomic groups narrate their life experiences. For instance, Low SES narratives often emphasize familial and spiritual struggles, Medium SES stories reflect routines and identity development, and High SES texts are dominated by professional and financial themes. These associations not only support the validity of our profiling approach but also illuminate how SES manifests in language use across diverse contexts.

6.1 Interpretability vs. Performance

One of the central contributions of this work lies in demonstrating that narrative profiling can strike a balance between predictive performance and interpretability. While large-scale classifiers may achieve comparable or higher accuracy with enough training data, they often lack transparency. In contrast, our profiling system produces topic-level justifications and visual diagnostics (e.g., heatmaps and t-SNE plots), making it a promising tool for practitioners in education, health, and social policy.

6.2 Future Work

Building on the current profiling framework, future work will focus on advancing its practical deployment and expanding its capabilities. First, we aim to develop a fully integrated SES profiling application that consolidates narrative intake, theme detection, and SES inference into a streamlined platform. This system would enable educators, counselors, and researchers to extract insights from life narratives with minimal technical overhead.

Second, we will construct a structured knowledge graph that encodes connections between SES-aligned themes, extracted topics, psychological markers, and evidence-based interventions. This graph will serve as an interpretive backbone of the system, supporting semantic reasoning and contextual understanding of user profiles.

Finally, leveraging the knowledge graph, we plan to design a recommendation layer capable of suggesting tailored support resources—such as mental health services, financial coaching, or academic tools—based on the individual's expressed themes and inferred SES. These recommendations would bridge narrative understanding with actionable support, enhancing the system's social utility.

Longer-term directions include extending the framework to multilingual settings and exploring its application to diverse narrative formats, such as clinical interviews or reflective journals, thereby broadening its impact in global and policy-making contexts.

VII CONTRIBUTION

This study offers a significant advancement over our previously published SES profiling framework by presenting a more interpretable, generalizable, and thematically grounded system for analyzing narrative texts. The updated system addresses the limitations of the original version by introducing a robust topic filtering mechanism based on SBERT similarity, coherence scoring, and SES-specific variance, which collectively enhance topic diversity and social relevance.

Importantly, the system captures psychologically and sociologically informed themes that are distinctive across SES groups. This thematic alignment is demonstrated through qualitative insights and supported by interpretability tools such as heatmaps and scatter plots, which visualize topic-narrative relationships across SES levels. In an out-of-distribution binary classification task, the profiling system achieves 59.5% accuracy—surpassing the best classical model (Random Forest) from our earlier work, which attained 56.0%. While not designed explicitly for classification, this result highlights the generalizability and strength of the profiling-based representation.

Beyond predictive performance, the system surfaces interpretable topic clusters that align with SES-relevant constructs such as mental health, financial hardship, legal literacy, and life satisfaction—drawing meaningful connections to established theories in psychology and sociolinguistics. Finally, this work lays the foundation for future developments including knowledge graph construction and personalized recommendation services, bringing the vision of narrative-driven SES support systems closer to practical realization.

VIII ETHICAL AND SOCIETAL IMPACT

Our study aims to support socioeconomic equity by uncovering underrepresented challenges and strengths in personal narratives. However, modeling SES based on text entails ethical risks. There is potential for misclassification, especially among marginalized groups, which could reinforce harmful stereotypes if deployed irresponsibly. Furthermore, despite efforts to anonymize and aggregate content, narrative data may still contain sensitive experiences, necessitating careful data handling and consent. We explicitly discourage the use of this system for high-stakes decisions such as admissions or resource allocation without human oversight. Instead, we envision its role as a tool for augmenting qualitative analysis, raising awareness about structural disparities, and guiding supportive interventions grounded in empathy and context.

IX LIMITATIONS

While the proposed SES profiling system introduces significant advancements in interpretability and generalization, several limitations remain. First, the model was trained and evaluated on a geographically localized dataset from St. Louis, which may constrain its applicability to broader or more diverse populations. Additionally, the system's reliance on a manually defined set of themes and curated keywords may introduce bias and limit the discovery of emergent or context-specific topics. The threshold used for semantic similarity in theme assignment, set empirically at 0.5, may not be optimal across all domains or datasets, and a more adaptive thresholding strategy could improve performance. Furthermore, although the system shows improved performance in OOD settings, its classification accuracy (59.5%) is still modest compared to fully supervised models evaluated in-domain. The SES classification scheme itself, based on a three-tier structure using standard deviation cutoffs, may oversimplify the complex, multidimensional nature of socioeconomic status. Another limitation is the variability in narrative lengths and verbosity across SES groups, which may inadvertently bias topic modeling and similarity scoring. Lastly, the use of pre-trained language models such as SBERT introduces the risk of inherited biases, particularly in the interpretation of sensitive socioeconomic themes. These limitations point to important directions for future enhancement of the system and broader generalizability.

ACKNOWLEDGEMENTS

This work was supported by Computer Science Dept. at Lyle school of Engineering, Southern Methodist University (SMU). We would also like to thank SMU Office of Information Technology team for their assistance in using SMU AI SuperPOD.

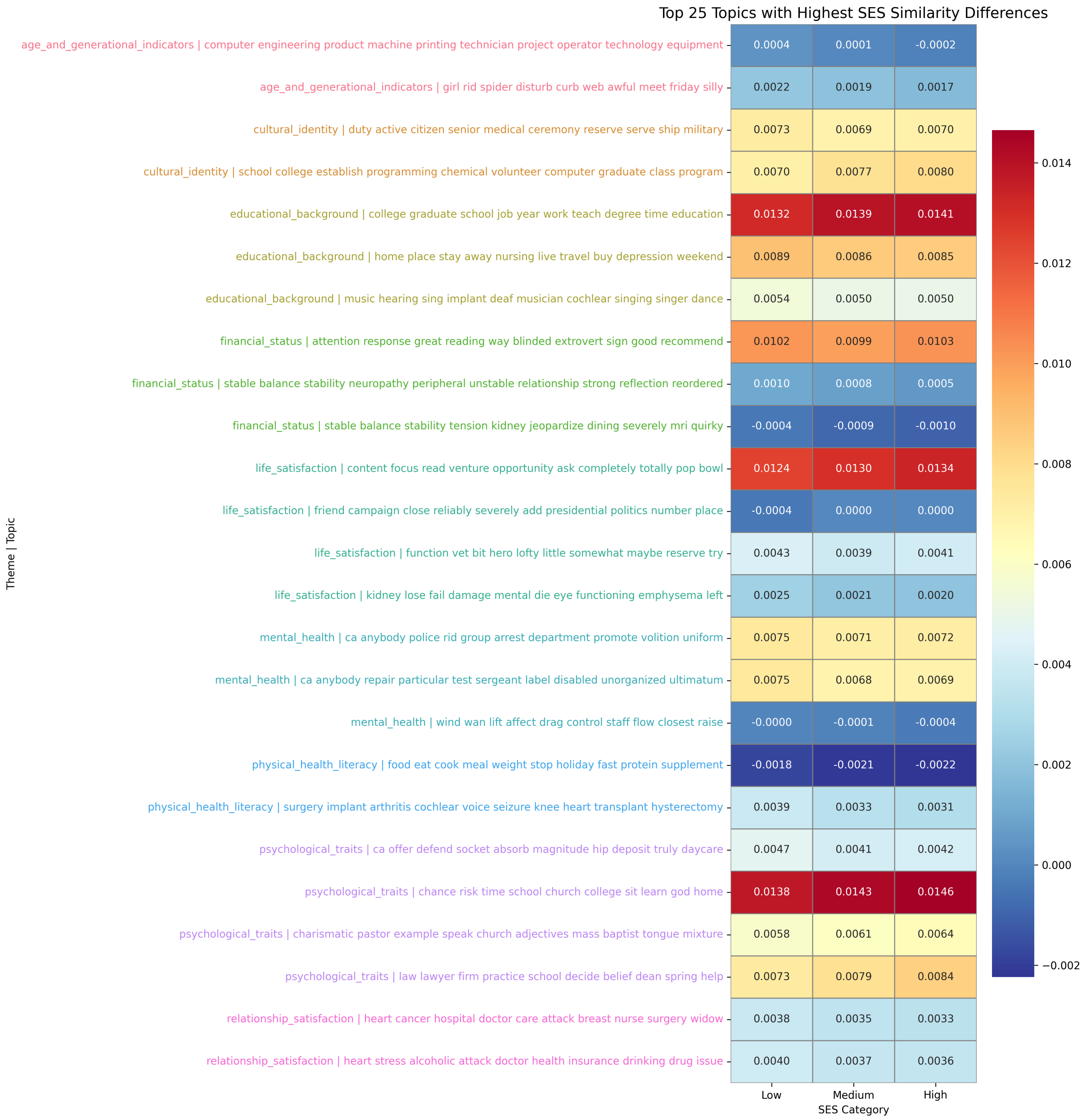

Figure 1: Top 25 topics with highest similarity differences across SES groups. Rows represent theme-topic combinations; colors reflect SBERT-based similarity between SES narrative embeddings and topic vectors.

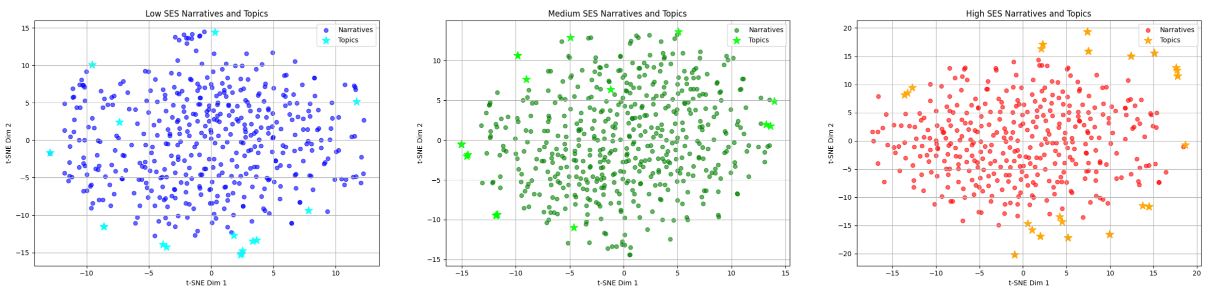

Figure 2: t-SNE visualizations of narratives and topic embeddings across SES groups: (a) Low SES (topic embeddings shown as cyan stars), (b) Medium SES (lime stars), and (c) High SES (orange stars).

A APPENDIX

A SES-ALIGNED TOPICS

The following tables present the filtered topics for each SES group. Topics are grouped according to the SES group they are most semantically aligned with, based on Sentence-BERT cosine similarity.

Topic
pressure, blood, seizure, stroke, heart, peer, diabetic, medication, attack, diabetes
artist, draw, heart, good, layout, great, art, wayward, stub, teenage
charge, deal, negotiation, dilemma, portion, album, electronics, supply, perspective, attempt
walk, lead, stand, leg, hook, turn, landmine, raise, shoulder, hand
weight, overweight, lose, belly, thread, impossible, necessarily, leg, stress, surgery
art, building, design, artist, roofing, printing, range, remodel, schematic, skid
walk, minister, movement, church, read, chaplain, black, faith, ordained, lottery
heart, failure, kidney, diabetes, congested, water, diagnose, dialysis, enlarged, multitude
insurance, car, truck, sexy, gas, ice, company, auditor, automobile, boating
cancer, surgery, breast, chemo, diagnose, liver, nerve, knee, disease, radiation
duty, active, citizen, senior, medical, ceremony, reserve, serve, ship, military
bear, clothes, grow, garden, yard, shopping, clean, healthy, food, mall
religion, catholic, christian, jewish, god, tradition, church, sin, muslim, faith
computer, engineering, product, machine, printing, technician, project, operator, technology, equipment
girl, rid, spider, disturb, curb, web, awful, meet, friday, silly
doubt, test, prejudice, inch, bad, friend, truth, sign, eighth, ambition
chance, risk, time, school, church, college, sit, learn, god, home
food, eat, hotel, stamp, grocery, body, store, meal, dog, lunch
open, door, close, plant, bar, week, job, store, place, christmas
ai, gon, game, input, computer, recognition, play, card, exploring, slap
major, type, cancer, sort, large, kind, surprised, body, disappointment, mass
wan, drive, boat, highway, downsize, tire, vehicle, bridge, vest, water
stable, stability, balance, ability, talent, skill, wherewithal, formative, exceptional, knowledge
woman, lady, weary, black, abstinent, spanish, thief, airport, organizer, heritage
tax, law, nation, return, number, arrive, taxpayer, cow, improper, freezer
mouth, drinking, drug, taste, attribute, smoking, gun, gland, remove, salivary
drug, smoke, alcoholic, drinking, alcohol, juvenile, jail, drink, court, prison
content, focus, read, venture, opportunity, ask, completely, totally, pop, bowl
country, english, italian, culture, speak, overseas, travel, language, learn, race

Table 5: Topics most aligned with Low SES group (keywords only).

References

Nahed Abdelgaber, Labiba Jahan, Nino Castellano, Joshua Oltmanns, Mehak Gupta, Jia Zhang, Akshay Pednekar, Ashish Basavaraju, Ian Velazquez, and Zerui Ma. AI assistant for socioeconomic empowerment using federated

Topic	Low Similarity	Medium Similarity	High Similarity
heart, cancer, hospital, doctor, care, attack, breast, nurse, surgery, widow	0.0038	0.0035	0.0033
stable, balance, stability, tension, kidney, jeopardize, dining, severely, mri, quirky	-0.0004	-0.0009	-0.0010
number, independent, significant, documentation, aggressive, attractive, setting, happen, contractor, capable	0.0088	0.0090	0.0091
travel, play, game, sport, enjoy, cook, dinner, walk, music, love	0.0074	0.0076	0.0076
ca, anybody, police, rid, group, arrest, department, promote, volition, uniform	0.0075	0.0071	0.0072
estate, property, house, real, buy, rent, sell, mortgage, payment, condo	0.0015	0.0014	0.0014
lawyer, attorney, young, judge, practice, sue, case, law, litigation, divorce	0.0058	0.0058	0.0059
home, place, stay, away, nursing, live, travel, buy, depression, weekend	0.0089	0.0086	0.0085
function, vet, bit, hero, lofty, little, somewhat, maybe, reserve, try	0.0043	0.0039	0.0041
account, company, business, switch, position, owner, manager, bank, checking, facility	0.0013	0.0013	0.0012
ca, offer, defend, socket, absorb, magnitude, hip, deposit, truly, daycare	0.0047	0.0041	0.0042
attention, great, instance, mind, lark, delve, map, understood, flush, acknowledge	0.0088	0.0086	0.0088
drug, hiv, alcohol, alcoholic, medication, therapist, therapy, effect, dope, spread	0.0075	0.0074	0.0075
art, paint, building, artist, performing, painting, build, color, museum, tile	-0.0009	-0.0009	-0.0011
heart, blood, leg, surgery, lead, pressure, stroke, motive, attack, weight	0.0031	0.0027	0.0027
ceremony, race, celebration, convention, anniversary, celebrate, occasion, wedding, august, invite	0.0030	0.0031	0.0030
friend, lunch, good, dinner, hang, close, family, friends, remain, friendship	0.0051	0.0053	0.0050
food, eat, cook, meal, weight, stop, holiday, fast, protein, supplement	-0.0018	-0.0021	-0.0022
movie, music, theater, guitar, bassoon, video, art, voice, play, sang	-0.0019	-0.0021	-0.0022
church, catholic, bible, religious, faith, seminary, priest, spiritual, lutheran, sunday	0.0053	0.0053	0.0055
film, shoot, stereo, camera, art, bassoon, video, performing, freelance, music	-0.0000	-0.0002	-0.0003
surgery, heart, arthritis, knee, injury, leg, transplant, brain, disease, headache	0.0021	0.0018	0.0019
ship, army, military, scud, navy, knock, heavy, gps, missile, hull	-0.0049	-0.0051	-0.0052
charismatic, pastor, example, speak, church, adjectives, mass, baptist, tongue, mixture	0.0058	0.0061	0.0064

Table 6: Topics most aligned with Medium SES group.

learning. In Mika Hämäläinen, Emily Öhman, Yuri Bizzoni, So Miyagawa, and Khalid Alnajjar, editors, *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities*, pages 490–501, Albuquerque, USA, May 2025. Association for Computational Linguistics. ISBN 979-8-89176-234-3. doi: 10.18653/v1/2025.nlp4dh-1.42. URL https://aclanthology.org/2025.nlp4dh-1.42/.

Steven Bird and Edward Loper. NLTK: The natural language toolkit. In *Proceedings of the ACL Interactive Poster and Demonstration Sessions*, pages 214–217, Barcelona, Spain, July 2004. Association for Computational Linguistics. URL https://aclanthology.org/P04-3031.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

Pierre Bourdieu. The forms of capital. In John Richardson, editor, *Handbook of Theory and Research for the Sociology of Education*, pages 241–258. Greenwood Press, New York, 1986.

Sven Buechel, Anne Buffone, Martin Seligman, and Lyle Ungar. Modeling empathy and distress in reaction to news stories. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 526–536. Association for Computational Linguistics, 2019.

Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

Ricardo JGB Campello, Davoud Moulavi, and Jörg Sander. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 160–172. Springer, 2013.

Rand D Conger, Katherine J Conger, and Glen H Elder Jr. Economic stress, coercive family process, and developmental problems of adolescents. *Child development*, 65(2):541–561, 1994.

Anjalie Field and Yulia Tsvetkov. Predicting latent attributes of twitter users with label regularization. In *Proceedings of the Second Workshop on NLP and Computational Social Science*, pages 1–6. Association for Computational Linguistics, 2017.

Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.

Alexander Hoyle, Carly McDuffie-Goodwin, and Jacob Eisenstein. Improving fairness in sociolinguistic nlp with explainable ai. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2704–2717, 2022.

Juliette M Iacovino, Joshua J Jackson, and Thomas F Oltmanns. The relative impact of socioeconomic status and childhood trauma on black-white differences in paranoid personality disorder symptoms. *Journal of abnormal psychology*, 123(1):225, 2014.

Motti Kelbessa, Estephanos Jebessa, and Labiba Jahan. Addressing educational inequalities of low ses students: Leveraging natural language processing for impact. In *Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '24, page 388–391, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400717604. doi: 10.1145/3652037.3663936.

Leslie Kish. A procedure for objective respondent selection within the household. *Journal of the American statistical Association*, 44(247):380–387, 1949.

Michael W Kraus, Paul K Piff, and Dacher Keltner. Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology*, 102(6):1376–1388, 2012.

Vasileios Lampos, Nikolaos Aletras, Jens K Geyti, Bin Zou, and Ingemar J Cox. Inferring the socioeconomic status of social media users based on behaviour and language. In *Advances in Information Retrieval: 38th European Conference on IR Research, ECIR 2016, Padua, Italy, March 20–23, 2016. Proceedings 38*, pages 689–695. Springer, 2016.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

Suniya S Luthar, Dante Cicchetti, and Bronwyn Becker. The construct of resilience: A critical evaluation and guidelines for future work. *Child development*, 71(3):543–562, 2000.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Thomas F Oltmanns, Merlyn M Rodrigues, Yana Weinstein, and Marci EJ Gleason. Prevalence of personality disorders at midlife in a community sample: Disorders and symptoms reflected in interview, self, and informant reports. *Journal of psychopathology and behavioral assessment*, 36:177–188, 2014.

Kenneth I Pargament. *The Psychology of Religion and Coping: Theory, Research, Practice*. Guilford Press, New York, 1997.

James W. Pennebaker and Cindy K. Chung. *Expressive writing: Words that heal*. iUniverse, 2013.

Daniel Preoţiuc-Pietro, Vasileios Lampos, Nikolaos Aletras, Eleni Aletrari, George Gkotsis, Arkaitz Zubiaga, et al. Studying the demographics of social media users: Modeling age, gender and income with social media. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 221–230. ACM, 2015.

Michael Röder, Andreas Both, and Alexander Hinneburg. Exploring the semantic coherence of topics. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 399–408, 2015.

Svitlana Volkova, Kyle Shaffer, Jin Yea Jang, and Nathan Hodas. Inferring user political preferences and demographics from youtube videos. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 74–80. Association for Computational Linguistics, 2016.

Xuhong Zhang, Jianbo Shen, Weiran Wang, Pengcheng Wang, and Zhou Yu. Interpretable multi-task learning for multi-modal emotion recognition and sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2752–2765, 2020.

Zixuan Zhang, Shuhan Liu, Yujia Mo, Yitong Liu, Zijian Liu, Yao Li, and Di Jin. Multi-task learning for social attribute prediction using pre-trained language models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 2468–2483, 2022.

Topic
future, television, communication, journalism, possible, skill, novel, action, commercial, speak
law, lawyer, firm, practice, school, decide, belief, dean, spring, help
heart, stress, alcoholic, attack, doctor, health, insurance, drinking, drug, issue
kidney, lose, fail, damage, mental, die, eye, functioning, emphysema, left
college, graduate, school, job, year, work, teach, degree, time, education
friend, campaign, close, reliably, severely, add, presidential, politics, number, place
music, hearing, sing, implant, deaf, musician, cochlear, singing, singer, dance
african, identity, culture, american, irish, politically, black, italian, hispanic, mexican
church, ministry, pastor, seminary, pastoring, congregation, community, preach, halftime, thankful
lawyer, attorney, case, dentist, judge, circuit, testify, firm, office, law
church, abuse, pain, heroin, treatment, sexual, drug, opportunity, nearby, youth
school, college, establish, programming, chemical, volunteer, computer, graduate, class, program
stable, balance, stability, neuropathy, peripheral, unstable, relationship, strong, reflection, reordered
bear, buck, poor, kid, highlight, bet, daughter, brunt, princess, reconcile
wind, wan, lift, affect, drag, control, staff, flow, closest, raise
type, major, sort, change, attack, merge, sobriety, heart, pastor, turn
attention, response, great, reading, way, blinded, extrovert, sign, good, recommend
food, eat, eating, cook, fat, meal, table, disorder, inspection, pantry
ca, anybody, repair, particular, test, sergeant, label, disabled, unorganized, ultimatum
funny, guy, restaurant, laugh, waitress, dinner, sneak, chairlift, snowdrift, blast
music, play, sing, shape, baseball, exercise, ball, voice, athlete, meditation
chance, risk, da, lord, grateful, pregnancy, live, house, turn, eager
tax, divorce, gift, estate, alimony, substance, reality, friend, support, federal
doubt, mind, real, official, hardest, fresh, worse, personnel, catholic, folk
surgery, implant, arthritis, cochlear, voice, seizure, knee, heart, transplant, hysterectomy

Table 7: Topics most aligned with High SES group (keywords only).