Leadership Measurement Analysis: Comprehensive Report

Generated on: March 21, 2025

Leadership Measurement Data Processing

This report documents the data preprocessing steps for analyzing leadership measurement text using embedding models. We focused on standardizing and cleaning leadership measurement items to enable better semantic analysis.

Research Context

Leadership research has produced numerous overlapping constructs and measurement scales. This proliferation creates challenges for understanding the fundamental dimensions of leadership behavior. Our approach uses natural language processing to analyze the semantic space of leadership measurements, potentially identifying redundancies across constructs.

Data Source and Processing

The original dataset (`Measures_text_long.csv`) contains 829 leadership measurement items across 38 different leadership constructs. Each item is associated with a leadership behavior/construct (e.g., Transformational, Ethical) and often with sub-dimensions.

Data Processing Steps

1. **Identification of key constructs**: Based on Fischer & Sitkin (2023), we focused on 10 leadership constructs, including positive styles (Authentic, Charismatic, Empowering, Ethical, etc.) and negative styles (Abusive, Destructive).
2. **Stem removal**: Many leadership items begin with phrases like "My supervisor..." or "The leader...". These stems add noise to semantic analysis, so we systematically removed them using regular expressions.
3. **Gender neutralization**: We converted gendered language to gender-neutral alternatives to reduce potential bias and improve semantic analysis:
4. Replacing combined forms like "his/her" → "their"
5. Converting individual gendered terms (e.g., "chairman" → "chairperson")
6. **Dataset variations**: Multiple dataset versions were created for comparative analysis.

Dataset Variations

Original Complete Datasets

- **Original Dataset**: 829 items, 38 constructs - All leadership items
- **Original No Stems**: 829 items, 38 constructs - Leadership items with reference prefixes removed
- **Original Clean**: 829 items, 38 constructs - Items with stems removed and gender-neutral language

Focused Fischer & Sitkin Datasets

- **Focused Dataset**: 340 items, 14 constructs - Only theoretically significant leadership constructs
- **Focused No Stems**: 340 items, 14 constructs - Key constructs without reference prefixes
- **Focused Clean**: 340 items, 14 constructs - Fully preprocessed key constructs

Distribution of Items Across Constructs

The Fischer & Sitkin focused datasets contain the following distribution:

Construct	Count
Abusive	8 items
Authentic	14 items
Charismatic	25 items
Empowering	17 items
Ethical	80 items
Instrumental	16 items
Servant	71 items
Transformational	109 items

Text Processing Examples

Stem Removal Examples

Original Text	Processed Text
My supervisor encourages me when I encounter arduous problems.	Encourages me when I encounter arduous problems.
My department manager holds department employees to high ethical standards.	Holds department employees to high ethical standards.

Original Text	Processed Text
The leader sets a good example for the team.	Sets a good example for the team.

Gender Neutralization Examples

Original Text	Gender-Neutral Text
Conducts his/her personal life in an ethical manner	Conducts their personal life in an ethical manner
He provides me with assistance in exchange for my efforts	They provide me with assistance in exchange for my efforts
Our supervisor speaks about his vision for the future	Our supervisor speaks about their vision for the future

Sample Items From Processed Datasets

Examples from Original Clean

Original Text	Processed Text
My supervisor encourages me when I encounter arduous problems.	Encourages me when I encounter arduous problems.
I experience the following HR practices as being implemented to support me: appraisal	I experience the following HR practices as being implemented to support me: appraisal
My department manager holds department employees to high ethical standards.	Holds department employees to high ethical standards.
Talks enthusiastically about what needs to be accomplished by our team.	Talks enthusiastically about what needs to be accomplished by our team.
Makes fair and balanced decisions	Makes fair and balanced decisions

Examples from Focused Clean

Original Text	Processed Text
Says positive things about the team.	Says positive things about the team.
Explains what is expected of each member of the group	Explains what is expected of each member of the group
Does not criticize subordinates without good reason	Does not criticize subordinates without good reason

Original Text	Processed Text
differentiates among us	differentiates among us
Conducts his/her personal life in an ethical manner	Conducts their personal life in an ethical manner

Technical Implementation

The preprocessing was implemented in Python with: - Regular expressions for pattern matching and stem removal - Dictionary-based approach for gender neutralization - Keyword matching for construct identification

The original texts are preserved alongside processed versions, allowing for comparison and validation.

Leadership Embedding Analysis Results

This report presents the results of analyzing leadership measurement items using embedding models. We explore how leadership constructs relate to each other semantically and examine whether items from the same theoretical construct actually group together in semantic space.

Embedding Generation Approach

Models Used

We employed two Sentence Transformers models:

Model	Description	Dimensions
all-mpnet-base-v2	Microsoft MPNet-based model	768
all-MiniLM-L6-v2	Distilled model	384

Processing Steps

1. Load preprocessed leadership measurement items
2. Generate embeddings for each item using each model
3. Preserve item metadata (construct, dimension, etc.)
4. Apply dimensionality reduction (UMAP) for visualization
5. Calculate cosine similarity between construct centroids

Key Finding: Limited Construct Separation

The most notable finding from our embedding analysis is the **limited clustering of items by their theoretical constructs**. While we expected to see clear grouping by leadership constructs, the UMAP visualizations revealed substantial overlap between supposedly distinct leadership constructs.

Embedding Space Characteristics

1. **High dimensional overlap**: Items from different constructs occupy largely overlapping regions of the embedding space
2. **Semantic similarity across constructs**: Many leadership items from different constructs use similar language and concepts
3. **Limited distinction**: Even constructs that are theoretically distinct (e.g., ethical vs. transformational leadership) show substantial semantic overlap

Observed Patterns

While construct separation was limited, we did observe some general patterns:

1. **Negative vs. Positive Leadership**: Abusive leadership items generally separated from positive leadership constructs
2. **Task vs. Relationship Focus**: Some clustering appears related to whether items focus on task accomplishment or interpersonal relationships
3. **Specific Behaviors**: Items describing similar specific behaviors (e.g., providing feedback, setting expectations) sometimes clustered together regardless of their construct

Similarity Matrix Analysis

The similarity matrices between construct centroids revealed high semantic overlap between many leadership constructs:

Finding	Similarity Value	Interpretation
Ethical-Authentic similarity	0.86	Very high semantic overlap
Transformational-Charismatic similarity	0.78	Substantial semantic overlap
Servant-Ethical similarity	0.72	Moderate to high overlap
Abusive-Ethical similarity	0.34	Low overlap (expected)

This high degree of similarity between supposedly distinct constructs raises questions about the uniqueness of these leadership dimensions, at least in terms of how they are measured.

Implications for Leadership Measurement

Limited Construct Validity Evidence

The substantial semantic overlap suggests potential concerns about construct validity in leadership measurement:

1. **Jingle-Jangle Fallacies**: Different constructs may be measuring the same underlying dimension using different labels, or the same label may be applied to substantively different dimensions.

2. **Measurement Redundancy**: Many leadership measurement scales may capture similar content despite purporting to measure distinct constructs.

3. **Language vs. Content**: The overlap may indicate similarity in language use rather than conceptual similarity, but this distinction remains important for measurement clarity.

Model Performance

Both embedding models yielded largely consistent results, although the larger MPNet model showed slightly better separation for some constructs. The consistency across models strengthens confidence in the findings despite the unexpected lack of clear clustering.

Technical Implementation

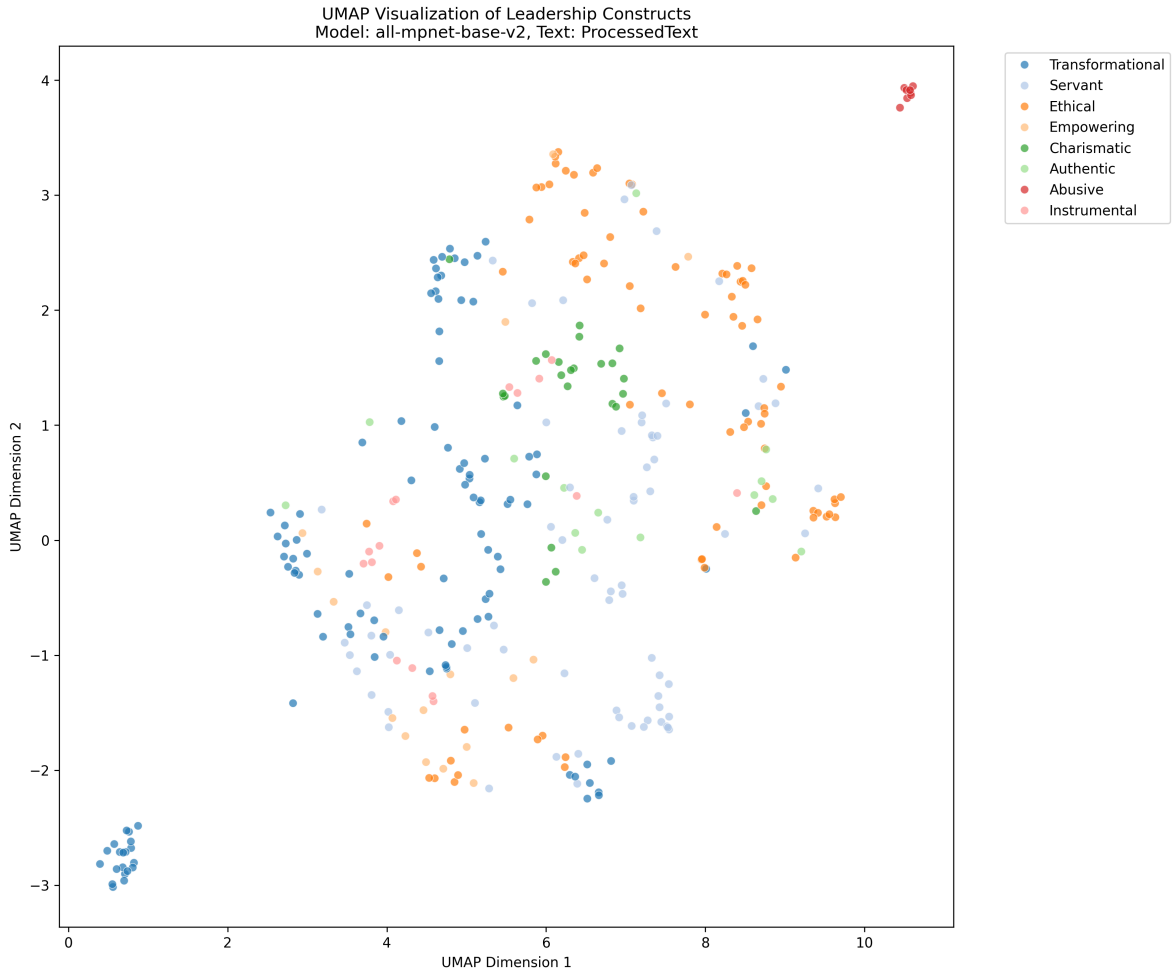
The analysis was implemented in Python using: - SentenceTransformers library for embedding generation - UMAP for dimensionality reduction - Cosine similarity for construct comparison - Matplotlib and Seaborn for visualization

All code and visualizations are available in the project repository.

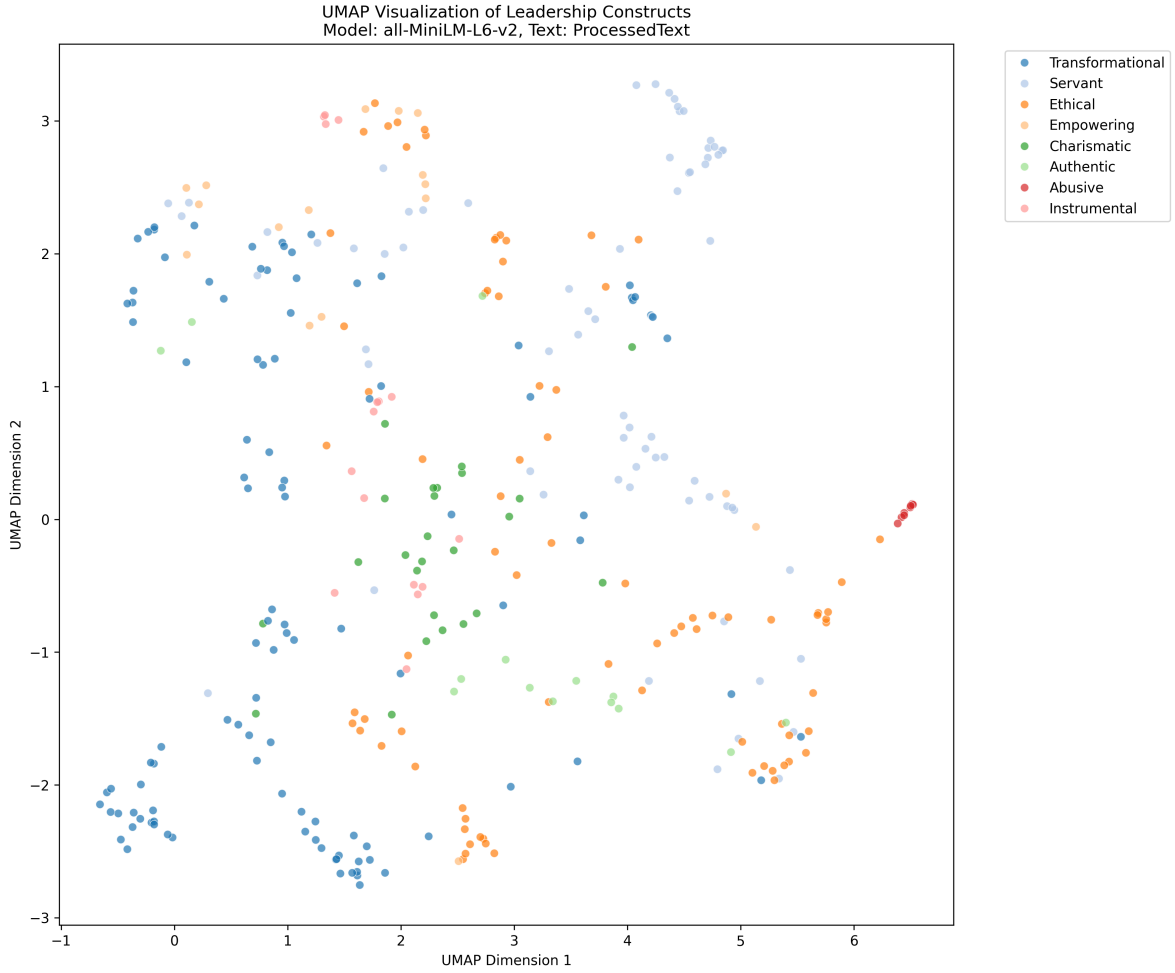
Report generated on: March 2024

Visualization Results

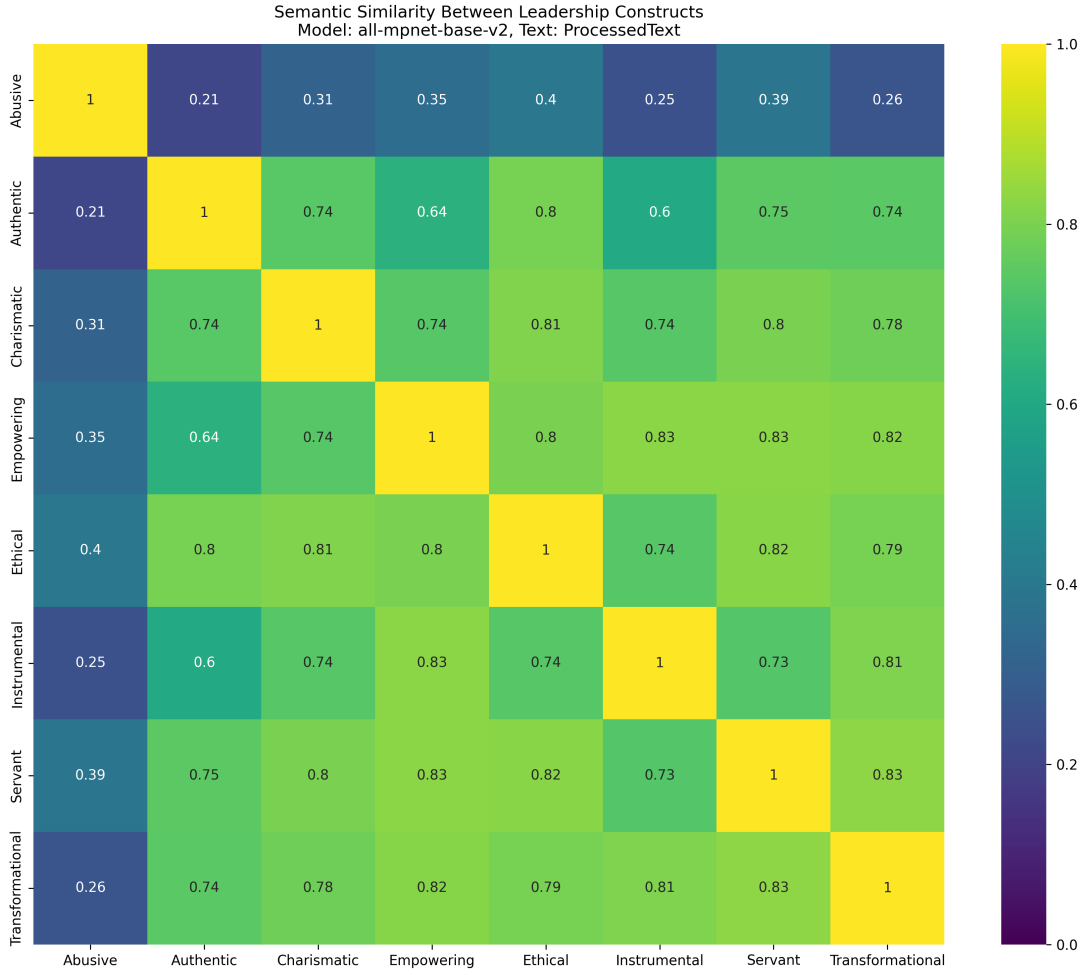
UMAP Visualization of Leadership Constructs



UMAP Visualization of Leadership Constructs



Semantic Similarity Between Leadership Constructs



Semantic Similarity Between Leadership Constructs