THE JOURNAL OF MATHEMATICAL SOCIOLOGY
2025, VOL. 49, NO. 1, 1–24
https://doi.org/10.1080/0022250X.2024.2395806

A workflow for analyzing cultural schemas in texts[1]

Marshall A. Taylor[a] and Dustin S. Stoltz[b]

[a]Department of Sociology, New Mexico State University, Las Cruces, New Mexico, USA; [b]Department of Sociology and Anthropology, Lehigh University, Bethlehem, Pennsylvania, USA

ABSTRACT
Concept class analysis (CoCA) is a method for recovering cultural schemas in texts using a combination of word embedding and community detection models. Like survey-based forms of schematic class analysis (SCA), however, interpreting results can be difficult. Some of these interpretive difficulties are applicable across types of SCA, while others are unique to CoCA. In this paper, we propose a complete workflow for interpreting and analyzing CoCA output. We use the case of social identity schemas in a collection of over 13,000 U.S. political blog posts to outline a number of interpretive and analytical strategies and a robustness check to make sense of the cultural schemas recovered from texts.

ARTICLE HISTORY
Received 8 March 2023
Revised 29 January 2024
Accepted 13 May 2024

KEYWORDS
Computational text analysis; concept class analysis; cultural schemas; cultural sociology; word embeddings

1. Introduction

Those who produce text rely on schemas – implicit memory structures learned via schematizing patterns in social interaction and experience (Boutyline & Soter, 2021; Hunzaker & Valentino, 2019; Leschziner & Brett, 2021; Wood et al., 2018). In turn, these patterned associations can be recovered from texts by attending to the latent regularities in word co-occurrences (Arseniev-Koehler & Foster, 2022; Morehouse et al., 2023). Concept class analysis (CoCA) (Taylor & Stoltz, 2020a) is such a method for measuring schemas in text. The method builds on survey-based schematic class analysis (SCA) – namely, relational class analysis (RCA) and correlational class analysis (CCA) – to find those latent groupings of documents (what we refer to as "conceptual classes") that have similar absolute values of weightings across a series of thematically-related semantic directions (e.g., feminine – masculine, liberal – conservative).

Documents within a conceptual class thus share "traces" of the same underlying schema since they share similar patterns of valuative opposition. CoCA

CONTACT Marshall A. Taylor mtaylor2@nmsu.edu Department of Sociology, New Mexico State University, 292B Science Hall, Las Cruces, NM 88003, USA

[1]A previous version of this paper was presented at the 118th Annual Meeting of the American Sociological Association in Philadelphia, PA, USA. We thank the session discussant, Neha Gondal, for insightful feedback on the paper. We also thank the editor and anonymous reviews at *The Journal of Mathematical Sociology for their feedback and commen* A replication repository for this paper can be found at: https://github.com/Marshall-Soc/CoCA_interpret

thus falls into the broad category of machine learning for text classification – with traditional forms of hierarchical and *k*-means clustering being the forebearers. CoCA differs, however, in that the algorithm partitions texts by how they assign more-or-less equal weight to the same semantic domains, *even if the valences of those weights run in opposing directions*.

Since these conceptual classes are those documents that reflect traces of the same schematic patterns and which might be said to be "effects" of the same underlying cultural schema in the text production process, sociologists of culture and cognition can use conceptual class assignments in a number of ways. For example, since the particular schematic associations one accrues over time are a function of lived experience and people with similar lived experiences are more likely to occupy similar positions in social space, a researcher may treat conceptual classes as outcome variables and test their relationships with theoretically-informed sociodemographic characteristics. Similarly, analysts may use conceptual class assignments in downstream tasks – e.g., as a predictor of other, non-textual document-producer attributes. Or, CoCA might be used to generate dyadic measures of schematic similarity to analyze the structure of text networks.

Interpreting CoCA output, though, can be difficult. Many of these interpretive difficulties are applicable across types of SCA – e.g., the need to empirically assess the extent to which inversion as a schematic pattern is happening within a class. Other hurdles are unique to CoCA, such as the possibility that a deep reading of a document might contradict that document's class assignment. Furthermore, like other forms of SCA, robustness checks are necessary to assess how well the class partitions fit the data.

In this paper, we propose a complete workflow for interpreting and analyzing CoCA output. We use the case of social identity schemas in a collection of 13,246 U.S. political blog posts to outline a number of interpretive and analytical strategies and a robustness check to make sense of the schemas recovered from texts.

2. Concept class analysis

This section offers a mathematical treatment of CoCA. Readers interested in a more in-depth theoretical treatment of how CoCA classes can be interpreted as schematic classes with underlying cultural schemas are referred to the original CoCA paper (Taylor & Stoltz, 2020a).

2.1. Semantic directions and concept mover's distance

CoCA begins with two matrices. The first matrix is a d-by-t document-term matrix, $\mathbf{M}$, where any given M_{dt} cell indicates the frequency of term t in document d. The second matrix is a w-by-v word embedding matrix, $\mathbf{E}$, where $\mathbf{w_t}$ is the tth word vector in vocabulary $\mathbf{w}$, and where $\{w_1, w_2, w_3, \cdots, w_n\} \in \mathbf{w}$. The length of $\mathbf{v}$–where $\{v_1, v_2, v_3, \cdots, v_k\} \in \mathbf{v}$–is equal to the dimensions of the word embeddings, which is typically between 100 and 300. The cosine between any two word embedding vectors in $\mathbf{E}$ – $\mathbf{w_{t(a)}}$ and $\mathbf{w_{t(b)}}$, respectively – indicates the extent to which words $w_{t(a)}$ and $w_{t(b)}$ occupy similar semantic positions in the embedding space. The cosine between the two word vectors is denoted with $\cos(\mathbf{w_{t(a)}}, \mathbf{w_{t(b)}})$.

We next use $\mathbf{E}$ to estimate a series of concept mover's distance (CMD) scores measuring the extent any given document engages with a polar end of a semantic direction (Arseniev-Koehler & Foster, 2022; Stoltz et al., 2024).[2] We create a semantic direction, $\mathbf{s}$, using the vector offset method. This involves subtracting a word vector in $\mathbf{E}$ from another word vector in $\mathbf{E}$ (Mikolov et al., 2013) – both of which are selected by the researcher on the basis of theory, intuition, etc. This creates a *juxtaposition pair* anchoring down each side of the direction. The cosine between other words in $\mathbf{E}$ and this semantic direction can then be interpreted as the extent to which a word is more contextually similar to one pole of that direction relative to the other pole.

Typically, the semantic direction is enriched by using multiple juxtaposition pairs, where one word in the pair is a synonym or descriptor for one pole, and the other word is a synonym or descriptor for the other pole. For example, for a gender semantic direction, one juxtaposition pair might be $(\overrightarrow{man}, \overrightarrow{woman})$, another might be $\left(\overrightarrow{male}, \overrightarrow{female}\right)$, and another might be $(\overrightarrow{men}, \overrightarrow{women})$ (Kozlowski et al., 2019; Taylor & Stoltz, 2020b). The anchor terms can be reused if they fit in more than one juxtaposition pair. We select these terms using a combination of prior sociological research, close reading, and intuition.[3]

[2]One could also use a semantic centroid; however, a semantic direction is generally easier to interpret. While measuring a focal word's position along a semantic direction would, for example, show whether it is more masculine *as opposed to feminine*, a word's position in relation to a semantic centroid would show whether it is, for example, more masculine *as opposed to not-masculine*. While feminine may be contained within the set "not-masculine," the latter is far more encompassing. One could also define a "gender" centroid, for instance, which would measure the extent a text discusses gender when used with concept mover's distance.

[3]The question of which words to include in the juxtaposition pairs is a question that extends beyond CoCA and to the construction of semantic directions (and other semantic relations) in word embedding spaces more broadly (Stoltz et al., 2024). The are actually two questions here: First, how do we know that we have selected anchor terms that properly index our concept of interest, and second, how well are the anchor terms we selected measuring whatever direction it is actually measuring (Boutyline & Johnston, 2023, p. 3)? The first question is about anchor set validity; the second is about anchor set reliability. As Boutyline and Johnston note, the answer to the validity question is probably always to be answered on the grounds of theoretical and intuitive knowledge (2023, p. 3). However, in that same paper, Boutyline and Johnston propose three new metrics for assessing anchor set reliability – each derived from word analogy solution tasks. One specific metric, the mean *PairDir*, produces direction-specific reliability scores that are reasonably predictive of the direction-specific accuracy scores (the latter of which are derived by taking the Pearson correlation between the cosine similarity of a word and a semantic direction, on one hand, and the mean human-rated semantic differential score of that word, on the other). The mean *PairDir* metric, then, is a prime candidate for inclusion in word embedding packages/libraries as a way to help researchers determine their anchor sets. The mean *PairDir* algorithm is also easy enough to compute: it is simply the mean cosine similarity between vector offsets across all pairwise juxtaposition pair combinations.

The vector offsets are then summarized into a single vector in some way – e.g., the paired method, which we focus on here (Kozlowski et al., 2019; Taylor & Stoltz, 2020b). Using the paired method to find a semantic direction, $\mathbf{s}$, is of the following form:

$$\mathbf{s} = \frac{\sum_p^p (\mathbf{p_1} - \mathbf{p_2})}{|P|}, \tag{1}$$

where p is a juxtaposition pair in the total set P of juxtaposition pairs, $\mathbf{p_1}$ and $\mathbf{p_2}$ are the word vectors associated with the words in p, and $|P|$ is the cardinality of P. The cosine between any word in $\mathbf{E}$ and $\mathbf{s}$, $\cos(\mathbf{w_i}, \mathbf{s})$, varies in the $[-1, 1]$ interval. The cosine will be closer to 1 when $\mathbf{w_i}$ is more contextually similar to the pole of $\mathbf{s}$ defined by the $\mathbf{p_1}$ vectors and closer to -1 when $\mathbf{w_i}$ is more contextually similar to the pole of $\mathbf{s}$ defined by the $\mathbf{p_2}$ vectors. For example, if the $\mathbf{p_1}$ vectors represent words related to "man" and the $\mathbf{p_2}$ vectors represent words related to "woman," then $\cos(\mathbf{w_i}, \mathbf{s}) \approx 1$ means $\mathbf{w_i}$ is used in more discursively masculine contexts and $\cos(\mathbf{w_i}, \mathbf{s}) \approx -1$ means $\mathbf{w_i}$ is used in more discursively feminine contexts. Each word in the juxtaposition pairs must be words in $\mathbf{E}$.

The typical method to validate a semantic direction is to categorize a sample of words manually as likely to fall on either pole. For instance, one could categorize words like "executive" and "career" as more masculine, while words like "home" and "children" are more stereotypically feminine (Durrheim et al., 2022; Jones et al., 2020). Our semantic direction should correctly quantify this bias. If it does not, we should select more and/or different anchor terms for our juxtaposition pairs.

The next step is to estimate each document's (in $\mathbf{M}$) engagement with $\mathbf{s}$. This is done with CMD scores – an extension of word mover's distance scores (Kusner et al., 2015), which themselves are extensions of earth mover's distance scores (Rubner et al., 1998). CMD relies on the word mover's distance (WMD) algorithm – a standard metric for quantifying document similarities using word vector information. WMD finds the minimum "cost" of moving one document (d_i, represented in $\mathbf{M}$ with row vector $\mathbf{d_i}$) to another (d_j, represented in $\mathbf{M}$ with row vector $\mathbf{d_j}$). This cost is a function of two quantities, each specific to the relationship between a term in the first document, $d_{i_{t(a)}}$, and a term in the second document, $d_{j_{t(b)}}$: the cosine *distance*[4] (the complement to cosine similarity), $\cos_{dist}(\cdot)$, between any two words across the two documents and a weight indicating "how much" of a word $d_{i_{t(a)}}$ will travel to $d_{j_{t(b)}}$. The goal of the algorithm is to find a "flow matrix," $\mathbf{F}$, for each (d_i, d_j) pair of documents in $\mathbf{M}$, that satisfies the following (Kusner et al., 2015, p. 3)[5]:

[4]The original CoCA paper (Taylor & Stoltz, 2020a) has a typographical error on page 551, where it should specify cosine *distance* and not cosine *similarity* in the mover's distance calculation.

[5]Readers interested in the finer details of what the flow matrix looks like and how it is computed are encouraged to read the original WMD paper (Kusner et al., 2015) and the original concept mover's distance paper (Stoltz & Taylor, 2019).:

$$\Omega_{\mathrm{d_i,d_j}} = \min\left(\sum_{\mathrm{t(a),t(b)=1}} \mathrm{F_{t(a),t(b)}} \cos_{\mathrm{dist}}\left(\mathrm{w_{t(a)}}, \mathrm{w_{t(b)}}\right)\right), \tag{2}$$

where $\mathbf{\Omega}$ is the d-by-d WMD matrix and $\Omega_{\mathrm{d_i,d_j}}$ is the cost of transforming $\mathbf{d_i}$ into $\mathbf{d_j}$. In the regular WMD case, $F_{t(a),t(b)}$ is found under the constraints that the summation of row entry $\mathbf{F_{t(a),}}$ is equal to the relative frequency of $t(a)$ in $\mathbf{d_i}$ and the summation of column entry $\mathbf{F_{,t(b)}}$ is equal to the relative frequency of $t(b)$ in $\mathbf{d_j}$ (Kusner et al., 2015, p. 3). CMD, however, incorporates relaxed word mover's distance (RWMD), which finds Ω_{d_i,d_j} with the first constraint removed and again with the second constraint removed, and then takes the maximum of these two minimum costs as the entry in both Ω_{d_i,d_j} and Ω_{d_j,d_i}, thus making $\mathbf{\Omega}$ symmetric. Smaller values of Ω_{d_i,d_j} indicate that documents d_i and d_j are more contextually similar.[6]

CMD diverges from WMD in one important respect: document vector $\mathbf{d_j}$ is replaced with a "pseudo-document" vector, Θ. This pseudo-document contains only terms representing the concept of interest (Stoltz & Taylor, 2019) or a "pseudo-term" corresponding to a semantic direction (Taylor & Stoltz, 2020b). This involves adding a new row to $\mathbf{M}$ that consists of all 0s – except for the columns for the term(s) denoting the concept, which get(s) a 1. In the case of semantic directions – the focus here – this also involves adding a new column denoting that direction ($\mathbf{s}$) and then adding a 1 for that column in the pseudo-document row (Taylor & Stoltz, 2020b):

$$\mathrm{M}_{\Theta t} = \begin{cases} 1 & \text{if} \quad \mathrm{t = s} \\ 0 & \text{if} \quad \text{otherwise} \end{cases}. \tag{3}$$

The CMD score for d_i is then found with:

$$c'_{d_i} = c'_{d_i,\Theta} = \min\left(\sum_{t(a),s=1} F_{t(a),s} \cos_{\mathrm{dist}}\left(\mathbf{w_{t(a)}}, \mathbf{s}\right)\right). \tag{4}$$

The vector of document-specific distances, $\mathbf{c}'$, is converted into a vector of z-scores and then inverted so that larger scores indicate concept *engagement* – i.e., $c_{d_i} = z\left(c'_{d_i}\right) \times -1$. c_{d_i} is the semantic direction CMD score for document d_i, where larger positive scores indicate higher levels of concept engagement with the positive pole (e.g., "man"), higher negative scores indicate higher levels of engagement with the negative pole (e.g., "woman"), and a score of 0 indicates a document with a mean engagement level. Finally, we take a sample

[6]The WMD equation bears similarities to Boutyline, Cornell, and Arseniev-Koehler's (Boutyline et al., 2021, pp. 1427–1428) transmission chain simulation, which also uses word frequencies and cosine information in the workflow.

of documents from each extreme of the CMD engagement vector for close reading to validate the scores.[7]

2.2. Absolute correlations and modularity maximization

CoCA assumes that the document-by-CMD matrix – which we'll denote as a d-by-c matrix $\mathbf{C}$, following notation in the previous section – has a ≥ 2 column length. In other words, there should be at least two semantic direction CMD column vectors. This matrix $\mathbf{C}$ is then passed to the CCA algorithm (Boutyline, 2017) to find the conceptual classes and the schemas that underlie them.

Specifically, the d-by-c matrix $\mathbf{C}$ is converted into a symmetric d-by-d matrix ρ, where each ρ_{d_i,d_j} cell contains the absolute Pearson correlation between those two documents:

$$\rho_{d_i,d_j} = \left| \frac{\sum_{c=1}^{C} \left(d_i - \bar{d}_i\right)\left(d_j - \bar{d}_j\right)}{\sqrt{\sum_{c=1}^{C} \left(d_i - \bar{d}_i\right)^2 \sum_{c=1}^{C} \left(d_j - \bar{d}_j\right)^2}} \right|. \tag{5}$$

See Taylor and Stoltz (2020a) and Boutyline (2017) for discussions of how absolute Pearson correlation coefficients are good indicators of schematic similarity.

Finally, after optionally removing statistically non-significant absolute correlations,[8] partitioning ρ into conceptual classes is treated as a network community detection problem. The ρ matrix is conceptualized as a weighted undirected graph, and the partition that maximizes the graph's modularity score Q is returned (Boutyline, 2017; Goldberg, 2011; Taylor & Stoltz, 2020a). The Q value for a partition is found with (Newman, 2006, p. 5):

$$Q = \frac{1}{2m} \sum_{d_i,d_j} \left[\rho_{d_i,d_j} - P_{d_i,d_j}\right] \delta\left(g_{d_i}, g_{d_j}\right), \tag{6}$$

where ρ_{d_i,d_j} is the observed bivariate absolute correlation between documents d_i and d_j, P_{d_i,d_j} is the probability of some absolute correlation

[7]We could also manually label a sample of documents prior to estimating CMD, and calculate agreement metrics (e.g., Carbone & Mijs, 2022, p. 7)

[8]The norm across all three current forms of schematic class analysis – RCA, CCA, and CoCA – is to remove statistically non-significant edges from the respondent-by-respondent/document-by-document graph. This is because it quite unlikely to observe a perfectly zero relationality (in the case of RCA) or correlation (for CCA and CoCA) between any two respondents/documents. The graph, then, is very dense, and many of the edges hovering around zero might be noise (Goldberg, 2011, p. 1408). However, no work to current knowledge has assessed how sensitive the resulting schematic classes might be to α thresholds before partitioning. As a reviewer for this paper astutely points out, the probability that a given edge might fall below a given α threshold might be correlated with the number of column vectors used to compute the rowwise relationalities/correlations. Future work might attend to this question.

between d_i and d_j if the correlations in the graph were placed at random (while constrained to have the same weighted degree centralities of ρ), m is the summation of all absolute correlations (weighted edges) in graph ρ (i.e., $m = \frac{1}{2}\sum_{d_i,d_j} \rho_{d_i,d_j}$ (Newman, 2004, p. 7)), and $\delta(g_i, g_j) = 1$ if d_i and d_j are in the same class (g) and 0 if they are not. The P_{d_i,d_j} term is defined as:

$$P_{d_i,d_j} = \frac{\lambda_{d_i}, \lambda_{d_j}}{2m}, \tag{7}$$

where λ_{d_i} and λ_{d_j} are the weighted degrees for the d_i and d_j documents, respectively. In the case of CoCA/CCA, a higher Q indicates stronger absolute correlations between documents assigned to the same class than one would expect if the correlations were placed between documents at random. The higher the Q (with a maximum of 1), the more pronounced the community structure of the graph. All forms of SCA (RCA, CCA, and CoCA) use the leading eigenvector approximation method (see Newman, 2006).

CoCA then returns the partition $\mathbf{g}$ of length k–$\{g_1, g_2, \cdots, g_k\} \in \mathbf{g}$–that maximizes Q, and where every document d is allocated to one and only one g class. SCA techniques increase the covariance observed between column vectors within the classes (Goldberg, 2011, p. 1415); as such, each g can be interpreted and subsequently labeled with reference to its own inter-direction correlation matrix.

3. A workflow for CoCA

Our workflow begins, naturally, with selecting a corpus, estimating a series of theoretically-informed semantic direction CMDs, and using the resulting document-by-CMD matrix to create a CoCA model. After briefly outlining this process, we then move on to some interpretation strategies for making sense of the schematic classes. Specifically, using the case of social identity schemas in U.S. political blogging, we discuss constructing correlation networks for the classes, determining dyad and node distinctiveness in those networks, assessing inversion, and locating and reading representative blog posts. We then cover what can be done with the CoCA classes once they are interpreted – namely, schema-covariate analysis – before concluding with a robustness check for checking schema invariance.

All data, lists, and necessary packages to reproduce the results in the R statistical computing environment are available in the paper repository.[9]

[9]The replication repository can be found here: https://github.com/Marshall-Soc/CoCA_interpret

3.1. Data and analysis

3.1.1. Data

For our demonstration, we use the publicly available "CMU 2008 Political Blog Corpus" collected by Jacob Eisenstein and Eric Xing, consisting of 13,246 posts (Eisenstein & Xing, 2010). These posts were collected from six political blogs for the full year of 2008 and were selected based on "the Technorati rankings of blog 'authority,' ideological balance, coverage for the full year 2008, and ease of access to blog archives" (Eisenstein & Xing, 2010, p. 1). The collectors rated three blogs as "Liberal" (Digby, ThinkProgress, and Talking Points Memo) and three blogs as "Conservative" (American Thinker, Hot Air, and Michelle Malkin).

We used pretrained fastText English word vectors (Bojanowski et al., 2017) for the embedding matrix.

3.1.2. Estimating engagement with key dimensions

The first step in our workflow is estimating semantic directions and calculating each post's overall engagement with either pole of that direction. As these posts are from political blogs during the 2008 U.S. general election, we estimate each post's engagement with a few dimensions of social identity which political bloggers might consider salient for framing political actors in the electoral process. These include: character (moral-immoral), gender (masculine-feminine), race (white-nonwhite), age (young-old), and social status (influential-uninfluential). We want to emphasize here that this illustration should not be taken as a robust analysis of social identity schemas, news media framing, and the U.S. electoral process, but only as an example of how a researcher might carry out a CoCA project.

We follow the same procedure as above to estimate the directions and calculate each post's engagement with one side or the other of the dimension (juxtaposed pairs for each can be found in Table A1 and Table A2 of Appendix A).

3.1.3. Estimating CoCA

Finally, after semantic direction CMDs are estimated for our 5 social identity dimensions, we pass them to the CCA algorithm (see Eqs. 5 and 6). The result is five schematic classes with $Q = 0.41$.[10]

3.2. Interpretation strategies

We use four procedures to aid in the interpretation of each class: (1) visualizing schematic classes as correlation networks, (2) determining dyad and node distinctiveness, (3) assessing within-class inversion, (4) and reading representative texts. We demonstrate each in turn below.

[10]Clauset and colleagues (Clauset et al., 2004, p. 2) find that a $Q > 0.3$ usually points to the presence of communities.

3.2.1. Correlation networks for conceptual classes

Following Goldberg (2011) and Boutyline (2017), a first step in interpreting class solutions is visualizing each as a network, where the vertices are dimensions and the edges are weighted by the correlations between them. For our five classes, we remove non-significant edges (at $\alpha = 0.05$); and, following Boutyline (2017), p. 380), we also removed edges with an absolute correlation below 0.15. The networks are shown in the left columns of Figure 1. Correlation matrix versions of the networks are shown in the right columns. (The figure also includes class labels, but we will hold off on discussing these until the end of the "Close Reading of Representative Blogs" section since this is where we put all of the pieces together to derive the class labels).

From this, we can see that Class #1 and #2 show an "omnivorous" network structure, in which a document's engagement with one dimension is highly predictive of their engagement on all other dimensions. Class #3 and Class #5 show an "anything but" structure, in which one dimension (Race for Class #3 and Gender for Class #5) is largely isolated from most of the other dimensions. To the extent a document may engage with these dimensions, it is not predictive of its engagement with other dimensions. Finally, Class #4 shows a "separate spheres" structure, in which a document's engagement with Status is predictive of Character (and vice versa), while a document's engagement with Age is predictive of Race and Gender (and vice versa); i.e., there is a very clear distinction between ascribed and achieved social identities.

3.2.2. Determining dyad and node distinctiveness

To aid in interpreting what is semantically distinct about each class, we begin by create dyad grids. We display this as stacked dyads in Figure 2. Specifically, we show which edge is most unique to a class as compared to other classes, by highlighting (as black and wider) which class has the largest absolute value for each unique dyad. Furthermore, we mark (as black) whichever node has the highest weighted degree centrality within each class. (Note that edge thickness/darkness may not perfectly match the edge thickness/darkness in Figure 1).

We can see that Gender has the highest degree centrality of any node in Class #1 and Class #2. Therefore, we know that Gender is important for organizing these omnivorous classes. We can also see that Class #1 has the highest correlations between Character and Gender, Character and Age, and Gender and Status of any class – specifically, Gender is negatively correlated with Character to a greater extent than any other class (e.g., engagement with masculinity correlates with engagement with immorality, and vice versa).

Turning to Class #4, our separate spheres class, we can clearly see the lack of ties between ascribed and achieved identities. Furthermore, Race is central to the cluster of ascribed identity dimensions, with the negative correlation between Race and Gender being the most distinctive of this class.

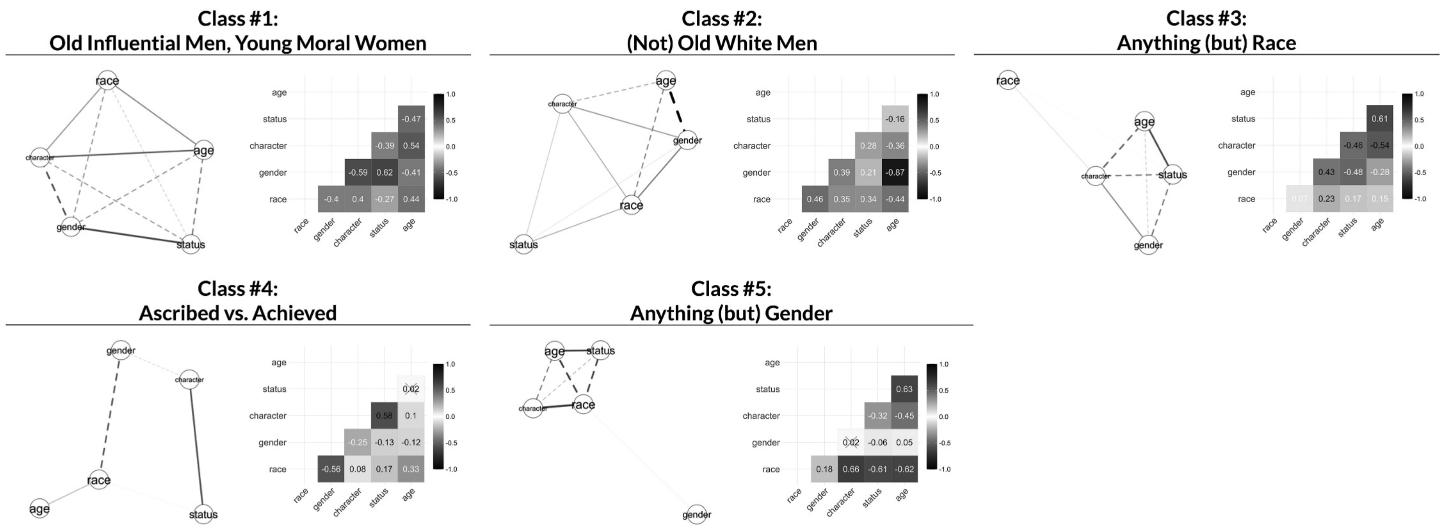

Figure 1. CoCA class inter-dimension correlation networks and matrices. The correlation networks in the left columns plot all statistically significant Pearson correlation coefficients ($\alpha = 0.5$, two-tailed). Correlations where $|r|_{ij} < 0.15$ are not visualized. Positive correlations are shown with straight edges; negative correlations are shown as dashed edges. Darker and thicker edges represent stronger bivariate correlations. The networks are spatially arranged using the Fruchterman-Reingold force-directed algorithm so that dimensions with stronger correlations are placed closer together and those with weaker correlations are "repulsed" from one another (with a repulsion factor equal to 1.86). The matrices in the right columns plot the same information, but in matrix form and with the cell colors weighted by the absolute value of the correlations. Xs in the cells indicate $p \geq .05$. $N = 2{,}395$ (class #1); $N = 4{,}650$ (class #2); $N = 1{,}859$ (class #3); $N = 2{,}489$ (class #4); $N = 1{,}853$.

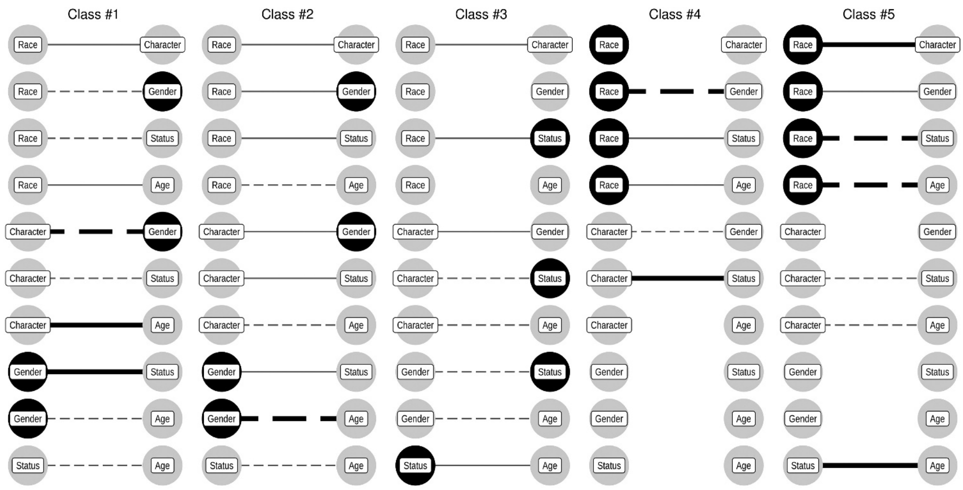

Figure 2. CoCA class dyad and node distinctiveness. The dyads include all statistically significant Pearson correlations ($\alpha = 0.05$, two-tailed). Correlations where $|r|_{ij} < 0.15$ are also not visualized. Positive correlations are shown with straight edges; negative correlations are shown as dashed edges. Black and thick lines are those that are most distinct across all classes for that dyad, determined by the greatest absolute value from zero. Black nodes are those that had the highest within-class weighted degree centrality. N = 2,395 (class #1); N = 4,650 (class #2); N = 1,859 (class #3); N = 2,489 (class #4); N = 1,853 (class #5). The class labels are as follows: "old influential men, young moral women," "(not) old white men," "anything (but) race," "ascribed vs. achieved," and "anything (but) gender." Edge thickness/darkness may not perfectly match the edge thickness/darkness in Figure 1.

For Class #5, an anything-but (Gender) class, Race is not only the highest within class centrality, but also included in three (of four) distinct dyads. Therefore, any interpretation of Class #5 will focus on the documents' engagement with the Race dimension. In other words, how blogs in this class talk about gender is not predictive of engagement along any other dimension, but how they talk about race is very predictive of other dimensions.

Finally, consider Class #3, an anything-but (Race) structure, which has no distinctive ties – meaning that its dimensions are relatively weakly correlated. We can interpret this as low overall semantic organization (Martin, 2002; Van Gunten et al., 2016). We can also see that the node with the highest degree is Status – as compared to the other classes for which Race and Gender are central.

3.2.3. Assessing inversion

Inversion is the extent to which any two document vectors (i.e., document scores across semantic directions) are diametrically opposed. At its extreme, inversion is a correlation of -1 between two documents. When interpreting CoCA output, just as with CCA, researchers should take into account the role of inversion in their class solutions. As Boutyline (2017, pp. 371–373) notes, inversion may not make sense as an indicator of shared schemas in some application domains. Furthermore, even when inversion does make sense as a relevant linear transformation, this does not mean that inversion is present in any sort of meaningful sense. For instance, looking at the Class #3 correlation network in Figure 1 might say this class includes blog posts that "engage the concepts of influence, goodness, femininity, whiteness, and youth," as well as those that "engage the concepts non-influence, wrongness, masculinity, non-whiteness, and old age." To make this interpretation, however, one must assume that there is some degree of non-trivial sign change in the inter-document vector transformations. Could it be, then, that Class #3 consists of documents that mostly engage in only one of those two patterns?

We propose one simple method for assessing class-specific inversion. Since negative inter-document correlations present the clearest case of inversion, an initial assessment of inversion involves simply counting the number of negative correlations in each class-specific inter-document correlation matrix. We do that for our 5-class CoCA model in Table 1. Specifically, for each class, we generated the inter-document correlation matrix (the same one used to partition the classes, but before taking the absolute value), removed any statistically non-significant correlations (at $\alpha = .05$, two-tailed), and then tallied the number of negative correlations. For example, 346,128 of the 736,734 statistically significant correlations in Class #3 are negative – i.e., about 47%, or almost half of the document correlations. Of course, it is important to note that two documents in the same class can exhibit *some degree* of inversion and still have a positive r_{ij} – it all depends on how (dis)similar they are, in the aggregate, across the K dimensions they are being compared. Nonetheless, negative inter-document correlations present clear cases of inversion; this sort of descriptive analysis offers a useful baseline.

Though helpful, using these percentages to make decisions about whether or not to use inversion language when interpreting a schematic class can still be difficult because of the lack of benchmarks about what percentages constitute trivial, normal, or above-average amounts of inversion. While we hope

Table 1. Inversion per class.

	# of Negative $\rho_{ij\|C}$	Total $\rho_{ij\|C}$	% of Negative $\rho_{ij\|C}$
Class #1	509,192	1,178,528	43.2%***
Class #2	1,828,200	4,262,578	42.9%***
Class #3	400,650	838,820	47.8%***
Class #4	346,128	736,734	47.0%***
Class #5	433,696	1,001,612	43.3%***

The first column lists the total number of statistically significant ($p < .05$) inter-document Pearson correlation coefficients per class that are negative. The second column lists the total number of statistically significant correlations regardless of sign. The third column lists the class-specific proportion of statistically significant Pearson correlations that are negative. The significance levels are derived from one-sample difference-of-proportion z-tests of the null hypothesis that the true population percentage is equal to or less than 10%. The class labels are as follows: "old influential men, young moral women," "(not) old white men," "anything (but) race," "ascribed vs. achieved," and "anything (but) gender."
***$p < .001$ (right-tailed tests).

that these substantive benchmarks may emerge as new SCA studies are published, we can – if we assume that our documents are an unbiased random sample of possible documents – test the null hypothesis that the true percentage of negative correlations is equal to or less than some percentage that we treat as indicating trivial inversion. Say, for example, that we treat 10% as the percentage at or below which we agree that inversion in a class is too small to warrant references to inversion when interpreting a schematic class. We can then use a simple one-sample difference-of-proportions z-test to test the null that the true proportion of negative correlations in the population of documents from which our sample was randomly selected is equal to or less than 10%. The results of these hypothesis tests are also shown in Table 1.[11]

In the present case, nearly 50% of the inter-document correlations comprising each of the five classes are negative. It would seem that inversion is present enough within the classes to warrant including inversion in the schema interpretation task.

3.2.4. Close reading of representative Blogs

Drawing on the prior methods to summarize the underlying semantic structure of each schematic class, we select posts which are particularly high (or low) on salient dimensions within each class. Here, we focus on only a couple examples for illustration purposes.

Consider Class #2, which was an omnivorous network with a core focus on Gender and with a relatively high amount of inter-document negative correlation (43% – our conservative estimate of inversion). Below are excerpts from two posts from Class #2 – the first among the highest engaging the "masculine" pole of the dimension, and the second engaging the "feminine" pole:

> Some of you may remember me writing about my conservative Dad–retired military, very right wing, very political. He's the type of guy who thinks General Curtis Lemay had the right idea about bombing Hanoi back into the stone age. … McCain should be my Dad's dream candidate–naval war hero POW, military obsessive, "go-to-hell" macho attitude. … I assumed the macho military thing would trump all. If they've lost my Dad, they really are in trouble. (Parton, 2008)

> Yesterday in her interview with NBC, Gov. Sarah Palin (R-AK) told NBC's Brian Williams that she rejected the "label" of being called a feminist … From a Sept. 30 interview with CBS's Katie Couric: COURIC: Do you consider yourself a feminist? PALIN: I do. A feminist who believes in equal rights. (Terkel, 2008)

As another example, consider Class #5. Race was central in this class, forming a tight (positive) relationship with Character, as well as (negative) ties to Status and Age. When reading through the posts in Class #5 with the highest engagement with the "non-white" end of the Race dimension, a clear pattern emerges where bloggers

[11] Of course, these tests are no substitute for substantive benchmarks for "how much" correlation-level inversion is a little, normal, or a lot. In fact, we even have to assume a percentage as indicating low inversion in order to even perform the hypothesis test! These sorts of benchmarks will (hopefully) become more clear as more SCA studies are published – across RCA, CCA, and CoCA.

write about the United States' relationship with other regions and countries: in particular, those in the Middle East and Latin America. For example, here is an excerpt from a blog post from Class #5:

> John McCain has come out with this new Spanish-language ad in Florida, usuing [sic] a tried-and-true tactic for courting Miami Cubans: Linking your opponent to Latin American Marxism, in this case Castro-ally Hugo Chàvez. (Kleefeld, 2008)

Exploring this dimension reveals the ways that Race (estimated as a white to nonwhite dimension) is coded into political discourse such that presumably domestic issues (e.g., the housing crisis) are associated with "white" while foreign relations are coded as "nonwhite."

Taking these various interpretation strategies together, we might arrive at the following substantive labels for Class #1 through #5, respectively: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender."

3.3. Schema-Covariate Analysis

The schematic associations a person accrues over time is a function of their lived experiences, and people will share similar schemas to the extent they have similar such experiences. Since people who occupy similar positions in social space are more likely to engage in similar practices (Bourdieu, 1990), researchers using a schematic class analysis technique typically examine how schematic class membership varies as a function of other sociological covariates. This work, in the case of RCA/CCA, usually involves using a multinomial logit/probit model[12] to regress the (categorical) class membership variable on a matrix of sociodemographic variables. So long as corpus covariates are available, this same logic applies to document classes derived from CoCA. Furthermore, in addition to prior experiences, schematic associations may be activated by objects of present attention. Therefore, we can also use features of the texts – such as people or organizations referenced – as predictors of class membership.

To illustrate this, we ran a multinomial logit regression model with our five-class categorical measure as the outcome variable and the following independent/control variables: publication date (specifically, the number of days into the year a post was written), ideological leaning (liberal or conservative, based on the rating of the blog source), whether or not Barack Obama, John McCain, Hillary Clinton, Sarah Palin, Joe Biden, or one of the Bush figures[13] were mentioned in the post (each getting their own indicator variable, where 0 = no and 1 = yes),[14]

[12]Or a binary logit/probit model, if the CoCA results in just two classes.
[13]Most of the references are George W. Bush, but there is likely a margin of error here.
[14]We used Named-Entity Recognition to determine the absence or presence of particular people in each post. Specifically, we used the entity() function in the trinker R package, which uses the Apache OpenNLP library to identify persons, organizations, and locations. We also lightly cleaned the output by hand to verify our key presidential and vice-presidential candidates were identified.

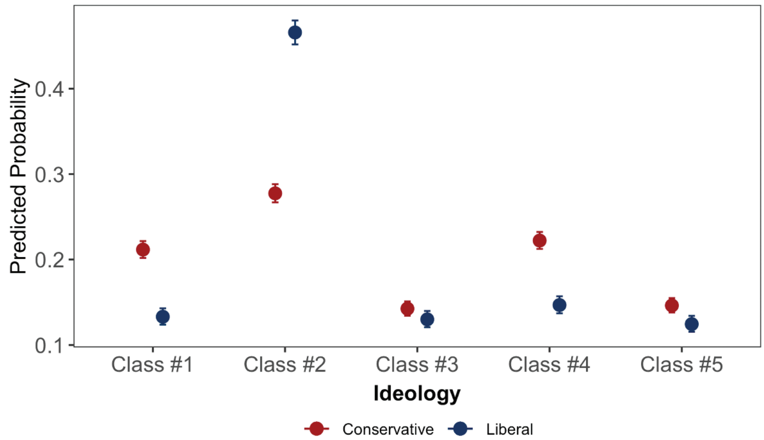

Figure 3. Predicted probability of schematic class membership by blog source political ideology. Predicted probabilities are from a multinomial logit model. All class-specific predicted probability contrasts are statistically significant at least at $p < .001$ (two-tailed). The prediction equation used the sample median for the publication date, the mean for the mentioned-persons variable, and averaged over the Obama, McCain, Clinton, Palin, Biden, and Bush indicator variables. Bars are 95% confidence intervals. The class labels are as follows: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender."

and the number of unique people mentioned in the post. We include an interaction term between ideology and publication date, and between ideology and each political figure attention variable. The regression table is provided in Appendix B as Table B1. It is important to note that these models are meant to illustrate the workflow and should not be interpreted as theoretically-driven models of differences in social identity schemas.

Figure 3 plots the predicted probabilities of a post being in each schematic class as a function of political ideology, separated by political ideology (further details on the prediction equation specification are in the figure note). The liberal-conservative predicted probability contrast (using a Wald test) is statistically significant for each class (at least $\alpha = .001$, two-tailed). The plot suggests that conservative posts are more likely than liberal posts to occupy Classes #1, (somewhat) #3, #4, and (somewhat) #5, while the opposite is true for Class #2. That said, it appears that the highest membership probability for both conservative and liberal posts goes to Class #2 ($0.27 \leq P_c \leq 0.29$ and $0.45 \leq P_l \leq 0.48$).[15]

The left panel of Figure 4 shows predicted probabilities by whether the blog post mentioned Clinton and/or Palin vs. Obama, McCain, Biden, and/or Bush.[16]

[15]Class #2 is also the largest class with 4,650 posts – about 35% of the corpus documents. It is also the most equitably distributed class across ideological lines, with 2,100 conservative posts (45% of the class) and 2,550 liberal posts (55% of the class).

[16]The model used to derive gender coefficients was different from the models controlling for each candidate by name. In this "gender model," the independent variable specification was: whether the blog post mentioned either female figure (Clinton or Palin), the number of unique people mentioned in the post, political ideology, and publication date. The regression table is Table B2 in Appendix B.

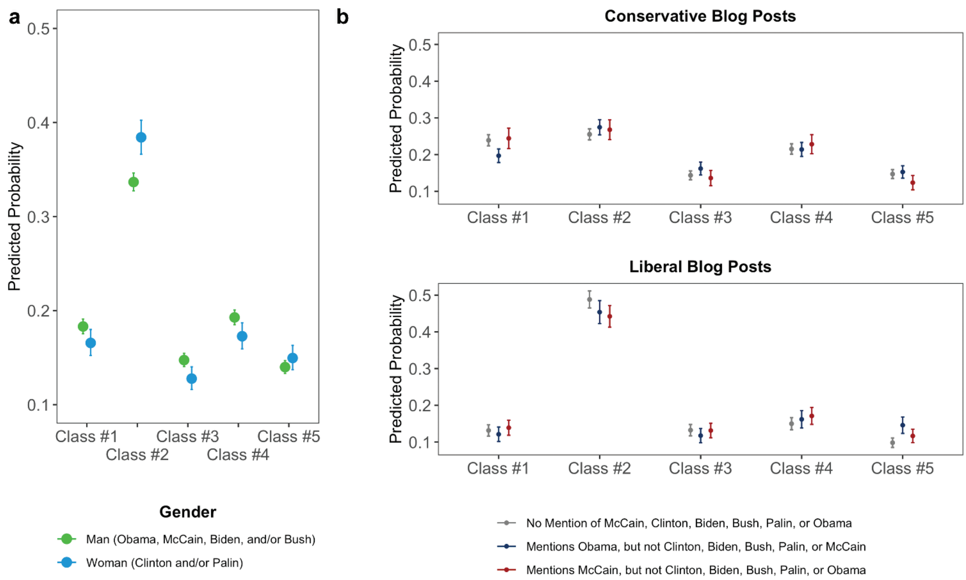

Figure 4. Predicted probability of schematic class membership by (a) gender of mentioned political figures and (b) blog source political ideology and post-level candidate attention. Note: Predicted probabilities are from a multinomial logit model (one for the left panel; one for the right panel – see footnote #16). In both models, probabilities were estimated with the median publication date and the mean for the mentioned-persons variable. For the right-hand panel models, the "name-dropping Obama" predictions, the McCain, Clinton, Palin, Biden, and Bush indicator variables were set at 0. For the "name-dropping McCain" predictions, the Obama, Clinton, Palin, Biden, and Bush indicator variables were set at 0. Bars are 95% confidence intervals. the class labels are as follows: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender."

Given our class interpretations, we should expect gender to be predictive of membership in Class #1 and/or Class #2, which correspond to the "Old Influential Men, Young Moral Women" and "(Not) Old White Men" social identity schemas. As the left panel shows, mentioning *either* these men or women is associated with a higher probability of a post being in Class #2, but especially mentioning Clinton or Palin.[17] The right-hand panels plot predicted probabilities of a post giving attention to specific candidates (Obama or McCain), separated by ideology of the post (details on the prediction equation specifications are in the figure note). Multiple insights could be gleaned from this plot, one being that posts focusing attention on Obama are not different from posts that focus attention on McCain when it comes to their likelihood of being in any given class.

Figure 5 plots show the predicted probabilities of class membership across 2008, separated again by ideology. None of the ideology-by-time interaction terms are statistically significant ($p \geq .05$) across any of the class membership logit contrasts; thus they mostly parallel predicted

[17] As a reviewer notes, it also possible that any given document has a higher probability of being in Class #2 because it is the highest frequency class and not because there is a larger association between Class #2 membership and gendered name-dropping. A more substantive analysis of these data would need to tease these possibilities apart.

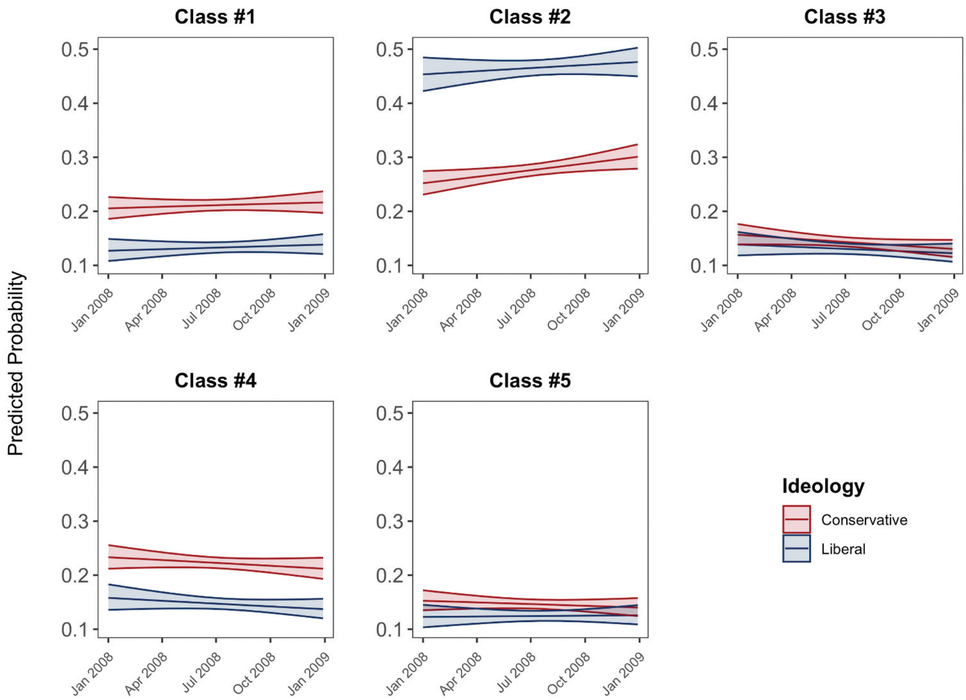

Figure 5. Predicted probability of schematic class membership by blog source political ideology and publication date. Note: None of the ideology-by-time interaction terms are statistically significant across any of the DV category contrasts. Predicted probabilities are from a multinomial logit model. Probabilities were estimated with the the mean for the mentioned-persons variable, and averaged over the Obama, McCain, Clinton, Palin, Biden, and Bush indicator variables. Bars are 95% confidence intervals. The class labels are as follows: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender."

time trends between liberal and conservative posts for each class. A takeaway from these plots might be that the likelihood of class membership does not vary across time for either liberal or conservative blogs.

Table 2 provides one final example. The table cross-tabulates class assignment and blog source. The table suggests a nonrandom association between schematic class and blog source ($\chi^2 = 2{,}545.68$; $p < .001$), meaning that individuals who write for these publications tap into different social identity schemas in their coverage of 2008 U.S. politics. Heather Digby's posts, for example, are disproportionately represented in Class #2 – that is, the posts that follow an omnivorous schema in which Gender is central and tends to be inversely associated with Age.

Table 2. Class-by-publication cross-tabulation.

	Conservative			Liberal			
	Am Thinker	Hot Air	Michelle Malkin	Digby	Think Progress	TP Memo	Row Total
Class #1	605	917	117	63	392	301	2,395
	(4.57)	(0.48)	(6.92)	(0.88)	(2.96)	(2.27)	
Class #2	864	1,042	194	1,563	436	551	4,650
	(6.52)	(11.80)	(7.87)	(1.46)	(3.29)	(4.16)	
Class #3	480	483	123	70	502	201	1,859
	(3.62)	(0.53)	(3.65)	(0.93)	(3.79)	(1.52)	
Class #4	699	789	171	64	446	320	2,489
	(5.28)	(0.48)	(5.96)	(1.29)	(3.37)	(2.42)	
Class #5	549	477	72	119	304	332	1,853
	(4.15)	(0.90)	(3.60)	(0.54)	(2.30)	(2.51)	
Col Total	3,197	3,708	677	1,879	2,080	1,705	13,246

χ^2 = 2,545.68*** (*df* = 20). Numbers in parantheses are percents of the total count. The minimum expected cell frequency is 94.71. The class labels are as follows: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender."
***$p < .001$.

3.4. Addressing schema invariance

Like any model, it is important to follow CoCA with robustness checks. We limit our discussion to graph partitioning sensitivities and propose the multiple group test of schema invariance as a useful procedure.

As Boutyline and Vaisey argue in their paper on belief network analysis (Boutyline & Vaisey, 2017, p. 1429), modularity maximization suffers from a "resolution limit" wherein it can systematically fail to pick up on small partitions and/or or large partitions. The consequence of this for CoCA (and CCA and RCA) is that maximizing modularity as the objective function for graph community detection can bias the model toward finding $k > 1$ schematic classes when, in fact, none (or only one) exists. It is necessary, then, to follow CoCA with a check against the possibility that there is schema invariance – i.e., only "one schema" – in the data.

Boutyline and Vaisey propose the multiple group test from structural equation modeling to do this (Boutyline & Vaisey, 2017, Appendix B). The logic behind the multiple group test here is to assess whether the observed data are better fit using a constrained model (where schematic classes are forced to have the same variances and covariances across the semantic direction CMDs in the CoCA model) or a fully saturated model (where each class is allowed have its own variances and covariances). The relative fit of each model can then be compared using standard information criteria and likelihood ratio tests (2017, Appendix B; 2020, Appendix B).

A multiple group test of schema invariance with our five-class solution on the political blogs is reported in Table 3. Model #1 is the constrained model; model #2 is the saturated model. If model #1 fits the data better than model #2, then there is evidence that a no-class solution more accurately produces the

Table 3. Multiple group test of schema invariance.

	$-2LL$	AIC	Δ_{2-1}AIC	BIC	Δ_{2-1}BIC
Model #1	174,390.32	174,470.3	–	174,770	–
Model #2	151,539.64	151,739.6	−22,730.7	152,488.8	−22,281.2

N = 13,236 for both models.

Table 4. Multiple group test comparing *k*-class partitions.

	$-2LL$	AIC	Δ_{5-k}AIC	BIC	Δ_{5-k}BIC
2 classes	165,478.74	165,558.7	−13,819.1	165,858.4	−13,369.6
3 classes	158,087.68	158,207.7	−6,468.1	158,657.2	−6,168.4
4 classes	155,001.72	155,161.7	−3,422.1	155,761	−3,272.2
5 classes	151,539.64	151,739.6	–	152,488.8	–

N = 13,236 for all models.

observed data (suggesting schema invariance); if the reverse is true, then there is evidence of schematic heterogeneity across the documents.

Both the AIC and BIC are smaller for model #2, indicating that allowing the schematic classes to have their own variances and covariances fits better than a model where they are constrained to not vary ($\Delta_{2-1}\mathrm{AIC} = -22,730.7$; $\Delta_{2-1}\mathrm{BIC} = -22,281.2$).

If the estimated k-class partition fits the data better than a no-class null model (as this one does), then we can also move on to comparing different k partitions since modularity maximization might result in diminishing returns for model fit with finer-grained groupings (DiMaggio & Goldberg, 2018, Appendix B; McDonnell et al., 2020, Appendix B). For example, the 5-class solution Q appears only marginally higher than the 4-class solution Q: 0.4129 vs. 0.4126. Is this a meaningful difference in terms of the resulting classes? We can extend the multiple-group test of schema invariance to assess this.[18]

Now we can compare multiple saturated models (instead of comparing a saturated model and a constrained null model). In other words, we can again use information criteria to see if higher-order partitions provide a better fit. Table 4 summarizes the results comparing the estimated 5-class solution against 4-, 3-, and 2-class solutions. The 5-class solution appears to fit the data better relative to each lower-order solution, per the $\Delta_{5-k}AIC$ and $\Delta_{5-k}BIC$ statistics. Were there evidence that a lower-order partition provided a better fit – pointing to diminishing returns in maximizing Q—we could simply have CoCA stop the modularity maximization procedure at an earlier iteration.

[18]Another robustness check useful for this purpose is the gap statistic, which DiMaggio and Goldberg extended to RCA (2018, Appendix B). However, the gap statistic computation procedure requires a considerable number of bootstrapped samples to create partitions from null distributions against which to compare the true partition, and this must be repeated for a series of *k* partitions. This means that calculating the gap statistic for even a small (e.g., four or five) set of partition solutions requires a considerable amount of computational power when the dataset is large.

4. Conclusion

We demonstrated a workflow for interpreting and analyzing Concept Class Analysis (CoCA) output – a kind of schematic class analysis adapted to the analysis of documents – using the case of social identity schemas in a collection of over 13,000 U.S. political blog posts. The workflow proceeds as follows: After selecting and preparing our corpus, we must (1) define two or more theoretically-driven semantic directions. Next, (2) we calculate concept mover's distance (CMD scores for each document in our corpus along each semantic direction). This produces a document-by-CMD matrix. We then (3) pass this matrix to the CCA algorithm to find conceptual classes. Next, we must interpret the results. This involves (5) visualizing schematic classes as correlation networks; (6) determining dyad and node distinctiveness with stacked dyads and weighted degree centrality, respectively; (7) assessing within-class inversion by counting negative correlations in each class; (8) and, a close reading of representative texts. Finally, we can turn to the analysis of schematic class membership by covariates of interest. These covariates may be external to the text – such as attributes of the individuals or organizations involved in producing a given text – or internal to the text – such as references to specific entities. Finally, it is useful to note that with slight modification, much of this workflow can be adapted to any schematic class analysis.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Marshall A. Taylor http://orcid.org/0000-0002-7440-0723
Dustin S. Stoltz http://orcid.org/0000-0002-4774-0765

References

Arseniev-Koehler, A., & Foster, J. G. (2022). Machine learning as a model for cultural learning: Teaching an algorithm what it means to be fat. *Sociological Methods & Research*, *51*(4), 1484–1539. https://doi.org/10.1177/00491241221122603

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Bourdieu, P. (1990). *The logic of practice*. Stanford university press.

Boutyline, A. (2017). Improving the measurement of shared cultural schemas with correlational class analysis: Theory and method. *Sociological Science*, *4*(15), 353–393. https://doi.org/10.15195/v4.a15

Boutyline, A., Cornell, D., & Arseniev-Koehler, A. (2021). All roads lead to polenta: Cultural attractors at the junction of public and personal culture. *Sociological Forum*, *36*(S1), 1419–1445. https://doi.org/10.1111/socf.12760

Boutyline, A., & Johnston, E. (2023). *Forging better axes: Evaluating and improving the measurement of semantic dimensions in word embeddings*. https://osf.io/preprints/socarxiv/576h3

Boutyline, A., & Soter, L. K. (2021). Cultural schemas: What they are, how to find them, and what to do once you've caught one. *American Sociological Review*, *86*(4), 728–758. https://doi.org/10.1177/00031224211024525

Boutyline, A., & Vaisey, S. (2017). Belief network analysis: A relational approach to understanding the structure of attitudes. *The American Journal of Sociology*, *122*(5), 1371–1447. https://doi.org/10.1086/691274

Carbone, L., & Mijs, J. (2022). Sounds like meritocracy to my ears: Exploring the link between inequality in popular music and personal culture. *Information Communication & Society*, 1–19.

Clauset, A., Newman, M. E., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111

DiMaggio, P., & Goldberg, A. (2018). Searching for homo economicus: Variation in Americans' construals of and attitudes toward markets. *European Journal of Sociology/ Archives Européennes de Sociologie*, *59*(2), 151–189. https://doi.org/10.1017/S0003975617000558

Durrheim, K., Schuld, M., Mafunda, M., & Mazibuko, S. (2022). Using word embeddings to investigate cultural biases. *The British Journal of Social Psychology / The British Psychological Society*. https://doi.org/10.1111/bjso.12560

Eisenstein, J., & Xing, E. (2010). *The CMU 2008 political blog corpus*. Carnegie Mellon University, School of Computer Science, Machine Learning Department.

Goldberg, A. (2011). Mapping shared understandings using relational class analysis: The case of the cultural omnivore reexamined. *The American Journal of Sociology*, *116*(5), 1397–1436. https://doi.org/10.1086/657976

Hunzaker, M. B. F., & Valentino, L. (2019). Mapping cultural schemas: From theory to method. *American Sociological Review*, *84*(5), 950–981. https://doi.org/10.1177/0003122419875638

Jones, J. J., Amin, M. R., Kim, J., & Skiena, S. (2020). Stereotypical gender associations in language have decreased over time. *Sociological Science*, *7*, 1–35. https://doi.org/10.15195/v7.a1

Kleefeld, E. (2008). *McCain Spanish ad: Obama wants to talk with anti-american Hugo Chávez*. https://web.archive.org/web/20080920083429/http://tpmelectioncentral.talkingpointsmemo.com/2008/09/mccain_spanish_ad_obama_wants.php

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949. https://doi.org/10.1177/0003122419877135

Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). From word embeddings to document distances. *Proceedings of the 32nd International conference on machine learning*, 37, 957–966.

Leschziner, V., & Brett, G. (2021). Have schemas been good to think with? *Sociological Forum*, *36*(S1), 1207–1228. https://doi.org/10.1111/socf.12767

Martin, J. L. (2002). Power, authority, and the constraint of belief systems. *AJS; American Journal of Sociology*, *107*(4), 861–904. https://doi.org/10.1086/343192

McDonnell, E. M., Stoltz, D. S., & Taylor, M. A. (2020). Multiple market moralities: Identifying distinct patterns in how consumers evaluate the fairness of price changes. *Socio-Economic Review*. https://doi.org/10.1093/ser/mwaa034

Mikolov, T., Yih, W.-T., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies*. 746–751.

Morehouse, K., Rouduri, V., Cunningham, W., & Charlesworth, T. (2023). *Traces of human attitudes in contemporary and historical word embeddings (1800–2000)*. https://www.researchsquare.com/article/rs-2922677/v1.

Newman, M. E. (2004). Analysis of weighted networks. *Physical Review E*, *70*(5), 056131. https://doi.org/10.1103/PhysRevE.70.056131

Newman, M. E. (2006). Finding community structure in networks using the eigen- vectors of matrices. *Physical Review E*, *74*(3), 036104. https://doi.org/10.1103/PhysRevE.74.036104

Parton, M. D. (2008). *The wisdom of the ancients*. https://digbysblog.net/?s=%22Some+of+you+may+remember+me+writing+about+my+conservative+Dad%22

Rubner, Y., Tomasi, C., & Guibas, L. J. (1998). A metric for distributions with applications to image databases. *Sixth international conference on computer vision* (pp. 59–66). Bombay, India: IEEE.

Stoltz, D. S., & Taylor, M. A. (2019). Concept mover's distance: Measuring concept engagement via word embeddings in texts. *Journal of Computational Social Science*, *2*(2), 293–313. https://doi.org/10.1007/s42001-019-00048-6

Stoltz, D. S., Taylor, M. A., & Dudley, J. S. K. (2024). A tool kit for relation induction in text analysis. *Sociological Methods & Research*. https://doi.org/10.1177/004912412412332

Taylor, M. A., & Stoltz, D. S. (2020a). Concept class analysis: A method for identifying cultural schemas in texts. *Sociological Science*, *7*, 544–569. https://doi.org/10.15195/v7.a23

Taylor, M. A., & Stoltz, D. S. (2020b). Integrating semantic directions with concept mover's distance to measure binary concept engagement. *Journal of Computational Social Science*.

Terkel, A. (2008). *Palin was a feminist before she wasn't*. https://web.archive.org/web/20090306023200/https://thinkprogress.org/2008/10/24/palin-feminist/

Van Gunten, T. S., Martin, J. L., & Teplitskiy, M. (2016). Consensus, polarization, and alignment in the economics profession. *Sociological Science*. https://doi.org/10.15195/v3.a45

Wood, M. L., Stoltz, D. S., Van Ness, J., & Taylor, M. A. (2018). Schemas and frames. *Sociological Theory*, *36*(3), 244–261. https://doi.org/10.1177/0735275118794981

Appendices

Appendix A. Juxtaposition Pairs

Table A1. Juxtaposed terms for race, age, and gender semantic direction.

RACE		AGE		GENDER	
white	black	young	old	man	woman
whites	blacks	youth	aged	men	women
european	african	current	ancient	he	she
caucasian	african	juvenile	elderly	him	her
white	latino	immature	mature	his	her
whites	latinos	junior	senior	his	hers
european	latino	novice	veteran	boy	girl
caucasian	hispanic	new	old	boys	girls
white	asian	green	seasoned	male	female
whites	asians	budding	withering	masculine	feminine
european	asian	fledgling	expert		
caucasian	asian	beginning	ending		

Table A2. Juxtaposed terms for character and status semantic direction.

CHARACTER		STATUS	
good	evil	influential	uninfluential
moral	immoral	honorable	dishonorable
good	bad	esteemed	lowly
honest	dishonest	reputable	disreputable
virtuous	sinful	distinguished	commonplace
virtue	vice	eminent	mundane
righteous	wicked	illustrious	humble
chaste	transgressive	renowned	prosaic
principled	unprincipled	acclaimed	modest
unquestionable	questionable	dignitary	commoner
noble	nefarious	venerable	unpretentious
uncorrupt	corrupt	exalted	ordinary
scrupulous	unscrupulous	estimate	lowly
altruistic	selfish	prominent	common
chivalrous	knavish		
honest	crooked		
commendable	reprehensible		
pure	impure		
dignified	undignified		
holy	unholy		
valiant	fiendish		
upstanding	villainous		
guiltless	guilty		
decent	indecent		
chaste	unsavory		
righteous	odious		
ethical	unethical		

Appendix B. Multinomial Logit Models

Table B1. Log odds estimates of class membership, relative to class #1.

	Class #2		Class #3		Class #4		Class #5	
	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$
Ideology	1.334	(0.138)***	0.502	(0.167)**	0.279	(0.161)	0.153	(0.172)
Day	0.000	(0.000)	−0.000	(0.000)	−0.000	(0.000)	−0.000	(0.000)
Name-Drops	−0.002	(0.007)	0.022	(0.007)**	0.013	(0.007)	0.022	(0.007)**
Cand. Attn.								
Obama	0.265	(0.075)***	0.316	(0.088)***	0.190	(0.079)*	0.234	(0.088)**
McCain	0.028	(0.084)	−0.078	(0.100)	0.036	(0.088)	−0.197	(0.101)
Clinton	0.288	(0.097)**	0.035	(0.115)	0.103	(0.103)	0.529	(0.115)
Palin	−0.069	(0.150)	−0.420	(0.199)*	−0.297	(0.166)	−0.379	(0.195)
Biden	−0.130	(0.158)	−0.215	(0.197)	0.156	(0.161)	0.074	(0.184)
Bush	0.179	(0.094)	0.124	(0.111)	0.216	(0.098)*	0.351	(0.107)***
Ideology ×								
Day	−0.000	(0.000)	0.000	(0.000)	−0.000	(0.000)	0.000	(0.000)
Obama	−0.255	(0.120)*	−0.351	(0.146)*	−0.028	(0.138)	0.246	(0.146)
McCain	−0.182	(0.122)	0.013	(0.148)	0.039	(0.139)	0.313	(0.150)*
Clinton	−0.107	(0.153)	−0.356	(0.193)	−0.582	(0.186)**	−0.060	(0.186)
Palin	0.117	(0.219)	0.608	(0.275)*	0.363	(0.252)	0.666	(0.266)*
Biden	0.046	(0.186)	0.731	(0.216)***	0.261	(0.205)	0.931	(0.197)***
Bush	−0.363	(0.127)**	−0.020	(0.152)	−0.468	(0.143)**	−0.632	(0.151)***
Constant	0.006	(0.082)	−0.503	(0.095)***	−0.095	(0.086)	−0.532	(0.095)***

Numbers in parentheses are standard errors. $-2LL = 40012.64$. AIC = 40148.64. Likelihood-ratio $\chi^2 = 827.72$*** (df = 64). McFadden pseudo-$R^2 = 0.020$. $N = 13{,}246$. The class labels are as follows: "old influential men, young moraL women," "(Not) old white men," "anything (but) race," "ascribed vs. achieved," and "anything (but) Gender." class #1 (""old influential Men, young moral women") is the reference class.

*$p < .05$.

**$p < .01$.

***$p < .001$ (two-tailed).

Table B2. Log odds estimates of class membership, relative to class #1.

	Class #2		Class #3		Class #4		Class #5	
	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$	$\hat{\beta}$	$se(\beta)$
Ideology	0.962	(0.054)***	0.388	(0.066)***	0.051	(0.063)	0.337	(0.066)***
Day	0.000	(0.000)	−0.000	(0.000)	−0.000	(0.000)	0.000	(0.000)
Name-Drops	−0.000	(0.007)	0.029	(0.007)***	0.018	(0.007)**	0.029	(0.007)***
Fem. Attn.	0.232	(0.064)***	−0.043	(0.078)	−0.009	(0.073)	0.167	(0.076)*
Constant	0.198	(0.065)**	−0.434	(0.077)***	−0.018	(0.071)	−0.608	(0.078)***

Numbers in parentheses are standard errors. $-2LL = 40266.63$. AIC = 40306.63. Likelihood-ratio $\chi^2 = 573.74$*** (df = 16). McFadden pseudo-$R^2 = 0.014$. $N = 13{,}246$. The class labels are as follows: "Old Influential Men, Young Moral Women," "(Not) Old White Men," "Anything (but) Race," "Ascribed vs. Achieved," and "Anything (but) Gender." Class #1 (""Old Influential Men, Young Moral Women") is the reference class.

*$p < .05$.

**$p < .01$.

***$p < .001$ (two-tailed).