Teaching Two Americas: A Machine Learning-Based Comparative Analysis of Political Framing in California and Florida Textbooks

Khetan, Sneha (School: North Carolina School of Science and Mathematics)

Curricula shape how societies transmit knowledge, values, and collective memory. In the United States, textbook adoption occurs at the state level, producing variation in what students learn about history, civics, and the world. This study presents a computational comparative analysis of 16 state-adopted middle school social studies textbooks from California and Florida (2017–2023), two of the largest and most ideologically divergent states in contemporary U.S. politics. A 2.2-million-word dataset was constructed from matched textbook pairs by the same publishers and subjects to control for pedagogical variation and isolate ideological differences in language. Text alignment analysis showed that 96% of content in matched editions was identical, reflecting publishers' reliance on a shared base text. The remaining 4%, comprising 603 pairs of divergent passages, contained systematic ideological differences and stylistic variation. To analyze these passages, I applied a transformer-based machine learning classifier, RoBERTa-POLITICS, trained on 3.7 million U.S. political news articles. Results revealed ideological polarization across states. California passages more often leaned progressive, while Florida passages leaned conservative or centrist in framing. Three forms of divergence were identified: additions or redactions (51%), rephrasings (35%), and stylistic edits (14%). Divergences were strongest in Civics Education, followed by U.S. History, with subtler contrasts in World History. More than half of the differences centered on structural racism, civil rights and activism, gender and sexuality, and Indigenous histories. Overall, these findings show that even small textual edits can systematically reshape ideological framing, revealing political influence in curricula.

Awards Won: