Separation of Church and State Curricula? Examining Public and Religious Private School Textbooks*

Anjali Adukia
University of Chicago
and NBER

Emileigh Harrison
University of Chicago

February 2024
Preliminary Working Draft. Please do not distribute.

Abstract

Curricular materials not only impart academic knowledge but also prepare children for citizenship by teaching them societal values. Consequently, changes in curriculum have been shown to impact children's worldviews and academic outcomes. Given the importance of what we teach our children, we know very little about the educational content children are exposed to outside the public school system. In particular, there is little to no research on what children are taught in religious educational settings. Our study begins to fill in this knowledge gap by comparing the content of influential religious textbooks with the content of state-adopted textbooks from Texas and California. Specifically, we apply and improve upon methods of computer-driven content analysis to quantitatively measure *what* topics and values are incorporated, *who* is represented, and *how* various topics, values, and people are portrayed within different sets of educational materials over time.

*Contact: Anjali Adukia, adukia@uchicago.edu; Emileigh Harrison, harrisone@uchicago.edu. For helpful feedback, we thank Ben Arold, Marianne Bertrand, Dan Black, Ethan Bueno de Mesquita, Michael Dinerstein, Steven Durlauf, Alex Eble, Sendhil Mullainathan, Thomas Wollmann, David Yanagizawa-Drott, and seminar participants at AEFP, APPAM, SREE, The University of Chicago, University of Michigan, and the Zurich AI and Economics conference. For excellent research support, we thank Sam Abers, Matt Bonci, Elisa Chen, Paula Dastres Gallardo, Benjamin Leiva, Ayush Raj, Celia Zhu, and other members of the MiiE Lab. For financial support, we thank the National Academy of Education, Spencer Foundation, and Becker Friedman Institute at the University of Chicago.

I Introduction

Curricular materials impart knowledge: they both teach academic topics and concepts and prepare children for citizenship by teaching them societal values. Additionally, they transmit messages about who inhabits spaces in society. Changes in curriculum have been shown to impact children's worldviews and later life outcomes (Cantoni et al., 2017; Riley, 2022; Arold, 2022). Despite the importance of what we teach our children, however, we know little about content of educational curriculum being used in public schools across the United States. This knowledge gap is even more pronounced regarding the curricula in non-public settings, such as religious private schools or home schools. In light of the increasing diversion of tax dollars to these options, either through vouchers or other school choice programs, understanding the educational content to which children are exposed in these alternative settings becomes ever more critical.

In this study, we examine content in textbooks used in public and conservative Christian private schools in the U.S. over the last several decades. We examine three main sets of content: First, we analyze *what* topics and values are being included. Second, we examine *who* are highlighted in these lessons. Third, we explore *how* different groups of people, topics, and values are discussed. We also advance computational social science methods by introducing new applications of natural language processing and computer vision tools.

Specifically, we examine the content of state-adopted textbooks from Texas and California and the content of textbooks used widely in Christian private schools and by home-schooling families. We apply and improve upon computational content analysis methods to measure what topics, character values, and groups of people from different race, gender, and sexuality backgrounds are present in educational materials and if present, how they are portrayed, and how this differs across educational settings over time.

To examine *what* is being communicated, we employ a large language model to summarize the topics and values present. We find strong similarities across public and religious

private curricula. We see a similar proportion of pages dedicated to character values such as responsibility, family, and perseverance. We also see a similar amount of content focused on topics such as plants, animals, geography, history, and government. In contrast to public school textbooks, religious textbooks, including those for science, place a significantly greater emphasis on faith and religious values. We also see that religious textbooks are more likely to discuss creationism than climate change.

When examining *who* is represented, we analyze the faces in images and identity-specific words in the text. From the images, we first use a face detection model described in Szasz et al. (2022) to identify the faces of all pictured characters (both in photographs as well as in illustrations). We then classify skin color by first segmenting the facial area of interest and then calculating a continuous measure of its darkness or lightness based on the color of each pixel in the segmented skin. Separately, we predict the putative race, gender, and age of each face using a feature classification model described in Szasz et al. (2022). From the text, we examine which famous individuals are highlighted and then examine the presence of different race and gender groups among famous people. Additionally, we assess text-based gender representation by examining the proportion of gender-specific words used, including gendered pronouns, terms, famous individual names, and other mentioned character names. Within religious private school textbooks, we see that pictured faces are depicted with lighter skin colors on average. We see similar presence of females and males in Texas and California public school textbook images and text. However, in religious private school textbooks, we see less female representation in both the images and text.

We also see very little LGBTQIA+ representation. Indeed, it is only in the California collection that we see any mention of famous individuals who are identified as LGBTQIA+: Harvey Milk and Chaz Bono. They each appear in a 3rd grade Social Studies textbook.

To understand *how* different identities are portrayed, we use word embeddings to measure the context surrounding different groups of people. In particular we utilize Word

Embedding Association Tests (WEAT) to gauge the proximity of these groups to various domains, including politics and family life. To facilitate this analysis, we introduce a novel placebo test to evaluate whether the word embedding associations we find are meaningfully large. Additionally, we apply semantic role labeling to discern and characterize the sentiment of actions attributed to different groups of people.

We again see similar patterns across the three collections. They each are more likely to describe females in connection with family and the home compared to males, and more likely to describe males in connection to politics and tools compared to females. The top words more associated with females and least associated with males predominantly revolve around appearance, family roles, or household duties such as sewing or groceries. In contrast, the top words associated with males and least associated with females revolve around military/fighting, leadership, and horses. In comparison to males, words and actions associated with females are more positive, but less active and less powerful.

When the two LGBTQIA+ individuals are mentioned in the California textbook, they are described as fighting for LGBTQIA+ rights, but there is little other accompanying discussion. However, in religious textbooks, LGBTQIA+ rights are presented as a threat to religious rights.

Our study contributes to a vast literature focused on the content analysis of children's literature and educational materials (Wade, 1993; Allen, Allen and Sigler, 1993; Clark, 2007; Braden and Rodriguez, 2016; Bickford and Knoechel, 2017; Crisp, Gardner and Almeida, 2018; Koss, 2015; Koss, Johnson and Martinez, 2018; Koss and Paciga, 2020; Adukia et al., 2023). Most closely related to our approach is a study by Lucy et al. (2020) which uses natural language processing tools to examine the text (but not the images) of a small sample of 15 Texas high school textbooks used between 2015 and 2017. They find that Latine people are rarely discussed, the most common famous figures are nearly all White men, and that women tend to be discussed in the contexts of work and the home. Another related paper

by Stevenson and Zlotnik (2018) focused on college level introductory economics textbooks finds that fictionalized women more likely to be involved in food, fashion, or household tasks. In a qualitative analysis from (Scaramanga and Reiss, 2018) on curriculum published by Accelerated Christian Education – a publisher in our sample of religious private school textbooks – revealed that churches and schools are often depicted as racially segregated in their materials. We expand on this work by applying computer-driven content analysis to examine both the images and text found in over two hundred K-12 textbooks used in a variety of contexts between 1980 to the present.

The computational tools that we develop and apply throughout our analysis provide a complementary understanding of educational materials and expand the scope of previous studies. Our work surpasses the number of books previously systematically analyzed and is able to both replicate qualitatively similar results found by previous studies in addition to documenting new findings. The larger sample and greater breadth of tools allows us to look at a broader set of questions such as how different skin colors are represented, which is an important dimension of representation not captured by the analysis of race alone (Keith and Monroe, 2016).

Our study also incorporates innovative uses of computer vision and natural language processing tools, thereby contributing to the computational social science literature in a manner similar to Caliskan, Bryson and Narayanan (2017); Garg et al. (2018); Kozlowski, Taddy and Evans (2019). This methodological advancement has significant implications for the broader field of computational social science.

The paper proceeds follows: Section II provides additional context to religious school choice in the United States, Section III describes the textbooks in our sample, Section IV discusses our text and image processing tools, Section V presents our results, and Section VI concludes.

II Parental Educational Investment Decisions and School Choice

The process of education includes the transmission of both knowledge and values. Educational content and instruction have been shown shape worldviews explicitly through the discussion of certain topics and values, and implicitly through who is represented and not represented (Cantoni et al., 2017; Riley, 2022; Arold, 2022). Consequently, parents' decisions about their children's education are driven by the values, cultural norms, and particular sets of knowledge they want to instill in their children.

There are two main ways that values are transmitted to children, the first is horizontal transmission from society and peers, typically provided by formal institutions such as public schools. The second is vertical transmission of values and knowledge from family and parents. If parental values differ enough from those of mainstream society, they may choose to remove their children from public school and either homeschool them or place them in a private school whose values are more aligned.

In today's day and age, we see a increase in conflict around what what to include – or exclude – from public school curricula. Recent years have witnessed significant shifts in public school content standards. On one hand, some areas have embraced greater inclusion; a notable example is California's 2011 FAIR Education Act, which mandates instruction in history and social science to include a variety of underrepresented ethnic and cultural populations in the teaching of California and United States history including both men and women, Native Americans, Mexican Americans, LGBTQIA+ individuals, and persons with disabilities. On the other hand, there has been a trend towards more restrictive content policies in other areas, often labeled as "educational gag orders." Between January and September 2021, for instance, legislatures in 24 U.S. states introduced 54 bills aimed at barring specific "prohibited" or "divisive" concepts from public K-12 and university curricula. These restrictions predominantly focus on limiting discussions around race, gender, and American history, as detailed in (Friedman and Tager, 2022). Notably, such restrictions

are not imposed on private schools and homeschooling families. If parents disagree with what public schools are teaching, they have the option to remove their children from public school and find an alternative education setting. In tandem with the increasing discussion around what we allow to be taught in public schools, we have witnessed a notable surge in the adoption and expansion of school choice initiatives, including voucher programs, charter schools, and homeschooling (Figure 1).

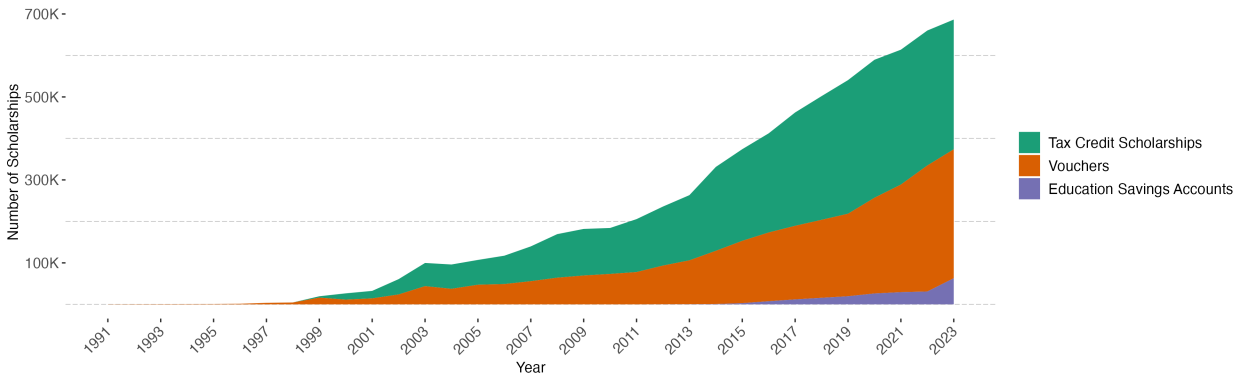
Despite these increases in school choice options, curricular requirements in the U.S. are minimal for private schools and homeschooling families. Most states simply require private school curricula to be "similar" to standards prescribed for public schools and/or require that certain subjects be taught. There are 18 states that either have no requirements or only require that students be taught the state constitution and state history. While states usually have similar curricular requirements on paper for homeschoolers as they do for private schools, there is rarely any enforcement for homeschooling families. This lack of regulation is especially surprising given how much political attention is given to what can and cannot be included in public school curricula.

The evidence on the effect of school choice programs within the United States on children's academic outcomes is mixed. Deming et al. (2014) find that school choice across public high schools benefits applicants' academic outcomes only when they gain access to higher-quality schools. Research on the effectiveness of school voucher programs, which allow students to direct public funds toward tuition payments at private schools of their choice, is also mixed. An analysis of the Milwaukee Parental Choice Program, one of the oldest voucher program in the United States, finds modest test score gains for voucher users but Abdulkadiroglu, Pathak and Walters (2015) find that parents in more rural settings such as Louisiana are using these programs to move their children to lower quality schools as measured by students' academic performance.

This suggests that parents may be leaving the public school system for more than just

better academic outcomes for their children; they may be looking for an educational setting that better aligns with their values and beliefs. We see evidence of this in survey data. In a survey of parents with children in private schools, 80% listed character or values instruction as a top desired school quality (Catt and Rhinesmith, 2017a). A desire for religious instruction is reflected in private enrollment patterns; most students educated outside the public school system are educated in a religious school environment (Appendix Figure BI). Indeed, In 2017, 76% of the 5.8 million private school students were enrolled in a religious school, with 49% enrolled in a conservative Christian school (PSS, 2015). Among those parents who chose to educate their children at home, or in informal settings, 75% reported a desire to provide moral instruction as a top reason they chose to homeschool (National Center for Education Statistics, 2022). Pre-COVID estimates of the number of homeschoolers who are religious, or for whom religion is a primary reason for homeschooling, range from over half to 90% (Kunzman, 2010; Bartholet, 2020). While academics still clearly play a role in parent's school choice decisions, they are not the only and sometime not even the most important factor.

FIGURE 1
Number of School Choice Scholarships Awarded to Students Over Time



Given the growing number of tax dollars being diverted from public schools to religious private schools and homeschooling families with little regulation on what they are taught and limited evidence that this diversion benefits students academically, it is important for

to us to understand what children are being taught in these private educational settings and how it compares to what children are taught public schools.

III Data

We digitize a novel data set of Reading, Science, and Social Studies textbooks designed for third and fourth grade. Our sample draws from three distinct sources, detailed in the subsequent subsections. As depicted in Table 1, our data set comprises over 75,000 pages originating from 261 different textbooks, amounting to a corpus of over 16 million words. To our knowledge, this is the largest data set of textbooks collected and digitized by an order of magnitude. Previous studies have typically digitized only a small number of textbooks, limiting their analyses to cross-sectional examinations. Our panel of textbooks affords a unique opportunity for longitudinal analysis across sets of curriculum and over multiple decades, setting our study apart in both scope and depth within the academic literature.

TABLE 1
Unique Textbooks in Our Sample

	Reading		**Social Studies**		**Science**	
	Books	*Pages*	*Books*	*Pages*	*Books*	*Pages*
Religious	55	10,677	13	4,234	17	4,703
California	22	9,268	22	8,948	36	10,774
Texas	60	13,142	22	11,093	14	4,001

III.A *Religious (Conservative Christian) Textbooks*

The religious textbooks in our sample include those published by three ideologically Christian textbook companies: Abeka, Bob Jones University Press (BJU), and Accelerated Christian Education (ACE). These textbooks are commonly used by homeschooling families and adopted by over a third of non-Catholic Christian schools receiving tax-payer funding (Klein, 2017). We focus on textbooks published by these three Christian publishers both because of their wide adoption, but also because there are many taxpayer dollars funding

schools which use these curricula. Notably, in Indiana, about 4,240 students received over $16 million in scholarships to attend schools that use Abeka or Bob Jones curriculum, according to 2016-2017 figures from the Indiana Department of Education (Klein, 2017). We digitize each textbook published by these publishers for the grades and subjects in our sample starting with books published in 1980 up until 2022. Note that most of curriculum published by ACE consists of workbooks called PACEs instead of textbooks, so we only including books from their Reading curriculum in our analysis.

III.B *Texas State-Adopted Textbooks*

We include Texas textbooks as a comparator group because historically, Texas has driven much of the textbook procurement market in the U.S. due to their large share of the educational publishing market. Texas is often depicted as being on the more conservative end of the political spectrum in the United States.

The state of Texas compiles a list of state adopted textbooks for all K-12 grades which is updated yearly. Before 2011, Texas public schools were only allowed to use funds to purchase books from this list. In 2011, this law changed to allow schools to use funds to purchase any textbooks, but most schools continued to purchase textbooks off the state adopted list. We digitized a convenience sample of approximately 50% of all textbooks on the yearly state adopted textbook list for textbooks in use between 1985-2010. Beginning in 2011, we digitized the top three textbooks purchased by school districts each year for each relevant grade and subject (reliable textbook purchasing data was not collected before 2011).

III.C *California State-Adopted Textbooks*

California also has a large influence on the textbook publishing market due to the size of its student population, making it another useful comparator group. It is typically associated with more progressive policies and is often posed as the political or substantive converse of Texas, which is typically considered to be more conservative in popular discourse.

The state of California compiles a list of state adopted textbooks for grades K-8 which

is updated every couple years for each subject[1]. We digitize the universe of books in the relevant grades and subjects which were adopted in California for all years in which the list of adopted materials are available. This includes adopted textbooks beginning in 1999 for social studies, 2000 for science, and 2002 for reading.

IV Methods

The majority of previous studies examining the content of textbooks have been conducted using manual content analysis which involves recording the presence of certain words or themes by hand. Because this process is time-consuming and therefore costly, resource constraints have limited the scope of such work. The need to analyze large corpora of educational materials led to our exploration of computer vision and natural language processing tools, which allow for a more cost-effective and systematic approach to content analysis. More recent studies which apply computational content analysis methods to examine the content of textbooks (Lucy et al., 2020) have only examined the text and not the images found in the textbooks. Our methods advance both image and text analysis in the computational social science literature along with recent papers by Adukia et al. (2023) which examines children's books, Ash et al. (2021) which examines newspapers, and Yanagizawa-Drott and Voth (2023) which examines yearbooks.

We focus on three main categories of content analysis: What, Who, and How.

***What* topics and values are being communicated.** To measure *what* these textbooks cover, we use a large language model to identify topics and values on a given page. We also use token counts to measure whether and how often specific words are mentioned.

***Who* is depicted.** We will measure *who* is highlighted in the images and text. To examine the photographs and illustrations in the textbooks, we use face detection to identify each pictured individual. We then classify skin color and predict features to measure how often people of different putative races, genders, and skin colors are represented in textbook images.

[1]Note that during the economic crisis of 2007, California did not make any changes to their state adopted textbook lists to save districts the expense of purchasing new textbooks.

To measure who is highlighted in the the text of the textbooks, we use token counts and named entity recognition to measure how often gendered words (including names, pronouns, and terms) are mentioned as well as how often famous people of different races, genders, and nationalities are represented.

***How* content is portrayed.** To measure *how* different people, topics, and values are discussed, we use word embeddings to identify words found in similar contexts to them and semantic role labeling to identify the actions that people of different genders and races are described as taking. We then analyze the sentiment associated with this content.

We describe the image and text analysis methods that we use for this exercise below and illustrate our overall methods pipeline in Figure 2.

IV.A Image Analysis

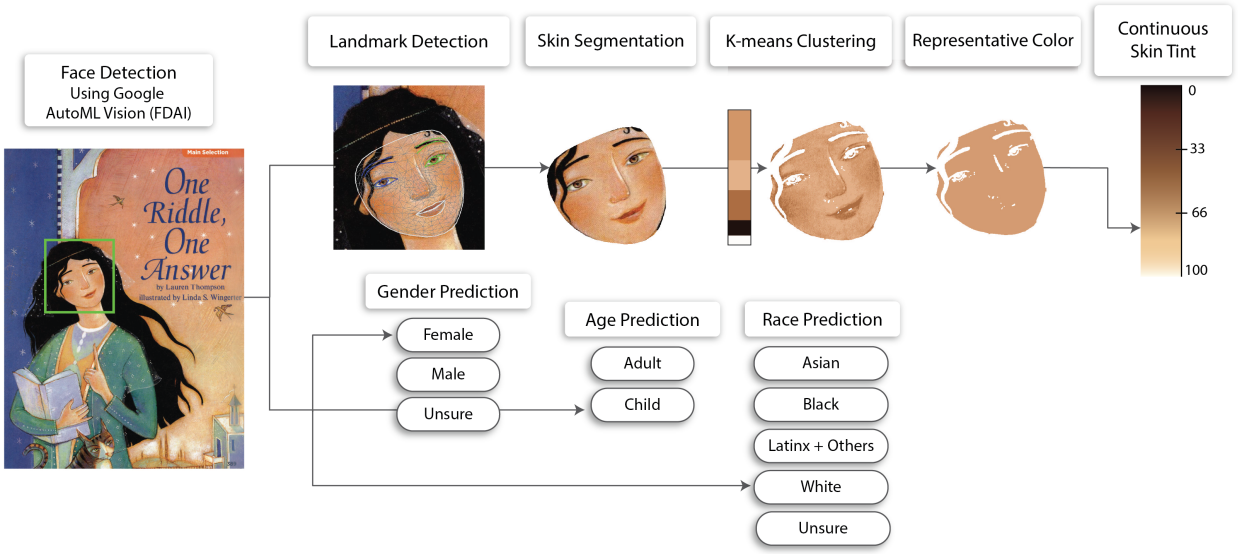
The process that we use to convert images into data is illustrated in Figure 2a. We describe in this section how we detect faces, classify skin color, and predict features such as putative race and gender.

IV.A.1 *Image Analysis: Face Detection*

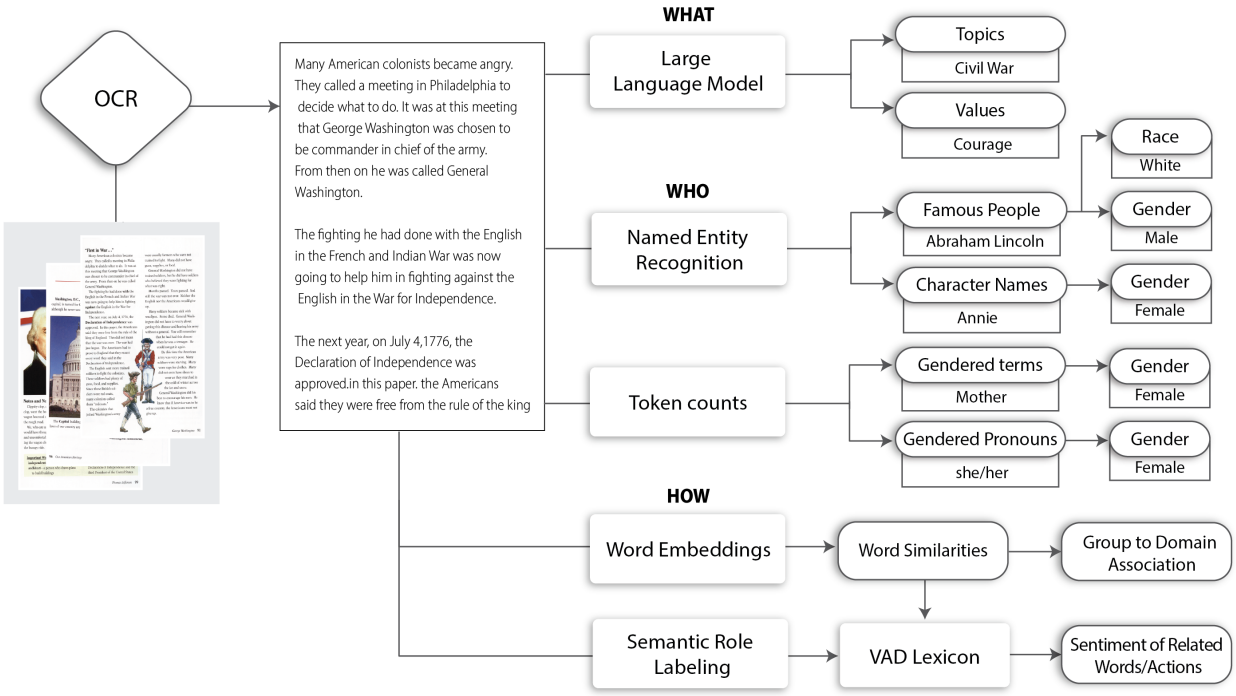
Our first step when measuring representation in images is to run face detection on every digitized page in our sample to identify all depicted faces. However, in the process of applying off-the-shelf face detection models, we discovered a significant drawback; while off-the-shelf methods are an excellent way to detect faces in photographs, they perform poorly when applied to illustrations. Since many of the images found in content targeted towards children contain illustrations, this presented a problem. To mitigate this issue, we trained a model which can better detect faces in both photographs and illustrations as described in Szasz et al. (2022). This model increased the number of detected faces by 250% in the sample of books. We apply this model to detect faces in the textbooks in our sample.

FIGURE 2
Methods Pipelines: Image and Text Analysis

(a) Image Analysis Pipeline



(b) Text Analysis Pipeline



Note: This figure illustrates the steps in our image analysis (top) and text analysis (bottom) pipelines beginning with the raw scan of each page in our sample of books.

IV.A.2 *Image Analysis: Skin Color*

We extract and classify skin color using a similar method as Adukia et al. (2023). We begin by identifying the facial area of interest in the pictured face by "segmenting" the skin. To do this, we use Google's MediaPipe Face Landmarker to detect facial landmarks and then to isolate the parts of the face which contain skin from other facial features. This helps us to eliminate hair, clothing, and other accessories. Then, we extract the dominant colors in the segmented skin by applying k-means clustering to group each pixel in the skin segmentation in RGB space and take the center of each cluster. Finally, we collapse the entire skin segmentation into a single "representative color" by taking a weighted average of the dominant skin colors (where the weights are determined by the size of the cluster each dominant color was drawn from). To construct a continuous measure of how light or dark the skin color of a detected face is, we convert the center of each cluster in RGB space and convert its color into the perceptually linear CIE L*a*b* color space. We take a weighted average the L* value of each center which ranges from 0 (darkest) to 100 (lightest). This results in a continuous measure of perceptual tint that serves as our measure of skin color. This process is illustrated in Figure 2a.

IV.A.3 *Image Analysis: Feature Classification of Putative Race, Gender*

We then wish to classify different features of pictured characters, specifically putative race and gender. We use an AutoML feature classification model that we trained using a manually curated data set which can more accurately predict features of illustrated faces such as the race, age, and gender (Szasz et al., 2022). We use this improved feature classification model for our analysis.

IV.B Text Analysis

The process that we use to convert images into data is illustrated in Figure 2b.

IV.B.1 *Text Analysis: Large Language Models*

Large language models (LLMs) are deep learning models that can recognize, summarize, translate, predict, and generate text. LLMs acquire these abilities by using massive amounts of data and computational resources during their training to achieve general-purpose language understanding and generation.

For our analysis we use OpenAI's GPT-4 model (from now on referred as GPT), which is an autoregressive pre-trained LLM that applies deep learning algorithms to generate what would be a human-like response to user input. This is done through the use of a transformer (i.e., neural networks) architecture, composed of encoding and decoding models. The encoding part takes the input given by the user, extracts its relevant features, "understands it" through its context, and creates a high-level representation of it. This representation is then passed as an input to the decoding model, where the final response is generated using information learned in the pretraining process.

In particular, we ask GPT to analyze single pages of textbooks and identify topics and values present in them. OpenAI's API allows us to define a specific LLM, a system role, and a *temperature* for each of our prompts. In this paper we use the `gpt-4-1106-preview` model, which allows a 16,000 token limit for total input and response. The parameter *temperature* defines the freedom or creativity we allow the model to have for its responses, and is set up to its default median value, 1.[2] Finally, we assign a role to the system to set its behaviour. In our analysis, we instruct the model to take on the role of "a knowledgeable and unbiased judge."

For each textbook page, we ask the model to identify the topics and intrinsic values conveyed in the page, a sentiment analysis rate, and flags for religious, evolution, climate

[2]The optimal value for the temperature parameter will usually depend on the type of task we ask GPT to do, and to the best of our knowledge there isn't a benchmark or standard for this. While higher temperature values give the model more room to hallucinate (create fake or made-up responses), lower ones make responses more deterministic, potentially limiting its flexibility to provide reasonable responses.

change, sexuality, racism, sexism, obedience, and stereotype content.

IV.B.2 *Text Analysis: Token Counts*

Token counts are one of the simplest natural language processing tools which allow us to count the number of times specific groups or topics are mentioned. We can use these counts to see how often words related to females (such as "she", "queen", "sister", etc) are mentioned relative to words related to males. These words include pronouns and other gendered terms. We can also use these counts to examine whether changes in science education content standards surrounding the theory of evolution translate into more frequent mentions of the topic.

IV.B.3 *Text Analysis: Named Entity Recognition*

Named Entity Recognition (NER) identifies "named entities" (proper nouns predicted to be names). And counts the number of times they appear in the source text. We combine these identified names with a pre-existing data set containing information from over 70,000 Wikipedia biographies (Yu et al., 2016) to identify famous people mentioned in the text such as Martin Luther King Jr. or Amelia Earhart. Using these data, we can examine the number of famous people mentioned by race, gender, occupation, birthplace, and sexual orientation.

We use the remaining named entities that were not matched to a famous person to measure the number of female and male names mentioned in the text. To do this, we extract the first name of each remaining named entity and estimate the probability that the name is female (or male) using data on the frequency of names by gender in the U.S. population from the Social Security Administration.

IV.B.4 *Text Analysis: Semantic Role Labeling*

Semantic role labeling (SRL) models identify the predicate-argument structure of a sentence. In other words, these models label the agent (active subject), main verb (predicate/action), and the patient (object or passive subject) in a sentence and the various

elements that accompany it (e.g. "who" (agent) "did" (verb) "what" to "whom" (patient)).[3] For example, in the sentence "In the North, President Abraham Lincoln asked men to fight to preserve the union" the *agent* in the sentence is "President Abraham Lincoln," the *action* in the sentence is "ask," and the *patient* in the sentence is the statement, "to fight to preserve the union."

We use SRL to help us characterize the actions taken by agents of different races and genders in our text. To do this, we first apply SRL to all the sentences in a given textbook to identify the agent and the action. Then we classify the agents' race and gender as well as the sentiment of each action.

To classify the race and gender of each agent, we start by using named entity recognition to extract any names in the text labeled as the agent. If any names are detected, we classify their associated gender and race (if possible) following the process in Section IV.B.3. We identify if the name belongs to a famous person by matching it with data on 70,000 Wikipedia biographies. If the name belongs to a famous person, we can classify the agent by the famous person's known race and gender. If the name does not belong to a famous person, we cannot classify the race of the person, but we can predict the gender of the agent based off their first name. If no names are found, we apply part-of-speech tagging to identify if the agent of the sentence contains a gendered pronoun or term. We follow the same process as described in Section IV.B.2 to classify gender. If the agent of the sentence contains a gendered pronoun (e.g. she), we use the gender of the pronoun to classify the gender of the agent. Similarly, if the agent of the sentence contains a gendered noun/term (such as brother or queen), we use it to classify the gender of the agent.

IV.B.5 *Text Analysis: Sentiment Analysis using VAD*

To characterize the sentiment of the actions taken by agents of different races and genders, we measure their valence, arousal, and dominance (VAD). Associated scores provide

[3]We implement Semantic Role Labeling using AllenNLP, a library for applying deep learning methods to NLP research (Gardner et al., 2018).

measures of different attributes of each action. Valence measures range from happiness to unhappiness, or positive to negative. Arousal measures indicate a level of "affective activation,' or active/excitement to passive/sleep. Dominance measures reflect the level of control of the emotional state: dominant/powerful to submissive/weak. We use the NRC VAD Lexicon which includes a list of more than 20,000 English words and their valence, arousal, and dominance scores (Mohammad, 2018). Scores range from zero to one. For example, as shown in Table 2, the words "teamwork" and "leader" both have relatively positive valence, but "teamwork" has slightly more positive association to it. "Leader" has slightly more active associations and meaningfully more powerful associations compared to "teamwork."

TABLE 2
Example Words and Their VAD Scores

	Valence *(Positive/Negative)*	Arousal *(Active/Passive)*	Dominance *(Powerful/Weak)*
Teamwork	0.918	0.490	0.773
Leader	0.833	0.582	0.926
Difference in Scores	**0.085**	**-0.092**	**-0.153**
Sing	0.858	0.606	0.420
Speak	0.765	0.472	0.500
Difference in Scores	**0.093**	**0.134**	**-0.08**

IV.B.6 *Text Analysis: Word Embeddings*

Word embeddings are vector representations of words which encode semantic meaning by incorporating information from the nearest neighbors (context) of that word. As a result, word embeddings can be used to examine not only whether a topic or group is present in our source text, but also the context in which it is discussed. Specifically, the closer two words are positioned together in high-dimensional space (as determined by the cosine similarity of two word vectors), the more often they are used within the same context.[4]

[4]We utilize word2vec from Python's Gensim library to estimate word embeddings. Our word2vec implementation employs the Skip-gram model architecture for training, which uses a given word to predict context words (words that appear within a certain window of the current word) in a sentence.

To examine the context in which females and males are portrayed, we apply a textual analysis variant of the Implicit Association Test (IAT) called the Word Embedding Association Test (WEAT) introduced by Caliskan, Bryson and Narayanan (2017). This word embeddings analogue of the IAT involves taking the estimated vectors for words related to a given gender group (e.g., "she" and "queen" belong to the "female" group) and calculating their average distance to words related to a given domain (e.g., "baby" and "marriage" belong to the "family" domain).[5] We then estimate a domain-specific parameter of "gender-skew" by comparing the average distance to a domain between males and females. To determine where differences in domain association between genders are statistically significant, we use a novel application of placebo tests.

We also examine the Valence, Arousal, and Dominance scores of words most closely associated with females vs. males using the same VAD lexicon we use in our Semantic Role Labeling analysis (Mohammad, 2018) as described in Section IV.B.5. Following Charlesworth, Caliskan and Banaji (2022), we do this by finding the average VAD score of words in the VAD lexicon that are mostly closely associated with female words relative to male words.

V Results

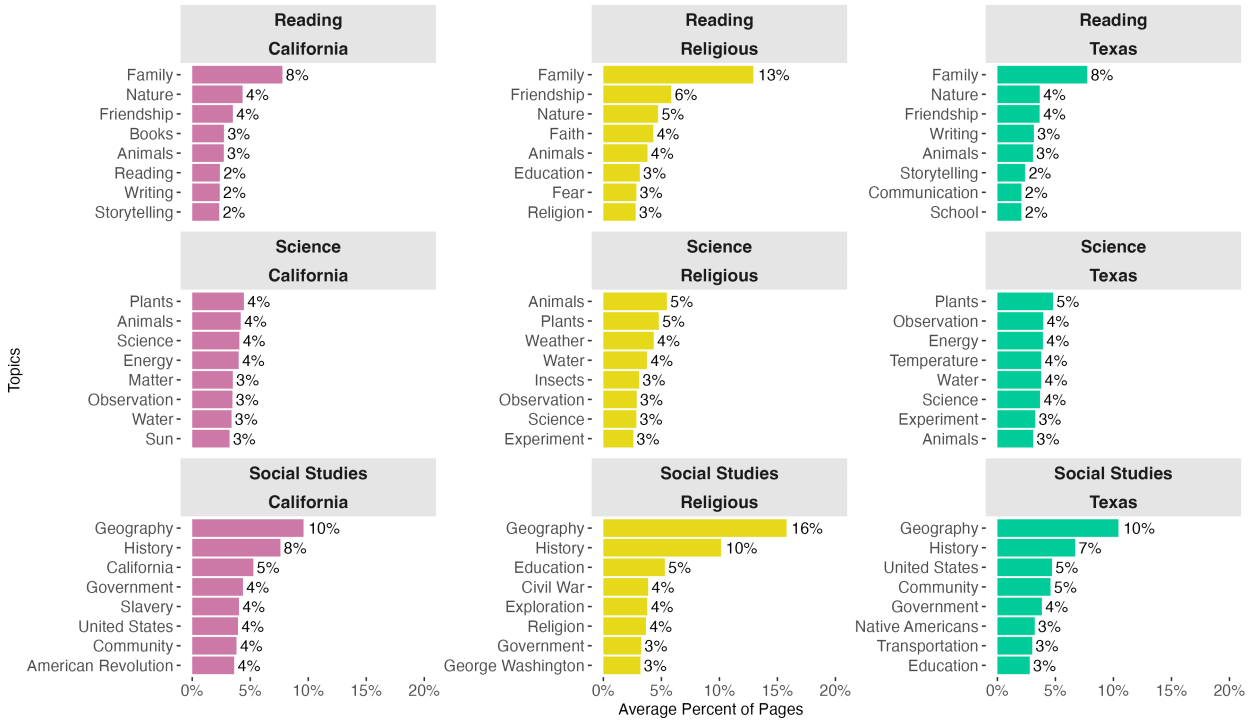
V.A What: Textbook Topics and Values

Parents choose schooling options, in part, because they want their children to be exposed (or not) to certain knowledge and values. In this section, we document what topics and values are being imparted to children, either implicitly or explicitly.

We first examine what topics are shared with children in each subject. We see that similar topics are discussed across collections within subjects with some notable exceptions unique to each collection. In Reading textbooks, the top three topics identified were the same for each collection: family, nature, and friendship. The religious collection is more likely to include narratives related to family than the collections of public school textbooks

[5]We include word lists for each group and domain in the appendix.

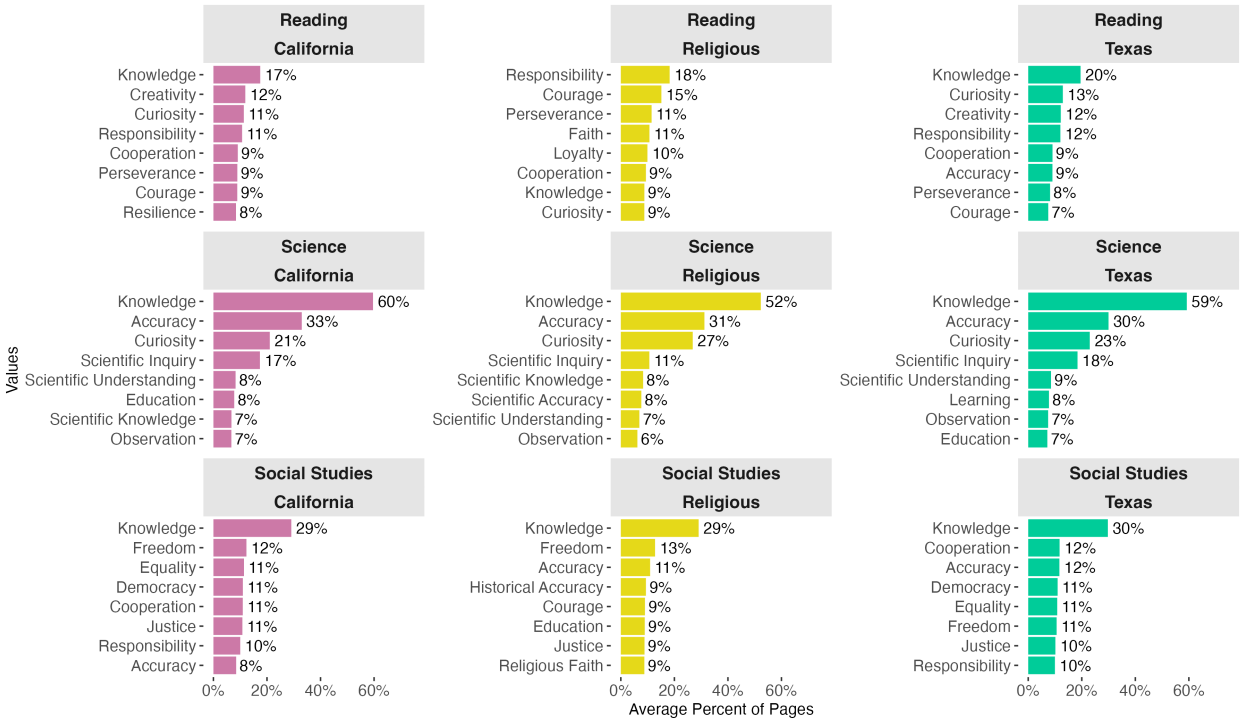
FIGURE 3
Top 8 Topics by Collection



(13% vs. 18%). The main important difference is that in the religious collection, "faith" and "religion" are among the most common topics identified. In Science textbooks, plants, animals, observation, and water are among the most common topics across all three collections and in Social Studies textbooks, geography, history, and government are among the most common topics for all three collections (Figure 3). Note that once again, we see Religion as a common topic in Social Studies textbooks for our collection of Religious textbooks.

The majority of parents who exercise school choice report a desire for "morals," "values," or "character" education as one of the top most important qualities when choosing a school (Catt and Rhinesmith, 2017b). We accordingly examine the values that are being espoused on the pages of the textbooks and find a similar pattern to the topics found of textbook pages. Specifically, the values we see are similar across collections with the exception of Faith and Religious Faith showing up as top values in the religious collection

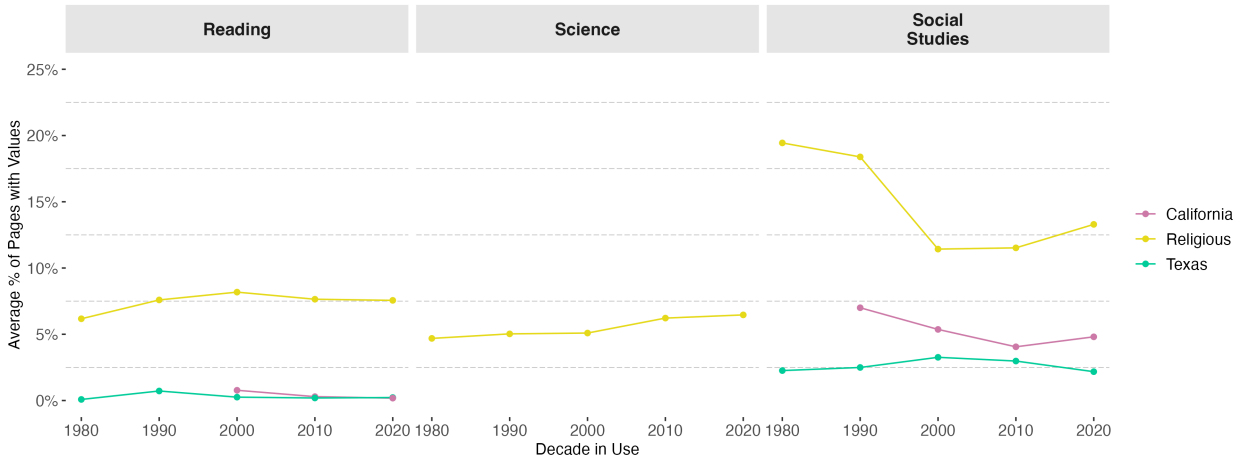
FIGURE 4
Top 8 Values by Collection



(Figure 4). We explore this further by combining values related to religion such as Religious Faith, Religious Freedom, Religious Beliefs, Religious Diversity, Religiosity, Religious Tolerance, Religious Devotion, etc. Using this group of values we find that religious values are present on 5%-20% percent of pages on average in a given subject and year in our collection of religious textbooks as compared to 0%-8% of pages in our collections of public school textbooks. Notably, in public school textbooks, religious values only appear in textbooks covering Reading and Social Science, but they are present on approximately 5% of pages on average from Science textbooks in the religious collection (Figure 5).

In contrast to the frequency of religious topics, controversial topics like climate change, creationism, evolution, racism, and LGBTQIA+ issues appear on less than 2% of the pages, as detailed in While all three collections mention evolution and climate change on a small percentage of pages in Science textbooks on average, creationism is exclusively mentioned in

FIGURE 5
Average Percent of Pages Containing Religious Values by Subject and Decade in Use



our religious collection. In fact, within this collection, creationism is mentioned more than climate change, occupying a larger percentage of pages on average.

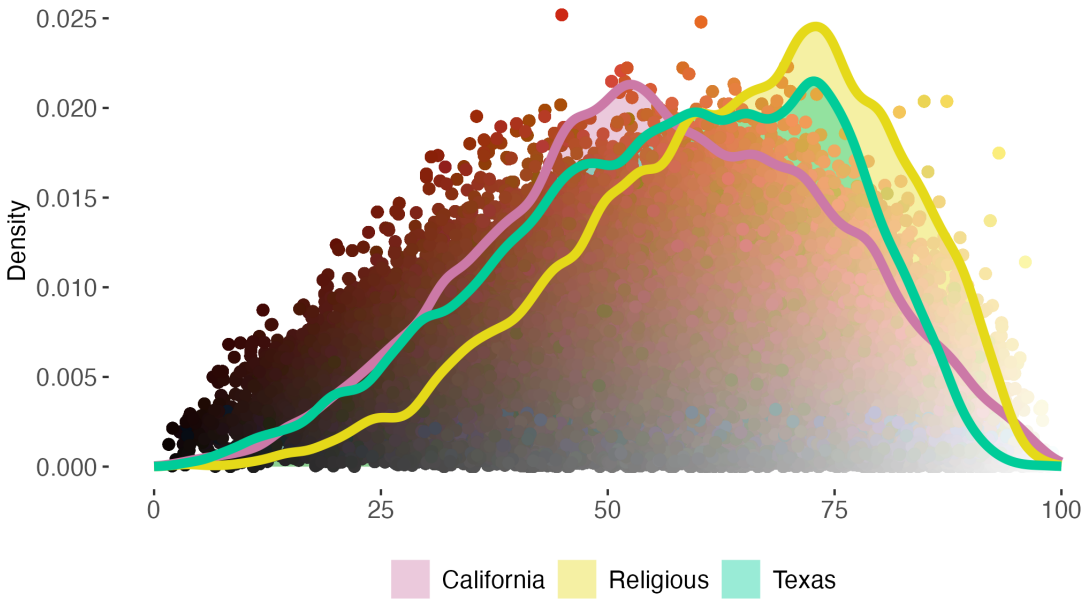
Furthermore, Figure BII highlights a notable divergence in the relative percentage of pages which mention slavery versus the Civil War. Specifically, Social Studies textbooks from our California collection allocate more pages to discussing slavery (approximately 5%) than the Civil War (around 1.5%). In contrast, in our collection of Religious Social Studies textbooks, the Civil War garners more attention (5% of pages) compared to slavery (3% of pages). Meanwhile, Social Science textbooks in the Texas collection exhibit a more balanced approach, dedicating roughly equal coverage to both topics.

V.B Who: Skin Color, Race, Gender, & Sexual Orientation

Skin color. In Figure 6, we show the distribution of detected faces in our textbook by skin tint (e.g. how light or dark their skin color is on a scale from 0 to 100). We plot these distributions over a scatterplot showing the skin color of every face in our sample. The x-axis for both the density and scatter plot is skint tint, but the y-axis for the scatterplot is the vibrancy of each skin color (where 0 means the face is monochromatic).

Overall, we see that the the distribution of faces in textbooks from our religious

FIGURE 6
Distribution of Skin Tint by Textbook Collection



collection are more likely to portray pictured characters with lighter skin than the Texas or California books. This is true even when conditioning on race; In the religious collection the skin color of faces is lighter on average than faces in California and Texas for all races/ethnicities except for Latine (Table 3). However, we do not see a similar clear pattern when comparing Texas and California. In California, for example, characters classified as being Asian or more likely to have lighter skin than characters classified as Asian in Texas; whereas this pattern is flipped for characters classified as being Black: they are shown with lighter skin in Texas compared to California.

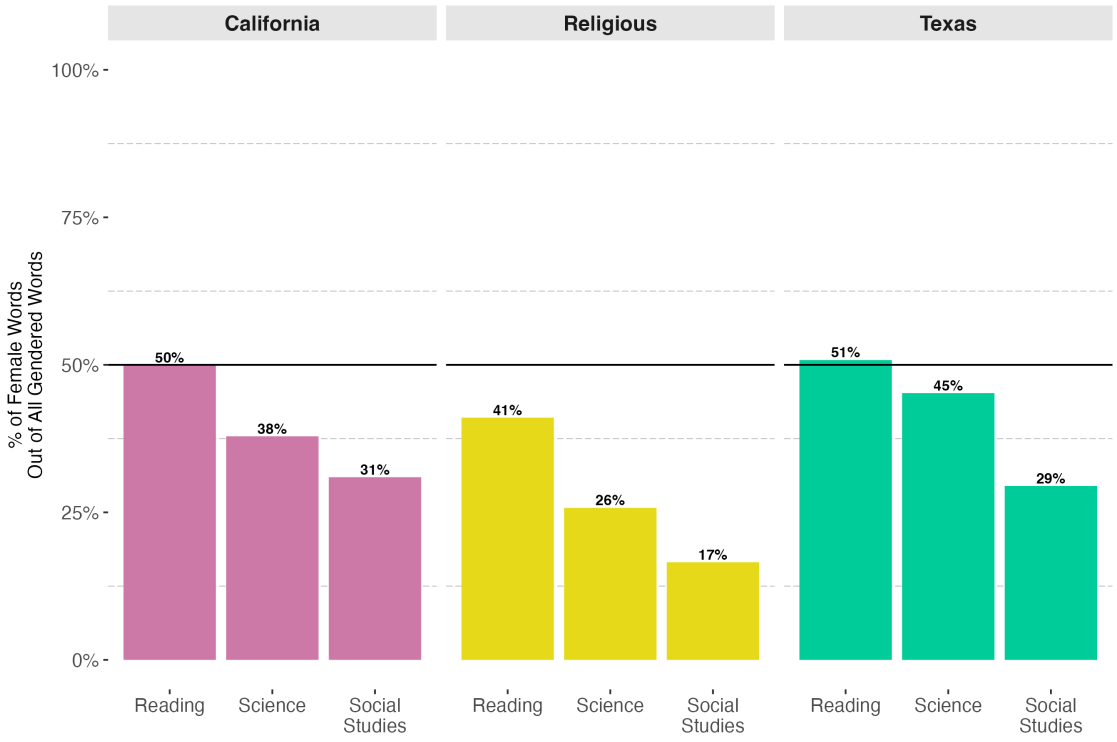
Gender. We find that females are less likely to be represented in both the text and the images of textbooks in the religious collection than in our collections of public school textbooks (Appendix Table AI). However, across all collections females are under-represented in relative to their share of the population. The disparity is much larger in the text than it is in the images. While females are depicted less often than males, females representation is close to parity in the images (Appendix Figure BIII); this is not the case for female representation

TABLE 3
Skin Color by Race

	Dependent variable:				
	Skin Tint				
	Asian	Black	Latine	White	Other/Unsure
	(1)	(2)	(3)	(4)	(5)
Religious	9.606***	4.970***	3.926	4.015**	7.060***
	(2.676)	(1.626)	(2.545)	(1.687)	(2.493)
California	7.699***	−3.465**	2.177	−0.444	5.532**
	(2.617)	(1.541)	(2.296)	(1.776)	(2.561)
Grade 5	−7.929***	−2.914**	−2.147	−5.248***	−6.946***
	(2.300)	(1.176)	(1.442)	(1.292)	(2.430)
Science	−8.043***	−8.629***	−3.613	−9.312***	−12.856***
	(2.945)	(2.078)	(2.378)	(2.033)	(3.411)
Social Studies	−3.519	−2.484*	−0.947	−2.088	−4.441**
	(2.241)	(1.337)	(1.557)	(1.482)	(2.122)
Year In Use	−0.159*	−0.148**	0.160**	−0.080	−0.074
	(0.094)	(0.060)	(0.069)	(0.057)	(0.088)
Constant	63.277***	59.275***	42.360***	68.369***	60.837***
	(3.283)	(1.745)	(2.851)	(1.842)	(2.847)
Observations	12,574	98,923	13,084	182,102	39,321
Adjusted R^2	0.161	0.068	0.033	0.089	0.109

in the text. The total proportion of female words over time ranges from approximately 20% in our collection of Religious textbooks to approximately 40% in our collections of public school textbooks, with these proportions being relatively stable over time (Appendix Figure BIV). Interestingly, when we disaggregate gender representation by subject, a distinct pattern emerges. Notably, in public school reading textbooks, females are mentioned as frequently as males (Figure 7).

FIGURE 7
Female Representation in Text By Subject



This pattern may be attributed to the greater flexibility publishers have in selecting the gender of characters for reading textbooks. Unlike history textbooks, where the narrative often revolves around historical events and figures from which women were historically marginalized—such as the consistent male dominance in the U.S. presidency—reading textbooks are not bound by these historical constraints. Consequently, discussions in reading materials can feature a more balanced gender representation. We don't see this same pattern in the religious collection of textbooks. As illustrated in Figure 7, females are consistently

underrepresented in the text across all subjects within the religious collection. The closest to gender parity is found in Reading textbooks, where female-related words constitute only 41% of all gendered words.

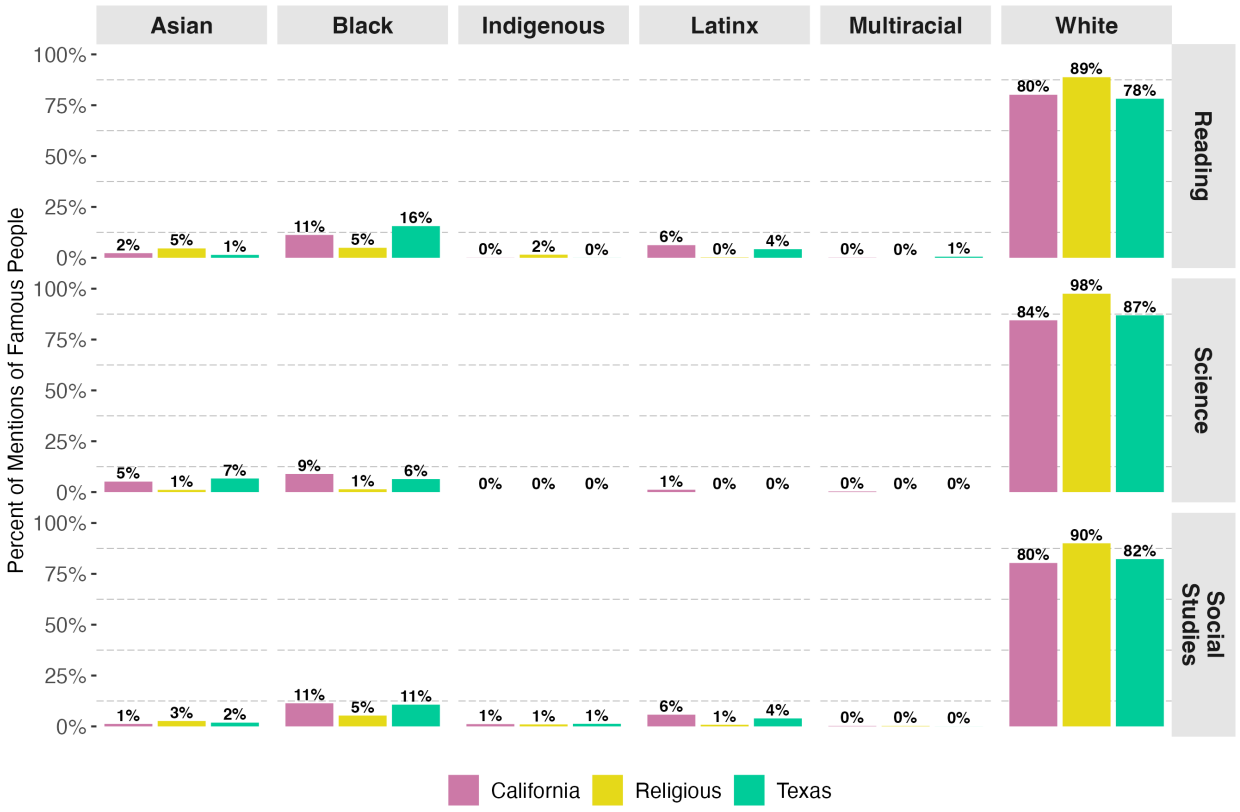
Race. A salient part of curricula are the famous individuals who are highlighted for children. History and science came from individuals, so in the conveyance of information, textbooks often include discussion of these individuals. They can serve as real-life characters in our histories, examples of leadership, and as role models to whom children may aspire to be similar. To measure how often famous individuals of different races/ethnicities are discussed, we apply Named Entity Recognition (NER) tools as discussed in Section IV. We find that the vast majority of famous individuals mentioned are White individuals (between 75%-98% across all collections and subjects as shown in Figure 8). This pattern is consistent over time with a slight decrease in famous White individuals represented in recent decades (Figure BVI).

One caveat of focusing on famous individuals only to measure the representation of race is that we are unable to capture the race of non-famous individuals in the text. However, we can look at how often both famous and non-famous individuals of different races are pictured in textbooks by examining the race classifications of detected faces in our sample. We find that in all collections across most subjects, most of faces depicted are White individuals. The exception to this can be found by examining Social Studies textbooks from Texas; here most of the faces are Black individuals (44% of faces) followed closely by White individuals (39% of faces) as shown in Appendix Figure BV.

LGBTQIA+ Individuals. There were only two books in our sample that we identified as discussing LGBTQIA+ topics. These books were identified using token counts of the words lesbian, gay, bisexual, transgender, queer, intersex, and asexual.

The first book is a 5th grade Social Studies textbook from our religious collection published in 2016. The relevant passage is taken from a section entitled "Gay Rights":

FIGURE 8
Race/Ethnicity of Faces by Collection and Subject



> "Another issue that troubled Americans was gay rights. During his time as president, Obama encouraged the growth of gay rights. He said that men should be able to marry men and that women should be able to marry women. When Obama first took office, many Americans disagreed with this view. By 2015, many agreed. Americans who disagreed faced problems. Many believed that God designed marriage to be between a man and a woman. They trusted that His plan benefited humanity. These Americans did not support gay marriage. They thought that it would harm society. For example, children in families where there is a same-sex marriage are denied having both a father and a mother. But in 2015, the Supreme Court ruled in favor of gay marriage. It had redefined marriage. Five justices voted in favor of gay marriage, but the other four justices voted against it and pointed out that this change threatened religious liberties. Some Christians who owned bakeries, flower shops, or photo studios were sued for not participating in a same-sex wedding. These businesses served all other customers, but refused to participate in same-sex weddings. Many of these business owners lost their court cases and were fined. This change in the United States raised other questions. Could religious schools continue to teach their views of marriage? Would schools and business owners come under more pressure to violate their beliefs on these issues? These questions remained unanswered."

The second book is a 3rd grade Social Studies textbook from our California collection that was published in 2019. The relevant passage is from a section entitled "Supporting Many Different Communities":

> "Many other Californians worked hard to make changes in their communities...
>
> ...Harvey Milk was a political leader and a gay activist in the 1970s. Gay men are attracted to other men. Milk was elected to a local government office in San Francisco in 1977. He was one of the first elected officials in history to tell the public that he was gay. Milk worked hard to protect gay and lesbian rights. Lesbians are women who are attracted to other women. Milk's work also made it easier for gay, lesbian, bisexual (those attracted to both men and women), and transgender groups to stand up for their rights.
>
> Chaz Bono is a transgender activist. The word gender refers to whether someone is male or female. Transgender people feel that their gender is the opposite of what it was at birth. Bono was born female, but always felt male. He changed his gender and today lives as a man. Bono and other activists fight against discrimination and stereotypes of their community."

The passage in the religious collection highlights the tension between LGBTQIA+ rights and religious freedoms. Given that this is the only representation of LGBTQIA+ individuals in our sample, children exposed to this curriculum are only seeing people in this community posed as a threat to the freedoms of others. The passage in the collection of California

textbooks gives definitions of what it means to be a gay, lesbian, bisexual, and/or transgender individual and highlights two LGBTQIA+ activists. While the sentiment of this passage is positive overall, by not discussing any LGBTQIA+ individuals of color or who identify as female, this passage fails to take an intersectional lens. We did not identify any passages in the Texas state adopted curriculum that discuss LGBTQIA+ topics, rights, or individuals.

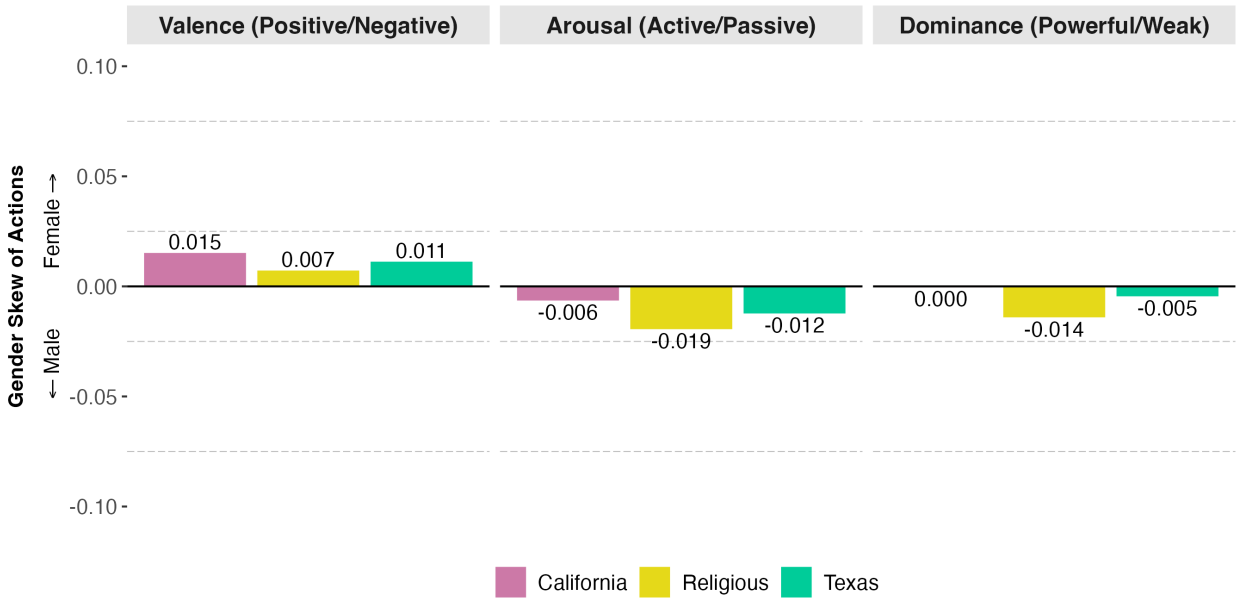
V.C How: Examining Context & Actions

In the preceding section, we identified disparities in who is present (and absent) from textbooks. However, it is crucial to examine not only how often different groups are mentioned but also the context and manner of their portrayal. In our previous discussion, we alluded to variations in the portrayal of LGBTQIA+ individuals across curricula through comparative analysis of passages. In this section, however, we conduct a systematic examination of the differences in the actions and context surrounding different groups of individuals, specifically focusing on race and gender.

Actions. We apply semantic role labeling as detailed in Section IV.B.4, to characterize the sentiment associated with actions attributed to various groups of people, we use the VAD lexicon to measure the sentiment of actions on three dimensions: valence (postive vs. negative), arousal (active vs. passive), and dominance (powerful vs. weak). The NRC Valence, Arousal, and Dominance (VAD) Lexicon includes a list of more than 20,000 English words and their valence, arousal, and dominance scores. For a given word and a dimension (V, A, or D), the scores range from 0 to 1 (Mohammad, 2018). For example, the verb "see" (VAD scores 0.635, 0.269, and 0.312 respectively) is more positive but less active and less powerful than the verb "say" (VAD scores 0.571, 0.31, and 0.482 respectively).

To investigate the presence of systematic differences in the portrayal of actions between females and males, we calculate the average valence, arousal, and dominance scores for actions associated with female agents and male agents in respective sentences. By subtracting these averages, we derive a gender skew metric for each sentiment dimension. In Figure 9, we

FIGURE 9
Average Difference in VAD Scores of Male vs. Female Actions



see that across all collections of textbooks, the actions taken by females are more positive, but less active and powerful than actions take by males. This pattern remains even when we focus only on sentences where the gender of agents is identified using only gendered pronouns and not names and gendered terms (Appendix Figure BVII) and when we focus on actions that are only taken by female agents vs actions that are only taken by male agents (Appendix Figure BVIII).

Context. In Section IV.B.6, we detail our use of word embeddings to analyze the contexts surrounding groups of people. We measure context in two different ways; through the words most closely associated with different groups of people and by testing whether groups of being are more likely to be found in the context of specific domains (such as the home or politics).

First, we identify the top relative words most closely associated a target group (such as females) and least associated with a comparison group (such as males). This approach follows a similar methodology outlined in Manzini et al. (2019) and Charlesworth, Caliskan

and Banaji (2022), where the "top relative words" are found by computing the group-to-word association between a single word and the target (or comparison) group, repeating for a given list of target words, and finding the top N words for which the difference in group-to-word associations between the target and comparison group is the greatest. More formally, group-to-word association is calculated using Equation 1 where the association between group G and word w can be calculated as $a(G, w)$:

$$a(G, w) = \frac{1}{|G|} \sum_{g \in G} cos(g, w) \tag{1}$$

where $cos(g, w)$ indicates the cosine similarity (e.g. distance in embedding space) between the two words embeddings g and w. We take the difference in word associations between the target and comparison groups to test whether a word w is "skewed" towards the target group (e.g. is more often found in the same context as the target group) as shown in Equation 2:

$$skew(w) = a(G_{target}, w) - a(G_{compare}, w) \tag{2}$$

Finally the top N relative words most closely associated a target group and least associated with the comparison group are the N words $w \in W$ with the largest values of $skew(w)$. We restrict the set of target words to be the 20,000 English words in the VAD lexicon for which we have valence, arousal, and dominance scores (Mohammad, 2018).

Table 4 shows the top 10 relative words most closely associated with Female and Males relative to each other for each collection of textbooks in our sample along with the average VAD scores for the top words. We see that Females are more associated with words related to clothing and males are more associated with words relating to the military. The sentiment of the top 10 relative words for females are more positive, but less active and powerful on average than top relative male words. These findings are in line with the patterns of sentiment in actions taken by females versus males, as discussed in the previous section.

TABLE 4
Top 10 Relative Words by Gender

		Mean		
Top Words Associated with Female (vs. Male) Pronouns		V	A	D
California	husband, dress, sew, quilt, kimono	0.665	0.309	0.386
Religious	husband, shawl, skirt, pale, quilt	0.603	0.342	0.316
Texas	dress, handwriting, fabric, groceries, lobby	0.675	0.332	0.392
Top Words Associated with Male (vs. Female) Pronouns		V	A	D
California	sword, horse, stallion, master, fearless	0.641	0.611	0.726
Religious	glorify, believed, army, future, faithful	0.727	0.578	0.705
Texas	horse, sword, lasso, general, king	0.604	0.559	0.665

This pattern is still present when we extend the our analysis to include the top 50 relative words by gender as shown in Figure 10.

We can test whether different groups have a greater association with different domains (such as politics or the home). To do this, we extend Equation 1 to calculate group-to-domain similarity between a group G and a domain D as shown in Equation 3 below:

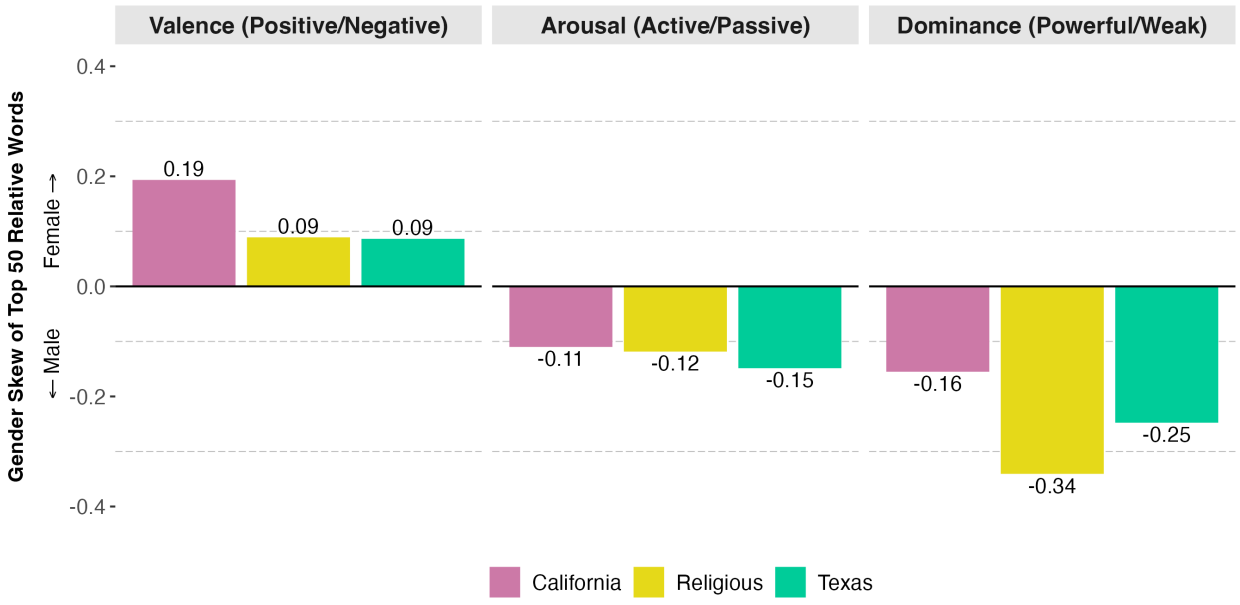
$$A(G, D) = \frac{1}{|G| \times |D|} \sum_{g \in G} \sum_{d \in D} cos(g, d) \tag{3}$$

Similarly, we extend Equation 2 to test whether a domain D is "skewed" towards the target group (e.g. if words in domain D are more often found in the same context as the target group G) as shown in Equation 4:

$$Skew(D) = A(G_{target}, D) - A(G_{compare}, D) \tag{4}$$

In figure 11, we show the gender skew of the Family domain (which includes words such as marriage, children, relatives, household, etc) and the Politics domain (which includes words such as government, army, democracy, capital, political, etc) is more skewed towards males. We test whether this measure of gender skew is statistically different from zero using a placebo test in which we randomly draw a set of words from the vocabulary of the textbook

FIGURE 10
Average VAD Scores of Top 50 Words by Gender

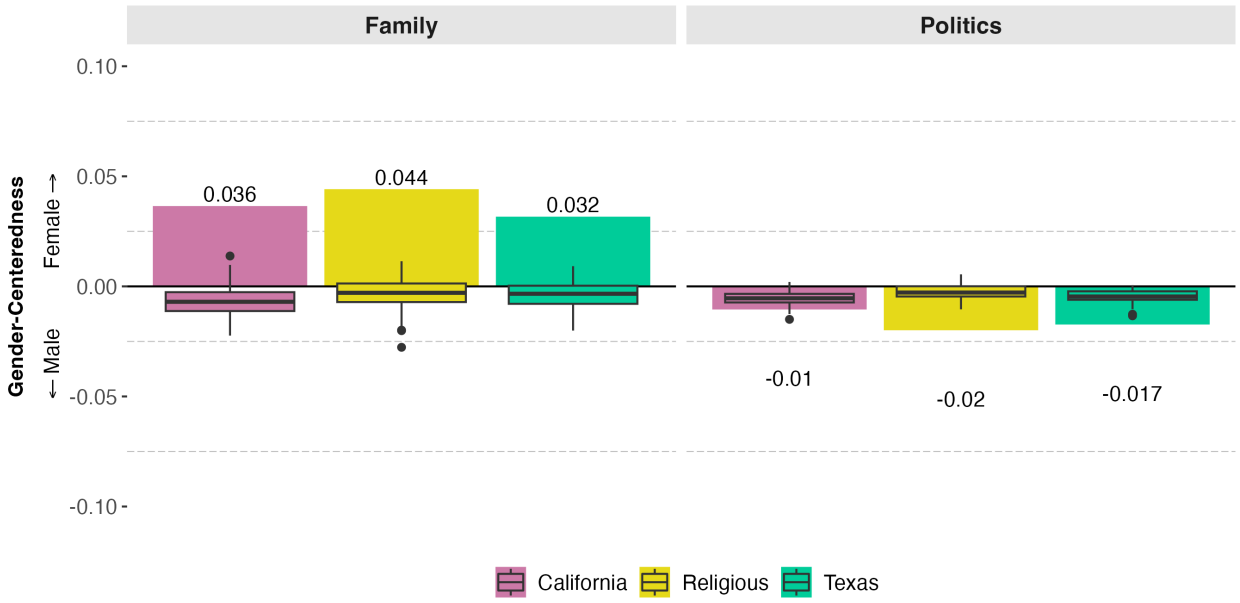


collection which is the same size and the list of words in the domain we are testing for gender skew in and calculate the skew for this "random" domain. We repeat this 1,000 times and plot the distribution of these random domain associations over the top of the bar graphs using a barplot where the edges of the barplot show the 10th and 90th percentile values of the possible gender skewness in a domain of random words. Compared to this distribution, we find that in all three textbook collections, the Family domain is skewed towards female, and this measure of skewness is statistically different than the skewness we would find in a random set of words. We find a similar (though smaller) gender skew in the Politics domain towards males for the Texas and Religious textbook collections, but this estimate is not statistically different than a the gender skew found in random set of words in the California textbook collection.

VI Conclusion

Previous literature has established the importance of what we teach our children on a variety of outcomes such as academic performance and career choices. As a result, it is

FIGURE 11
Gender Skew in the Family and Politics Domain



vital that we understand the representation of people and topics in the educational content we present to our children. Given the growing polarization in our society and the increase in demand for school choice options that more closely align with parental values, it's necessary to examine curricula used by children across different geographic regions both within and outside the traditional public school setting. However, traditional content analysis methods are costly and time consuming, which previously limited the scope of such analysis.

This paper demonstrates how advances in natural language processing and computer vision tools can be applied to measure how different topics and groups are portrayed in educational materials on a larger scale than was previously possible – such as within entire curricula. The methods introduced in this work also expand the set of tools available for education researchers, and social scientists more broadly, to measure representation in a variety of contexts. For example, advances in the measurement of skin color in images could be applied to measure the effect of student-teacher skin color congruence.

We find that textbooks commonly used in religious private school and homeschool

settings depict faces with lighter skin colors on average as compared to our sample of public school textbooks from Texas and California, more mentions of creationism than climate change, and a large proportion of pages dedicated to religious topics (even in Science textbooks).

These religious textbooks also have less representation of females in both the text and images in comparison to textbooks commonly used in public schools. However, even in public school textbooks females are under-represented relative to the share of the population. More importantly, across all collections, not only are females talked about less often than males, they are more likely to be mentioned in contexts that reproduce societal norms and disparities. For example, females are more likely to be talked about in the context of the family and less likely to be talked about in the context of politics. Words found in similar contexts to females are more positive but less active and powerful than words found in similar context to males.

There is a popular narrative espoused by political and media officials of polarization across the country and across different religious and ideological groups. While these methods may not pick up smaller mentions of differences, overall, what we see is not consistent with a narrative of polarization in terms of what is talked about. There are meaningful disparities in who is present (and absent), but these disparities remain similar in whether and how these groups are portrayed across public and private school textbooks.

In the midst of an increase in educational gag orders and statewide changes in the content standards of public school curriculum, our work shines a light on what messages children are receiving from informal and non-traditional curricula outside the public school system, which can have long-term implications for children's beliefs and decision-making.

References

Abdulkadiroglu, Atila, Parag A Pathak, and Christopher R Walters. 2015. "School vouchers and student achievement: Evidence from the Louisiana Scholarship Program."

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz. 2023. "What we teach about race and gender: Representation in images and text of children's books." The Quarterly Journal of Economics, 138(4): 2225–2285.

Allen, Ann M, Daniel N Allen, and Gary Sigler. 1993. "Changes in sex-role stereotyping in caldecott medal award picture books 1938—1988." Journal of Research in Childhood Education, 7(2): 67–73.

Arold, Benjamin W. 2022. "Evolution vs. Creationism in the classroom: The lasting effects of science education." Working Paper.

Ash, Elliott, Ruben Durante, Maria Grebenschikova, and Carlo Schwarz. 2021. "Visual representation and stereotypes in news media."

Bartholet, Elizabeth. 2020. "Homeschooling: Parent rights absolutism vs. child rights to education & protection." Ariz. L. Rev., 62: 1.

Bickford, John Holden, and Lori A Knoechel. 2017. "The historical representation of Native Americans within Ppimary-and intermediate-level trade books." The Councilor: A Journal of the Social Studies, 78(2): 4.

Braden, Eliza G, and Sanjuana C Rodriguez. 2016. "Beyond mirrors and windows: A critical content analysis of Latinx children's books." Journal of Language and Literacy Education, 12(2): 56–83.

Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. "Semantics derived automatically from language corpora contain human-like biases." Science, 356(6334): 183–186. Publisher: American Association for the Advancement of Science.

Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, and Y Jane Zhang. 2017. "Curriculum and ideology." Journal of Political Economy, 125(2): 338–392.

Catt, Andrew D, and Evan Rhinesmith. 2017a. "Why Indiana Parents Choose: A Cross-Sector Survey of Parents' Views in a Robust School Choice Environment." EdChoice.

Catt, Andrew D, and Evan Rhinesmith. 2017b. "Why Indiana Parents Choose: A Cross-Sector Survey of Parents' Views in a Robust School Choice Environment." EdChoice.

Charlesworth, Tessa ES, Aylin Caliskan, and Mahzarin R Banaji. 2022. "Historical representations of social groups across 200 years of word embeddings from Google Books." Proceedings of the National Academy of Sciences, 119(28): e2121798119.

Clark, Roger. 2007. "From margin to margin? Females and minorities in Newbery and Caldecott Medal-winning and Honor books for children." International Journal of Sociology of the Family, 263–283.

Crisp, Thomas, Roberta Price Gardner, and Matheus Almeida. 2018. "The all-heterosexual world of children's nonfiction: A critical content analysis of LGBTQ identities in Orbis Pictus Award books, 1990–2017." Children's Literature in Education, 49(3): 246–263.

Deming, David J, Justine S Hastings, Thomas J Kane, and Douglas O Staiger. 2014. "School choice, school quality, and postsecondary attainment." American Economic Review, 104(3): 991–1013.

Friedman, Jonathan, and James Tager. 2022. "Educational gag orders: Legislative restrictions on the freedom to read, learn, and teach."

Gardner, Matt, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson Liu, Matthew Peters, Michael Schmitz, and Luke Zettlemoyer. 2018. "Allennlp: A deep semantic natural language processing platform." arXiv preprint arXiv:1803.07640.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word embeddings quantify 100 years of gender and ethnic stereotypes." Proceedings of the National Academy of Sciences, 115(16): E3635–E3644.

Keith, Verna M, and Carla R Monroe. 2016. "Histories of colorism and implications for education." Theory Into Practice, 55(1): 4–10.

Klein, Rebecca. 2017. "Voucher schools championed By Betsy DeVos can teach whatever they want. turns out they teach lies."

Koss, Melanie D. 2015. "Diversity in contemporary picturebooks: A content analysis." Journal of Children's Literature, 41(1): 32.

Koss, Melanie D, and Kathleen A Paciga. 2020. "Diversity in Newbery Medal-winning titles: A content analysis." Journal of Language and Literacy Education, 16(2): n2.

Koss, Melanie D, Nancy J Johnson, and Miriam Martinez. 2018. "Mapping the diversity in Caldecott books from 1938 to 2017: The changing topography." Journal of Children's Literature, 44(1): 4–20.

Kozlowski, Austin C, Matt Taddy, and James A Evans. 2019. "The geometry of culture: Analyzing the meanings of class through word embeddings." American Sociological Review, 84(5).

Kunzman, Robert. 2010. "Homeschooling and religious fundamentalism." International Electronic Journal of Elementary Education, 3(1): 17–28.

Lucy, Li, Dorottya Demszky, Patricia Bromley, and Dan Jurafsky. 2020. "Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas US history textbooks." AERA Open, 6(3): 2332858420940312.

Manzini, Thomas, Yao Chong Lim, Yulia Tsvetkov, and Alan W Black. 2019. "Black is to criminal as caucasian is to police: Detecting and removing multiclass bias in word embeddings." arXiv preprint arXiv:1904.04047.

Mohammad, Saif M. 2018. "Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words." Melbourne, Australia.

National Center for Education Statistics. 2022. "Homeschooled Children and Reasons for Homeschooling." Condition of Education, Retrieved December 18th, 2023.

Riley, Emma. 2022. "Role models in movies: the impact of Queen of Katwe on students' educational attainment." Review of Economics and Statistics, 1–48.

Scaramanga, Jonny, and Michael J Reiss. 2018. "Accelerated Christian Education: A case study of the use of race in voucher-funded private Christian schools." Journal of Curriculum Studies, 50(3): 333–351.

Stevenson, Betsey, and Hanna Zlotnik. 2018. "Representations of men and women in introductory economics textbooks." Vol. 108, 180–85.

Szasz, Teodora, Emileigh Harrison, Ping-Jung Liu, Ping-Chang Lin, Hakizumwami Birali Runesha, and Anjali Adukia. 2022. "Measuring representation of race, gender, and age in children's books: Face detection and feature classification in illustrated images." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision.

Wade, Rahima C. 1993. "Content analysis of social studies textbooks: A review of ten years of research." Theory & Research in Social Education, 21(3): 232–256.

Yanagizawa-Drott, David H., and Hans-Joachim Voth. 2023. "Image(s)." Working Paper, University of Zurich.

Yu, Amy Zhao, Shahar Ronen, Kevin Hu, Tiffany Lu, and César A Hidalgo. 2016. "Pantheon 1.0, a manually verified dataset of globally famous biographies." Scientific Data, 3(1): 1–16.

A Appendix Tables

TABLE AI
Female Representation in Text and Images

	Dependent variable:	
	% Female Words	% Female Faces
	(1)	(2)
Religious	-0.115^{***}	-0.038^{***}
	(0.007)	(0.007)
California	-0.021^{***}	0.030^{***}
	(0.008)	(0.009)
Grade 5	-0.055^{***}	-0.013^{**}
	(0.006)	(0.006)
Science	-0.122^{***}	0.043^{***}
	(0.007)	(0.008)
Social Studies	-0.216^{***}	0.008
	(0.007)	(0.008)
Copyright Year	0.0001	-0.0002
	(0.0003)	(0.0003)
Constant	0.333	0.830
	(0.575)	(0.598)
Observations	2,521	2,434
Adjusted R^2	0.331	0.057

B Appendix Figures

FIGURE BI
Number of Students Enrolled in Private School by Religious Orientation Over Time

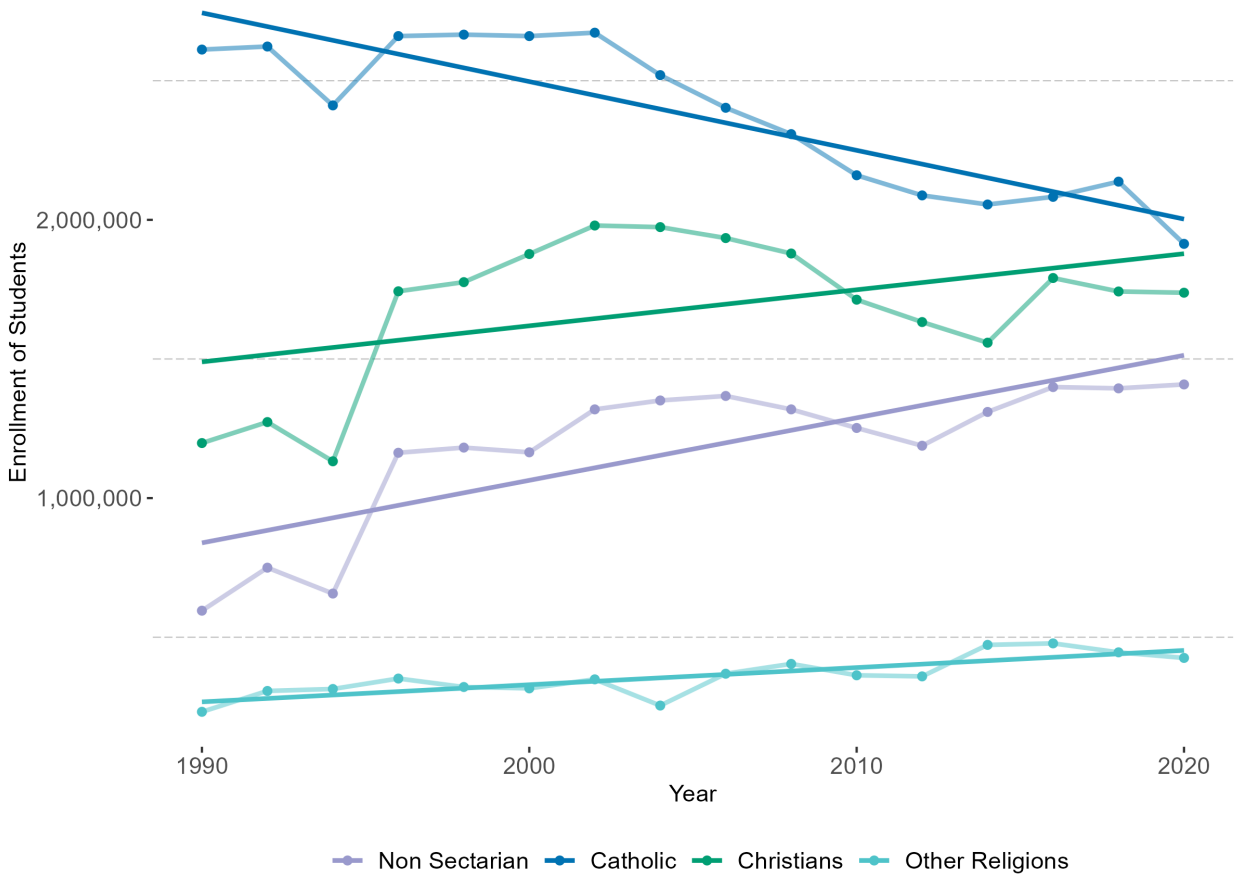


FIGURE BII
Average Percent of Pages Containing Contentious Topics

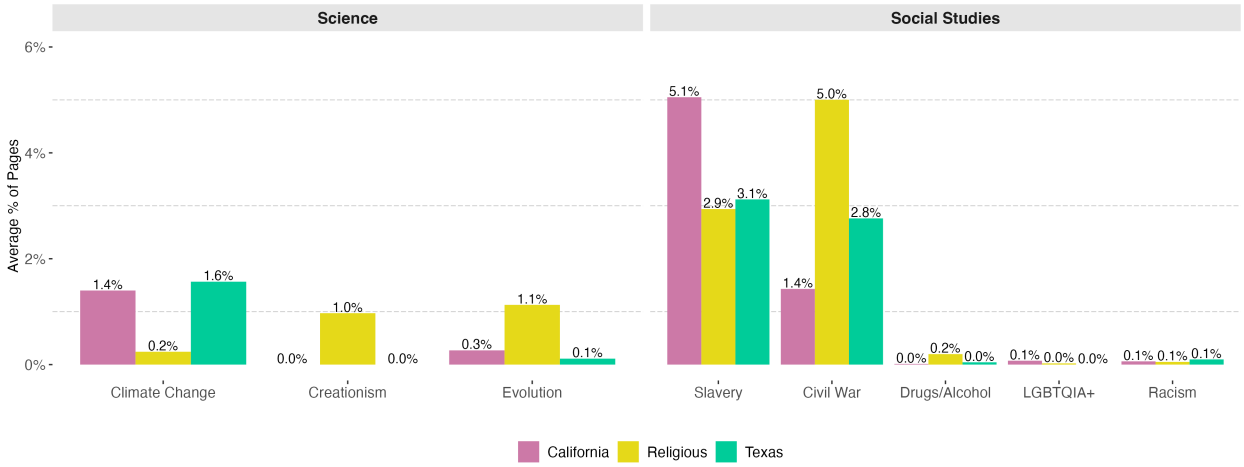


FIGURE BIII
Proportion of Faces Depicted by Gender and Collection

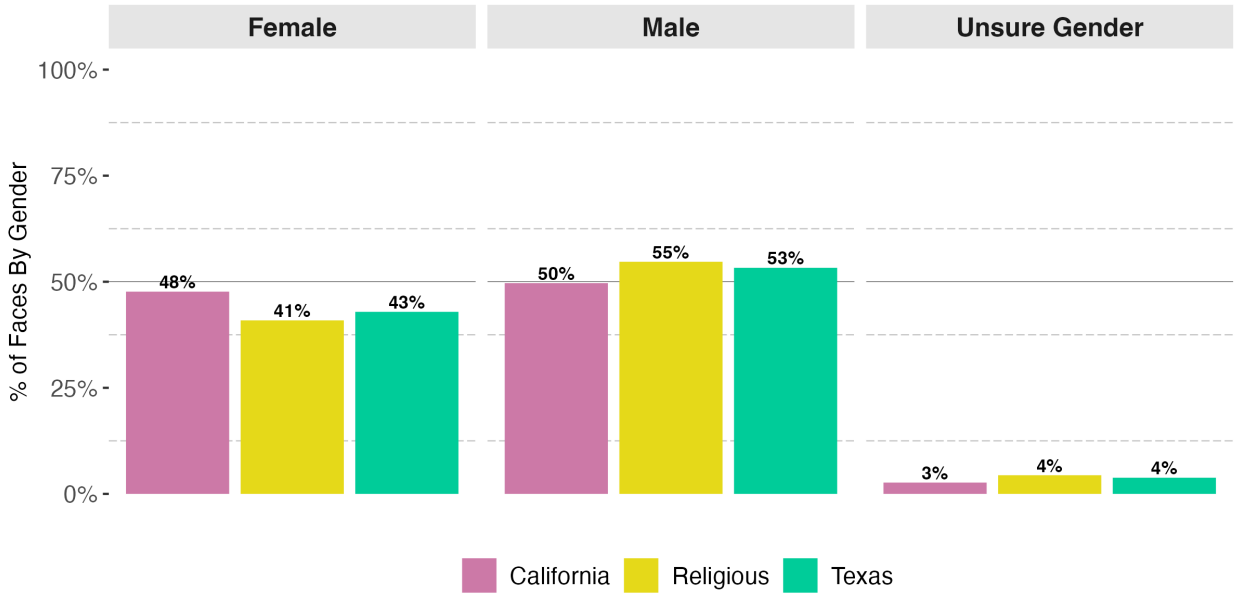


FIGURE BIV
Percent of Female Words by Subject Over Time

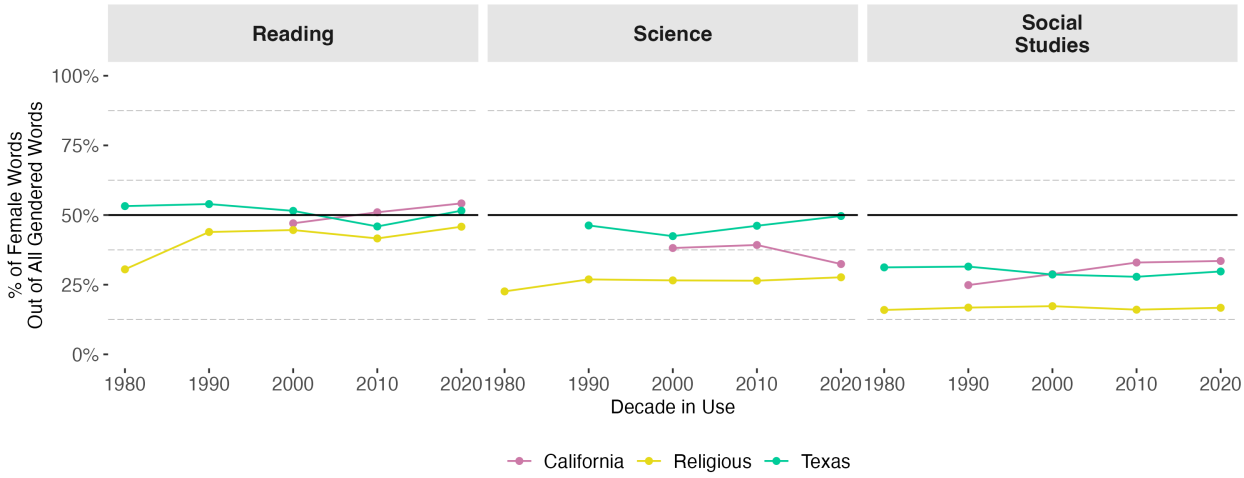


FIGURE BV
Race/Ethnicity of Faces by Collection and Subject

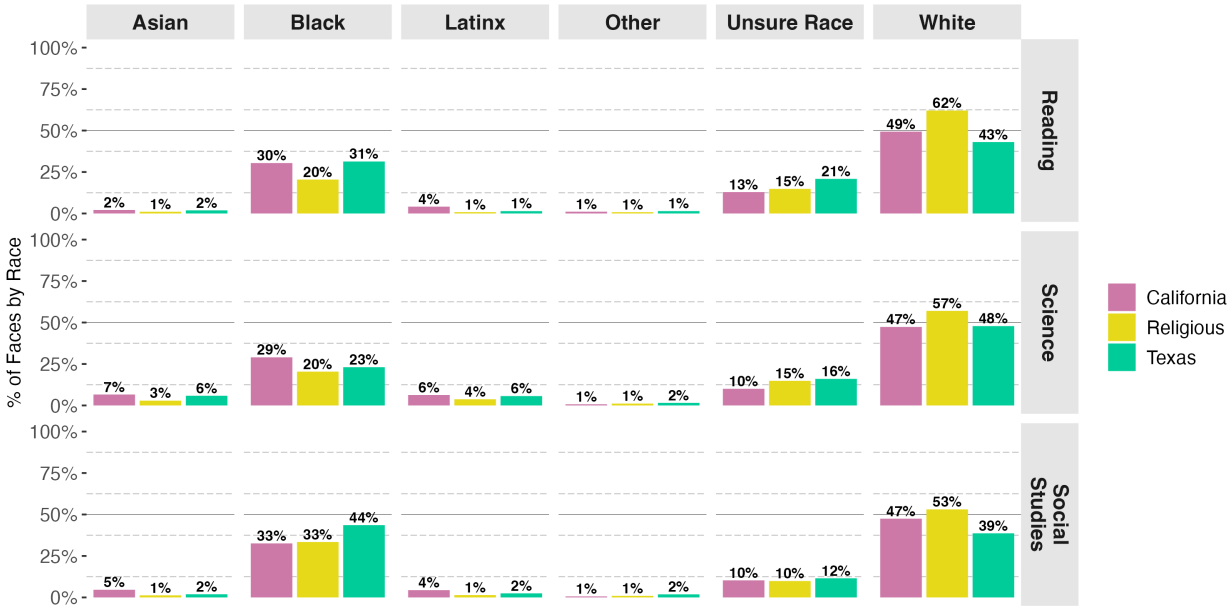


FIGURE BVI
Percentage of Famous Individuals By Race in Each Collection Over Time

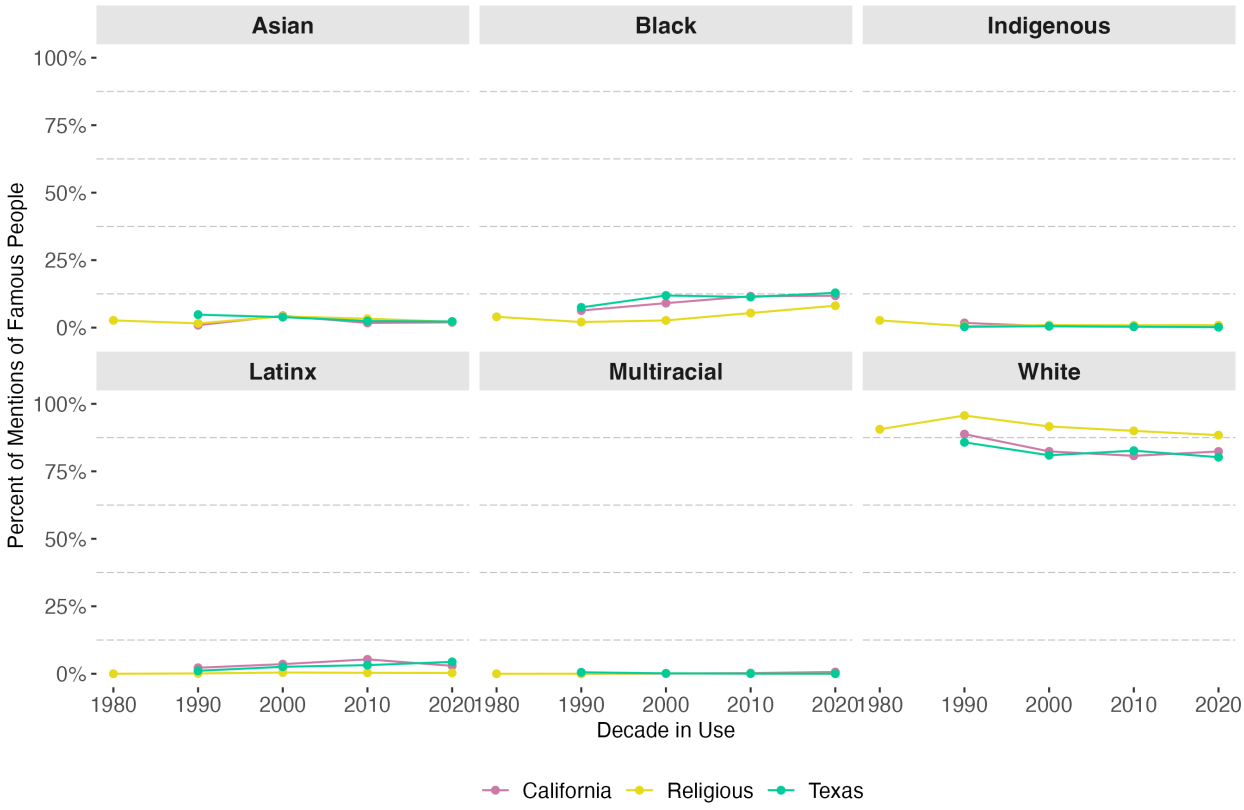


FIGURE BVII
Average Difference in VAD Scores of Actions Associated with Male Pronouns vs. Actions Associated with Female Pronouns

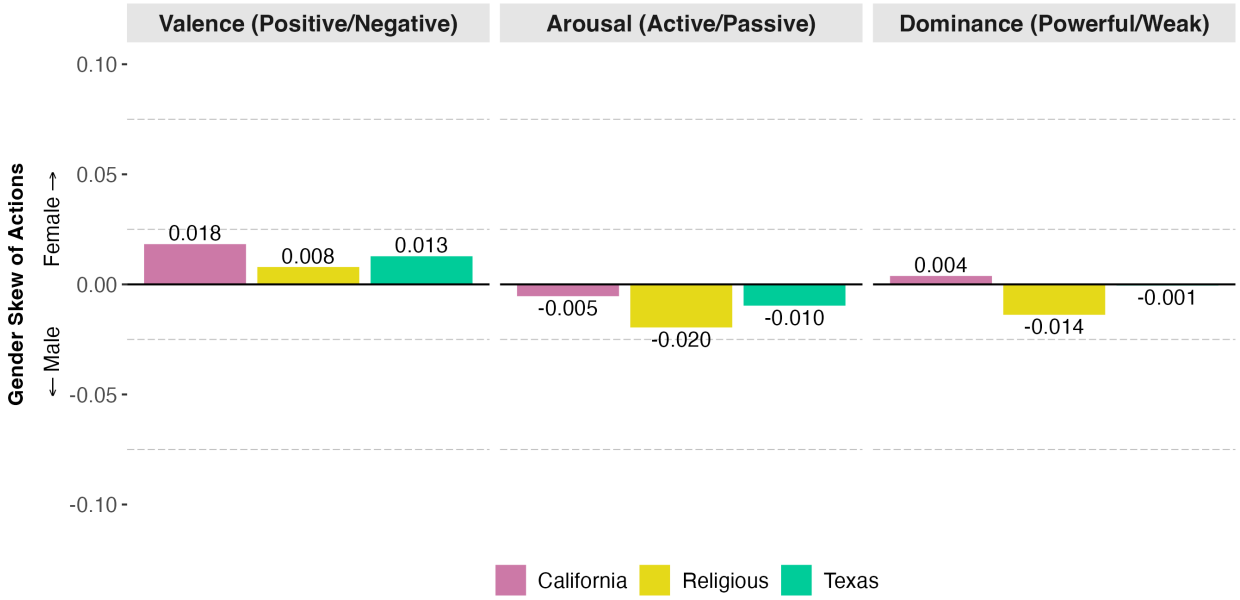


FIGURE BVIII
Average Difference in VAD Scores of Actions Taken Only by Males vs. Actions Taken Only by Females

