

SUPPLEMENTARY MATERIAL

IMAGINE ALL THE PEOPLE

Literature, Society, and Cross-National Variation in Education Systems

By CATHIE JO MARTIN*

World Politics

doi: 10.1017/S0043887118000023

Replication data for this article are available at:

Martin, Cathie Jo. 2018. "Replication Data for: Imagine All the People: Literature, Society, and Cross-National Variation in Education Systems." Harvard Dataverse, V1. doi: 10.7910/DVN/EBM4G6.

SUPPLEMENTARY MATERIAL

WORD SELECTION

With a few exceptions (additions noted below), I derive the dictionary word lists in each category by taking the top 200 words in my matched sets of novels (after removing stop words) and coding them into categories.

I added “education” to the list of education words derived from the top 200 words. I added “individual,” “liberal” and “independent” to the list of individual words because these go to the heart of the category. The findings remain essentially the same without these additions. I made no addition to the feeling words or political governance words categories, and all came from the top 200 words. For the societal and collectivist words, I added “social,” “collective,” and “communal.” For the upper class words, I did not include “sir” (which greatly increased the British frequencies), because children use this term to address teachers.

A large share of literature included in the HathiTrust and Google books collection have publication dates that differ from the initial publication of a work. Therefore, one cannot calculate the accurate use of a work in literature over time using the Google ngram viewer. I solve this problem by inserting the initial publication date in my volumes.

DATA COLLECTION AND SOURCES

My basic assumption is that an analysis of cross-national differences in literary artifacts should help us to grasp the complicated relationship between culture and cross-national differences in political system development.

I begin this study with the development of two corpora of literature. The Danish corpus of novels, poems and plays is drawn from the Archive of Danish Literature, a website for “classic Danish literature” that seeks to develop a comprehensive collection of leading authors’ work. (http://adl.dk/adl_pub/omadl/cv/OmAdl.xsql?nnoc=adl_pub). These files were supplemented by works from other lists of major Danish works, and I obtained the additional full-text files from HathiTrust. Here are the additional lists of great Danish works: http://kum.dk/fileadmin/user_upload/KUM_kulturkanonen_OK2.pdf; <https://sites.google.com/site/danskfag/litteratur-1/litteraturhistorisk-oversigt>

I constructed a parallel corpus of leading British work by identifying leading novels, plays and poems from multiple online lists. The lists are available upon request.

MACHINE LEARNING METHODS

Scholars using machine learning or natural language processing analyses in political science typically use either supervised learning techniques or unsupervised learning techniques. Supervised learning techniques may be used to classify works by organizing them into categories or to scale works by determining the position of a document on a linear scale, i.e. from left to right. Unlike hand coding, computational techniques may treat texts not as discourse but as

collection of words in form of word frequencies. This allows one to compare the appearance of words in diverse texts and to calculate confidence intervals and point estimates.

For example, Hopkins and King (2010) develop a “bag of words” technique with which they hand-code texts into set of exhaustive categories and use the output to estimate the proportion of documents into each category. Their goal is to get an accurate estimate of document categories and to represent unstructured text as structured variables that can be analyzed statistically. Therefore they try to develop a way to present natural language in terms of numerical variables, by presenting each word in the bag of words as a dichotomous variable – it either appears or does not appear. They begin with the small, hand-coded set, and then estimate the proportion of all population documents that fall into a given category. King, Pan and Roberts (2013) use machine learning techniques to reveal preferences for motivations of Chinese officials in censoring social media posts. They find that criticism or support of government is not the key trigger for censorship; rather censorship happens when posts mobilize collective action and social mobilization. The authors identify a priori 85 topic areas in three issue areas of sensitivity. They then randomly select posts to see if the posts are subsequently censored, and monitor posts in real time. 13 percent of posts are censored. They define a set of exhaustive & mutually-exclusive categories, code selected posts within each category for the “training set” and then infer the “test set” without hand coding. Their “Read Me” open source software program allows them to estimate the percent of posts in each category, and to categorize those supporting the state versus those opposing the state. Miller (2013) analyzes UN General Debate speeches to demonstrate the significant differences in the public discourse of colonial and non-colonial states, with the largest difference reflecting to expressions of victimization by colonial states. Initially totaling all words, the author then calculates the probability of a word appearing in each speech, plots the distributions of words, subtracts the probability of a word in one country from the word in the other country, and presents a chart to show these distribution of differences (or D values). For example, he calculates D values for various dimensions of victimization captured in words such as inequities, vulnerable, suffer, struggle, and unequal.

Laver, Benoit and Garry (2003) develop the Wordscore text scaling method. The scaling method can proceed in two ways: a priori policy positions versus inductive surveys. The inductive approach (unsupervised learning discussed below) seeks observed patterns in the texts showing a matrix of similarities and differences. The substantive meaning emerges through an a posteriori dimensions and one does not direct the process or set the meaning of the underlying policy dimensions in advance. But problem with this is that there are no objective criteria for choosing among competing interpretations of the spatial dimensionality. Therefore, Laver et al try to find a compromise, by identifying a priori words in advance and developing a spatial representation of the texts, but then changing the dimensions as they proceed. Thus, they identify “reference texts” (a priori) and then use “virgin texts” to make a posteriori corrections to the dimensions. They generate a score for each word (a priori) and use the word scores from the reference texts to make predictions about the virgin texts. But also use the findings from the virgin texts to refine the analysis. Each new finding in a virgin text increases the probability of finding more Bayesian possibilities. Baturo and Mikhaylow use the Wordscore text scaling method as well to explore followers’ responses to Putin’s statism versus Medvedev’s liberalism. Texts are placed on a single dimension, determined in advance, with two documents representing ideal types at each end of spectrum. Other texts are scaled according to these ideal types.

There are some methodological problems with supervised learning techniques. If the labeled set is not a random sample from the population, then it may not reflect the larger population. Moreover, it is very hard to find the true model (Hopkins and King 2010).

Thus scholars also use unsupervised-learning techniques such as probabilistic topic models. Rather than beginning with keywords, one may ask the computer to identify the main themes or topics in a collection of documents. A topic is a “distribution over a fixed vocabulary,” and documents have multiple topics, each with their own particular distribution of topics (Blei). There are probabilities of topics appearing in each document and of words appearing in each topic. With processes such as the latent Dirichlet allocation if one simply specifies the number of topics in advance, one can use a “topic model algorithm” to infer the hidden topic structure, and compute the distribution of topics that best captures the collection of words in a document. Machine learning defines a joint probability distribution over observed and hidden random variables. So one computes the conditional distribution of the hidden variables, given the distribution of the observed variables. This conditional distribution is the “posterior distribution.” In LDA, the observed variables are the words and the hidden variables are the topic structure. So the point is to infer the hidden topic structure from the documents, called computing the posterior distribution (Blei).

One issue using this technique is how to get the algorithms. 1) Sampling-based approaches approximate the posterior with samples. 2) Variational methods posits a “family of distributions over the hidden structure” and finds a member closest to this posterior. 3) One may look at the order of words, rather than simply the bag of words. “Topics generate words conditional on the previous words.” 4) Dynamic topic modeling means that rather than a topic being a simple distribution over words, a topic can be a “sequence of distributions over words.” The LDA may also incorporate feedback. Thus a Bayesian non-parametric topic model relaxes the assumption that the number of topics is known and fixed, and new documents reveal new topics (Blei; Catalinac).

Thus, Catalinac uses a probabilistic topic model, the Latent Dirichlet Allocation approach, in her analysis of topics in Japanese electoral manifestos before and after the 1994 electoral reform. She explores whether electoral reform makes candidates go from targeting the median voter (expected under Majoritarian systems) to groups of voters (expected with PR). Rather than specifying topics in advance, she uncovers the topics and estimates the probability that each document covers the topic. She then creates a term-document matrix: words in rows, document identifiers in columns, frequencies in cells. The word frequencies allow her to make inferences about the topics and to create estimates of probabilities that the topic will appear. She bases her assessment of particularistic versus programmatic topics on whether policies benefitted large or small groups (with particularistic obviously more geared to small groups.) She calculated the mean percent of discussions of particularistic and programmatic goods and compared the means before and after the reforms. She also used “wordfish,” which uses word frequencies to infer the locations of a document on an ideological scale (Slapin & Protesch)

In a similar mode, Nielsen (Richard, 2013, *The Lonely Jihadist*. Harvard Dissertation) uses Latent Dirichlet allocation (LDA), or topic modeling to think about why some clerics become jihadists whereas others do not. His major a priori task is to specify different fixed number of

topics in advance, and he chooses 5, 10, 20 and 100. He identifies clusters of correlated words within the corpus and found the 5 topics approach to be most useful in laying out the broad themes in Islam. All words are assigned to one of the 5 clusters, and scholars typically zero in on the top 20 words in each cluster to reduce complexity. Combination of techniques to get topic labels: frequency (but problem with words such as “Allah”) and exclusivity (words mainly appear in only one cluster, so may be a better indicator). Picks texts that are most representative of each cluster, and classifies authors according to similarities of writings.

My study has several important difference from scholarly research discussed above. First, the differences between the documents I wish to compare are preset and unambiguous. I do not need to scale or to sort the documents, as they have already been sorted into classifications for me. They constitute corpora of literature in different countries (the cross-national comparison) or works written at different points in time (the temporal differences), and my task is to compare the characteristics of these documents along dimensions that I have hypothesized from my assumptions about the cultural differences that might give rise to diverse types of education systems. Because the differences among sets of documents to be compared are so rigid, I felt confident using word frequencies.

Second, because I analyze literature instead of tweets or posts, there are likely to be a huge number of topics that have nothing to do with education. I use both supervised learning (with the Python Natural Language Tool Kit techniques for calculating word frequencies) and unsupervised learning (with probabilistic topic modeling) to analyze the corpora. I should note that I calculate word frequencies rather than tf-idf scores because I am not trying to categorize or scale documents within a corpus but rather am comparing frequencies across the corpora and across time. Tf-idf scores may be elevated either because the word is found in large numbers within a document or because the word is only found in selected documents. I assume that words conveying cultural norms about the individual in society, etc, will be found in a large percentage of the documents in each country, whereas the words that do not apply are found in fewer numbers of documents. Thus the gap in tf-idf scores between the two countries should be proportionally lower than the corresponding gap in word frequency scores.

I calculate frequencies of education words appearing in the entire corpora. To calculate the word frequencies, I stem the words in the corpora and take out the stop words. I construct snippets of 50 words around all of the education words that appear in the top 200 words of the selected 8 major volumes. The snippets are stemmed and filtered (i.e. stop words are taken out). The findings are not affected when the snippets include stop words, nor are the findings significantly affected if I choose snippets of 200 words instead of 50 words around the education words.

For the unsupervised topic modeling, I ended up specifying four models because this seemed to be the number that revealed the correct number of topics. Because the medium was literature, the topics had different contextual details; nevertheless, they tended to gravitate toward certain themes even within the different stories. For example, the early Danish ones concerned teaching students in Danish rather than Latin, studying Danish philosophers rather than German ones, and

This design has certain methodological challenges. One problem is that word frequencies alone cannot convey what a text is about. So whereas they are good at capturing broad changes across

countries and time, they are less adept at conveying nuances of the argument. This is why we have two other approaches: the unsupervised topic modeling and the qualitative close reading of selected major literary texts. The topic modeling gives us a better sense of what words are associated with education words, but the weakness is that these are probabilistic rather than deterministic, and topics change every time. Also there is a problem of knowing the right number of topics. The strength of the qualitative close reading of selected major literary texts is that this gives one a deeper understanding of meaning. The weakness is that one is not sure that one has chosen the correct books.

Another problem concerns the structure of the languages. It might be, for example, that a larger number of words are nouns in one language than the other, or that the tendency to form compound words in Danish may alter the findings. The comparison of the English word “give” and the Danish “giver” gives added confidence that the findings do not merely reflect linguistic differences. Moreover, the word frequencies vary in different directions across categories, and this gives added confidence that we are not simply observing linguistic differences.

There is also an issue of timing and genre. It might be that each country has a rise in collective goals when it is actively engaged in nation-building, but that this task comes at different points. Thus in Denmark nation-building is driven by external threats in the early 1700s and 1800s, whereas nation-building in Britain happens only with the expansion of mass suffrage after 1864. The nation-building moment has passed in Denmark by the end of the 19th century, which is why the countries converge. This argument helps to explain the differential timing of mass education, but does less to account for the differences in form – standardized academic track versus a large vocational track. Also the largest threat to Denmark was in 1864, when it lost most of Schleswig-Holstein. Yet the country became less focused on education after this (and substantially more focused on the industrial project.)

Finally, it is easier to demonstrate the differences in the representation of education in literary works occur across countries and time than it is to show decisively that these different depictions of the “self” have bearing on the development of diverse types of educational institutions. The larger book project explores authors’ engagement with policy-makers in constructing education reforms. I also look at the relationship between depictions of education-related issues in literature and depictions of these in political magazines.

READERSHIP

A set of questions about evaluating the impact of literature has to do with readership. Three questions: Did people read and who read? Did literature convey norms? Were policymakers specifically influenced by authors? Cultural artifacts will only reflect or construct the underlying culture if the works are read; and happily literary studies offer us much reflection on the question of readership.

First is the question of who read. My sample draws from list of identified classics, but how widely read were these works in their own time? Several bodies of readers are relevant to the hypotheses. At a minimum, the political class who makes public policy should have read literature, if fiction matters to perceptions of education reform. Yet to the extent that public

opinion shapes policymaking in pre-democratic regimes, a broader swath of the middle class and even the working class should be exposed to these works for the argument to hold.

I deliberately chose works that are identified (by various lists) as “classic works,” because these works were largely very well-known in their own era and have survived the test of history. There certainly is not a one-to-one relationship between literary output and public policy: that the works are now considered to be great does not ensure their contemporaneous popularity; moreover, their espousal of certain themes does not mean that their readers are influenced by these cultural assumptions. Therefore, the following paragraphs offer some background data on readership: who was reading these books and how they might be influenced by these text?

Fortunately, recent scholars of digital humanities have done remarkable work tracing the reading of texts, and there is ample evidence that fiction played a huge role in shaping the public’s perception of politics in the eighteenth and nineteenth centuries. My own goals (and capacities) are more modest, because as a political scientist, my focus is on the possible impacts of cultural artifacts on policy outcomes rather than on the patterns of use of cultural artifacts.

My broader research focuses on two types of readers. First, I explore how novels shaped intellectual elites’ views of politics and political reforms. Elite reading patterns give insight into how fiction influences policy decisions in authoritarian regimes, as literature was a site of political struggle and public opinion among elites. The literary community debated the grand issues in salons and taverns on a daily basis and this debate within the “republic of letters” led to social cohesion and shared values among the reading public. It facilitated debate among authors and spread their views to the broader reading public (Keen, 1999, 29-30). In short, Keen suggests that by the 1780s and 1790s, literature became the “single most effective means by which people could engage each other in a rational debate whose authority all governments would be compelled to recognize.” Thus literature was a kind of “group project” for projecting the groups’ interests onto the public consciousness so that “relations of power would give way to questions of morality.” (Keen 1999, 33). William Godwin reflected these views in 1793 when he wrote: “Few engines can be more powerful, and at the same time more salutary in their tendency, than literature. Without enquiring for the present into the cause of this phenomenon, it is sufficiently evident in fact, that the human mind is strongly infected with prejudice and mistake. The various opinions prevailing in different countries and among different classes of men upon the same subject, are almost innumerable; and yet of all these opinions only one can be true. Now the effectual way for extirpating these prejudices and mistakes seems to be literature” (cited in Keen, 1999, 28). Dissenters were particularly pronounced within this group (Keen, 1999 38). Romanticism at the end of the 18th century was a huge political project, and for Shelly, poets were de facto legislators (Keen 1999, 27).

Second, I explore how mass readers were exposed to the cultural assumptions of fictional works. For example, did these cultural assumptions play a role in how the masses thought about industrial institutions and democratic reforms in the nineteenth century? Of course, if the reading of classic works stayed within the elite class, literature may well have an impact on political reforms but be less influential on the development of the mentality of working class people.

In continental Europe, the revolution to read and educate in the vernacular – still a topic in early

Danish literature in the 1700s – began in the sixteenth century with the Protestant reformation, Similarly, Henry VIII sought the printing of all English books to occur in London (Gilmont, 2003, 215-17). Reading was elevated in Elizabethan England, but the Protestant regime and civil war caused elites to fear literacy among the common man. The Glorious Restoration in 1688 was associated with diminished literacy. Yet the Society for Promoting Christian Knowledge was formed in 1699 to establish charity schools (Altick, 1954, 30-1).

The reading public started to expand steadily in the early 18th century. Feather (1988, 90-91) suggests that 50 to 60 percent of men were functionally literate by the middle of the 1700s. Watt's estimate is lower, calculating that newspaper readers were at most one in eleven in the mid-1700s, but that readership continued to expand to the end of the century. Two groups of people with more limited means were likely to be readers: apprentices and household servants (Watt, 36, 47). Mostly religious works read during the 18th century, but authors such as Defoe and Richardson crossed genres (Watt 50), and there is evidence that the middle and increasingly the working classes engaged with the leading fiction of the times. Keen believes that there was already a sizable working class readership by the late 18th century (Keen, 1999, 37). Certainly by the 1850s, the reading public fully encompassed a mass audience, and this was a period when there was a reading public of between 5 and 6 million people (Altick, 1954, 4-6.) Uncle Tom's Cabin, the biggest literary phenomenon ever, sold 150,000 copies in the first six months (Altick, 1954, 6). Some believe that the Victorian writers were read by a more unified audience, compared with the fragmented readership for twentieth-century authors; Altick believes that publishers delivered cheaper fare to the working class, but some works such as Dickens' *Household Words* was read very widely (Altick, 1954, 17-20). Indeed, in the mid-nineteenth-century, working-class autodidacts were a veritable movement and authors such as Dickens were like the rock stars of our times (Cordner). Just think of the dock workers, upon Dickens' arrival home from America, shouting up, "What ever happens to Little Nell?"

The centrality of reading is reflected in book sales. It is beyond our scope to lay out book sales for all of the volumes in the corpora; however, the data for leading novels are instructive. Altick estimates the English population to be 6-7 million in 1750. Fielding published 6,500 copies of *Joseph Andrews* in thirteen months, and sold 5000 copies of *Amelia* in the first week (Altick, 1954, 49). The low sales in part reflected the high prices of books; however, this changed in 1774, when the publisher John Bell created a much cheaper venue on coarse paper that was the frontrunner of cheap reprinted six-penny volumes, and John Cooke followed suit with a series of British classics (Altick 1954, 54).

Robinson Crusoe was first written in 1719, before the huge growth of the book trade beginning around 1730, when first editions could total up to 10,000 copies (Feather 1988, 90-91). Yet eleven editions had been issued by 1759 ("Editions of Robinson Crusoe in English," 1936, 22). Readership of Robinson Crusoe expanded to encompass a broader cross-section of people when the novel was reprinted in the *Original London Post* (Watt, 42). In 1774, the publisher John Bell created a cheap book on coarse paper that was the frontrunner of cheap reprinted six-penny volumes, and John Cooke followed suit with a series of British classics (Altick 1954, 54).

The original issue of David Copperfield totaled a sale of 25,000 books, and 83,000 copies of the penny version of David Copperfield sold in three weeks in 1871. Dickens's publishers had sold

4,239,000 works by 1882 in England alone (Atlick, 1957, 384.) Hardy's *Jude the Obscure* sold 20,000 copies in the first three months (Altick, 1986, 238). Readership also expanded through the reprinting of cultural artifacts. For example, Ryan Cordell and David Smith have developed a website allowing readers to trace the circulation of "viral" texts through reprinting during the nineteenth-century. (See also Matthew Jockers).

A second issue concerns whether authors had any impact on views of norms. Scholars of literature widely agree that literature is crucial to the construction of social problems and norms. Moreover, the cultural impact of nineteenth-century authors seems greater than that of individual writers today, simply because there were fewer authors and they did not have to compete with television and internet in their construction of social problems. Literary scholars emphasize the importance of literature in constructing and revealing shifting views of children. For example, one sees over time a transformation in the depictions of boys, attitudes toward schooling and hegemonic theories about discipline (Parille). Children's literature conveys the values of adult society, norms of appropriateness, and widely accepted cultural standards, and these works are constructed in formulaic writing to reach the broadest audiences (Apol, 62). One also sees cultural artifacts reinforcing shifting norms, as McCallum (2000) points out in her analysis of the shifting views of women and work revealed in successive film versions of *Little Women*. One also sees shifts in depictions of schools. Reed suggests that schools had an anti-child ethos in the early nineteenth-century, as caning and flogging were prevalent, bullying, negative depictions in literature (Reed, 61-2). Yet schools became more favorably portrayed with time; for example, Matthew Arnold depicted a prefect system for social control and this made public schools more gentlemanly. Also rising importance of sports in the public schools and declining importance of classical education, opposition of scholars and sportsmen: students were oriented toward an idea of moral integrity and self-control (Reed 66).

Finally is the question of whether policymakers were directly influenced by authors. It is beyond the scope of this paper to lay out fully this relationship, yet considerable evidence suggests that parallel discussions about education transpired in cultural and political circles and that figures in each realm interacted. In many cases, politicians who wrote the reforms either drew on the texts or were directly influenced by prominent authors.

Ties between Danish authors and political figures are deep and enduring, and writers were very much part of the national project. Holberg was commissioned by the Danish king to produce plays in the Danish language, and his contribution to Danish society and the political economy cannot be over-emphasized (M. Hammerich, "Bidrag til en skildring af Holberg"). Other leading writers, such as Falster, were leaders in the movement to elevate the Danish language, as part of a state-building project.

In the nineteenth century, Georg Brandes (leader of the Modern Breakthrough literary movement) was brother to Edvard Brandes, a famous politician of Venstre (the Liberal Party) who helped to form the Radikale Venstre party. Georg Brandes delivered a famous talk in 1871 at Copenhagen University that criticized the old romantic traditions, gave voice to the new movement, and signaled the beginning of Brandes' quest to modernize Danish literature. The central ideas of the lecture were later embodied in Brandes' work, "Hovedstroemninger i det 19de Aarhundredes Litteratur" (Main Currents in the Literature of the 19th century.) Georg and

Edvard Brandes together formed the leading modern Danish newspaper, *Politikken*.

The Modern Breakthrough movement attacked traditional structures of church and aristocracy, and were deeply influenced by Darwin and Nietzsche. Brandes believed Nietzsche's claim that institutions for governance and morality largely reinforce power relations and maintain social control. Georg Brandes saw literature as central to political struggle and viewed himself as a radical who sought personal, individual liberation. Brandes disliked British Victorian literature as too close to middle-class values, too closely tied to the project of parliamentary democracy, and too respectful of religion and royalty (Skilton, 1980, 37-43). Brandes distrusted parliamentary reforms and believed that economic emancipation was far more vital than political suffrage (Skilton, 1980, 40-2.) Both the Modern Breakthrough authors and the Radikale Venstre had much more benign views toward labor and sought alternative structures to the church and the aristocracy.

SELECTED BIBLIOGRAPHY

Altick, Richard. 1957. *The English Common Reader*. Chicago: University of Chicago Press.

Altick, Richard. 1954. "English Publishing and the Mass Audience in 1852." *Studies in Bibliography* 6, pp. 3-24.

Apol, Laura. "Tamings and Ordeals: Depictions of Female and Male Coming of Age in the West in Turn-of-the-Century Youth's Companion Serials." *The Lion and the Unicorn* 24 (1 January): 61-80.)

Baturo, Alexander and Slava Mikhaylov. 2013. "Life of Brian Revisited: Assessing Informational and Non-Informational Leadership Tools." *Political Science Research and Methods* 1 (1 June): 139 – 157.

Blei, David. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55 (4 April): 77-

Brantlinger, Richard. 1998. *The Reading Lesson*. Bloomington: Indiana University Press.

Buckland, Adelene, Beth Palmer. 2013. *Return to the Common Reader: Print Culture and the Novel 1850-1900*. Ashgate.

Catalinac, Amy. 2015. "From Pork to Policy." *Journal of Politics* 78 (1): 1-18.

Cavallo, Guglielmo and Roger Chartier eds., 2003. *History of Reading in the West*. Boston: University of Massachusetts Press.

Cordell, Ryan. 2015. "Viral Textuality in Nineteenth-Century US Newspaper Exchanges." Vernoica Alfano and Andrew Stauffer eds. *Virtual Victorians: Networks, Connections, Technologies*.

Cordner, Sheila. *Education in Nineteenth-Century British Literature*.

“Editions of Robinson Crusoe in English.” 1936. *The Yale University Library Gazette* 11 (2 October): 19-23.

Feather, John. 1988. *A History of British Publishing*. Psychology Press.

Gilmont, Jean-Francois, 2003. “Protestant Reformations and Reading.” In Cavallo, Guglielmo and Roger Chartier eds., 2003. *History of Reading in the West*. Boston: University of Massachusetts Press.

Hammerich, M. “Bidrag til en skildring af Holberg”

Hopkins, David and Gary King, 2010, “A Method of Automated Nonparametric Content Analysis for Social Science,” *AJPS* 54 (1 Jan): 229-247.

Keen, Paul. 1999. *The Crisis of Literature in the 1790s*. Cambridge University Press.

King, Gary, Jennifer Pan and Margaret Roberts. 2013. “How Censorship in China Allows Government Criticism but Silences Collective Expression.” *American Political Science Review* 107 (2 May): 326-343.

Laver, Michael, Kenneth Benoit and John Garry. 2003. “Extracting Policy Positions from Political Texts Using Words as Data.” *The American Political Science Review* 97 (2 May): 311-331.

McCallum, Robyn. 2000. “The Present Reshaping the Past Reshaping the Present: Film Versions of Little Women.” *The Lion and the Unicorn* 24 (1 January): 81-96.

Miller, Manjari. 2013. *Wronged by Empire*. Stanford University Press.

Nielsen, Richard. 2013. *The Lonely Jihadist*. Harvard Dissertation)

Parille, Ken. 2011. *Boys at Home*. University of Tennessee Press.

John Reed. 1974. “The Public Schools in Victorian Literature.” *Nineteenth-Century Fiction* 29 (1 June): 58-76.)

Skilton, David. 1980. *The Activist Critic. A Symposium on Political Ideas, Literary Methods & International Reception of Georg Brandes*. Hans Christian Hertel, and Sven Moeller Kristensen, eds. Copenhagen: Munksgaard.

Slapin, Jonathan and Sven-Oliver Proksch. 2008. “A Scaling Model for Estimating Time-Series Party Positions from Texts.” *American Journal of Political Science* 52 (3 July): 705–722.

Wittmann, Reinhard. “Was there a Reading Revolution at the End of the Eighteenth Century?” in Guglielmo Cavallo and Roger Chartier eds., *History of Reading in the West*.

DIFFERENCE OF PROPORTION RESULTS

TABLE ONE DIFFERENCE OF PROPORTIONS RESULTS

(FREQUENCIES OF INDIVIDUALISM WORDS IN EDUCATION SNIPPETS)

Period	Danish Proportion	British Proportion	Difference	Z Score	Significance level
1720-1770	0.13	0.2682	0.1382	-15.624	$P < 0.0001$
1770-1820	0.0827	0.2953	0.2126	-29.278	$P < 0.0001$
1820-1870	0.0767	0.245	0.1683	-48.284	$P < 0.0001$
1870-1920	0.1067	0.2113	0.1046	-40.851	$P < 0.0001$

The differences between all periods exceed the critical Z score of 1.96 and are therefore statistically significant.

TABLE TWO DIFFERENCE OF PROPORTIONS RESULTS
(FREQUENCIES OF FEELING WORDS IN EDUCATION SNIPPETS)

Period	Danish Proportion	British Proportion	Difference	Z Score	Significance level
1720-1770	0.7691	1.6434	0.8743	-25.441	$P < 0.0001$
1770-1820	1.0216	1.7464	0.7248	-25.065	$P < 0.0001$
1820-1870	1.0213	1.6548	0.6335	-39.243	$P < 0.0001$
1870-1920	1.139	1.6324	0.4934	-33.227	$P < 0.0001$

The differences between all periods exceed the critical Z score of 1.96 and are therefore statistically significant.

TABLE THREE DIFFERENCE OF PROPORTIONS RESULTS
(FREQUENCIES OF SOCIETY AND COLLECTIVE GOAL WORDS IN SNIPPETS)

Period	Danish Proportion	British Proportion	Difference	Z Score	Significance level
1720-1770	0.8338	0.3907	0.4431	16.557	P < 0.0001
1770-1820	0.9242	0.3787	0.5455	25.899	P < 0.0001
1820-1870	0.8752	0.4252	0.45	37.408	P < 0.0001
1870-1920	0.7122	0.5062	0.2060	25.325	P < 0.0001

The differences between all periods exceed the critical Z score of 1.96 and are therefore statistically significant.

TABLE FOUR DIFFERENCE OF PROPORTIONS RESULTS
(FREQUENCIES OF POLITICAL GOVERNANCE WORDS IN SNIPPETS)

Period	Danish Proportion	British Proportion	Difference	Z Score	Significance level
1720-1770	0.5132	0.4405	0.0727	1.673	P = 0.0942
1770-1820	0.5014	0.4332	0.0682	2.929	P = 0.0034
1820-1870	0.5776	0.3309	0.2467	27.745	P < 0.0001
1870-1920	0.313	0.3517	- 0.0387	-4.492	P < 0.0001

The differences between the last three periods exceed the critical Z score of 1.96 and are therefore statistically significant.

TABLE FIVE DIFFERENCE OF PROPORTIONS RESULTS
(FREQUENCIES OF UPPER CLASS WORDS IN EDUCATION SNIPPETS)

Period	Danish Proportion	British Proportion	Difference	Z Score	Significance level
1720-1770	0.0398	0.6739	- 0.6341	-11.757	$P < 0.0001$
1770-1820	0.4578	0.6233	- 0.1655	-6.603	$P < 0.0001$
1820-1870	0.4407	0.6389	- 0.1982	-14.297	$P < 0.0001$
1870-1920	0.293	0.4093	- 0.1163	-12.830	$P < 0.0001$

The differences between all periods exceed the critical Z score of 1.96 and are therefore statistically significant.