

John Hazelton (Jch5nb@virginia.edu)

DS 5001

December 17, 2021

Introduction

The following analysis explores the works of six different Christian authors prominent across six centuries from the 1500's to present day 2000's, including John Calvin, John Bunyan, Jonathan Edwards, Archibald Alexander, C.S. Lewis, and John Piper. Their works include (in chronological order): *Letters of John Calvin – Volume I*, *The Institutes of the Christian Religion*, *The Pilgrim's Progress*, *The Holy War*, *Religious Affections*, *Select Sermons by Jonathan Edwards*, *Evidences of the Christian Religion*, *Outlines of Moral Science*, *The Problem of Pain*, *The Four Loves*, *Don't Waste Your Life*, and *Fifty Reasons Why Jesus Came to Die*. This report seeks to address the question of how Christian literature has changed over the years from the Protestant Reformation to the present. Specifically, this includes an exploration of potential changes in topics and how they may be tied to social and political happenings, changes in overall message polarity and prevalence of emotions, and changes in how key terms such as “God”, “Christ”, “heaven”, and “sin” are represented.

Analysis of Descriptive Statistics & Distance Measures

In compiling a vocabulary table and some descriptive statistics for each term across the documents, the most prevalent terms (that are also meaningful) are clearly “god”, “christ”, and “lord”. When breaking down vocabulary by author, however, a few other terms appear towards the top of the rankings, including “love”, “spirit”, and “nature” for Jonathan Edwards, “moral”, “reason”, “truth”, and “mind” for Archibald Alexander, and “love”, “glory”, and “joy” for John

Piper. These rankings hint at some of the themes discovered below as a part of principal components and topic modelling. When comparing each of the different works using cosine distance, many of the most-similar pairings are books by the same author, as expected. Ignoring these same-author pairings, a few that stand out as highly similar are Calvin's *Institutes of the Christian Religion* with Edwards' *Religious Affections* and *Select Sermons*, and Edwards' *Select Sermons* with Piper's *Fifty Reasons Why Jesus Came to Die* (Appendix Plot 1). Like these, many of the top pairings surprisingly stretch across multiple centuries, suggesting changes in style and focus of the literature may not be linear, but rather cyclical in nature, with certain themes or topics resurfacing centuries later. This idea is reinforced by the hierarchical clusters generated using the same cosine distance metric (Appendix Plot 2). Books are nearly grouped in pairs by author, as expected, but are not grouped in chronological order, indicating an inconsistent change in similarity over time. For example, in the second level of clusters, John Piper's two books, written in the early 2000's, are grouped with John Calvin's *Institutes of the Christian Religion*, written in 1541. Furthermore, Piper's works are situated furthest from C.S. Lewis', which were released no more than 50 years prior. The differences between these authors, and their relation to time periods, are more thoroughly explored through topic modelling and word embeddings below.

Analysis of Topics & Word Analogies

Conducting principal component analysis (PCA) on these documents, it becomes clear the first two components are significant in identifying the main differences between books, as together they explain nearly 55% of the overall variance in the data (Appendix Plot 3). Visualizing these components, it is clear much of the separation between books is due to Calvin's *The Institutes of the Christian Religion* being drastically different than the others on this

feature (Appendix Plot 4). Removing this data point, groupings become clearer for the remaining documents, with Edwards having separation from the others, while Lewis and Alexander's works seem to form a cluster (Appendix Plot 5). In plotting components two and three, Calvin's and Edwards' books separate from the rest on either ends of the graph (Appendix plot 6). Looking at the positive and negative loadings for these components, the biggest difference in Calvin's (and some of Edwards') 1500's–1600's works becomes clear in key terms such as “catholic”, “protestant”, “persecutions”, and “anabaptist” (meaning member of the radical Protestant Reformation) (Appendix Plot 7). These components are likely representative of a focus on the radical movement away from the Catholic Church at the time, and related persecutions of those individuals. This literature is a stark contrast from works written in centuries following, which likely focus more on fleshing out the meanings and beliefs behind Protestant religion, as opposed to the social and political events that surrounded its forming in the 1500's.

Using Linear Discriminant Analysis (LDA) to look more specifically at topics differentiating these documents, a few things become clear. Calvin's most prominent topics don't seem to be largely shared with any of the other authors, further supporting the idea that Calvin's works are most different because they occurred at the time of the Protestant Reformation and focused on the social and political happenings surrounding that movement (Appendix Plot 8, representative of how most of the topic plots between Calvin and other authors look). Alexander and Edwards share the most common topic between one another's works, which consists of terms such as “affections”, “evidence”, “power”, and “mind” (Appendix Plot 9). C.S. Lewis shares a similar topic with both of them, consisting of many of the same terms (Appendix Plot 10). “Affections”, used as a term to define how individuals are converted to Christianity through the power of the Holy Spirit, seems to largely categorize the focus of these topics and authors.

These works all seem to be centered around both conversion to Christianity and apologetics/defense of Christianity. This conclusion makes sense, as Edwards' *Religious Affections* talks about how emotion and intellect (hence "evidence" and "mind"), together with "affections" play a role in how individuals come to find Christianity. This book was released at the time of the First Great Awakening in the early 1700's, followed by the works of Alexander and Lewis in the 1800's and 1900's, when there was an explosion of Christian apologetics writing.

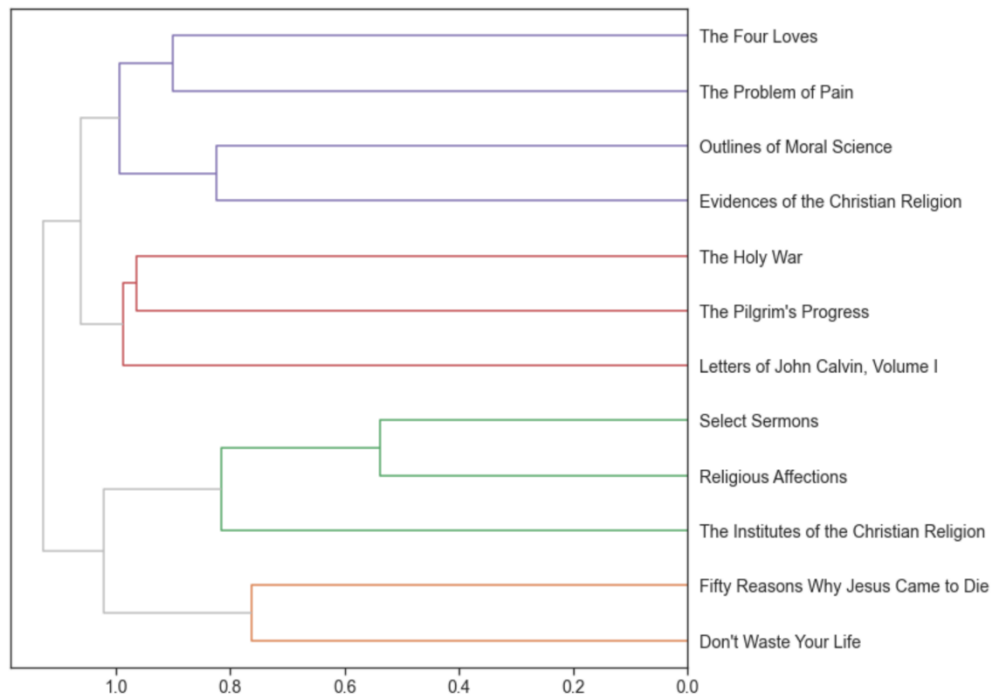
Looking at word embeddings for the various authors, we see that "god" and "christ" are closely related for most of the authors, along with terms such as "great" and "holy". Looking at the differences in analogies, though, Calvin equates "god" and "christ" to words like "declares" and "promise", whereas Edwards and Alexander equate them to terms like "apostle", "reason", and "testimony". Apostles being essentially the first Christian missionaries, this, along with "reason" and "testimony" further emphasizes the shift in the 1700's-1800's towards Christian evangelism, using reason and logic, along with testimonials as support and evidence for the religion. Looking at analogies for "sin", Calvin equates it with terms like "flesh", used commonly as a term to separate sin from the sinner's true heart or soul. He also equates it with "salvation". Alternatively, more contemporary authors Lewis and Piper don't seem to make this distinction or separation from the sinner, equating "sin" with "man" and "people". Piper does, however, make the similar connection to salvation, with the related term "heaven". Overall, the analogies for these authors suggest many of the same terms are related to "god", "christ", and "sin", but the subtle differences indicate a few nuanced ways they present these concepts.

Appendix

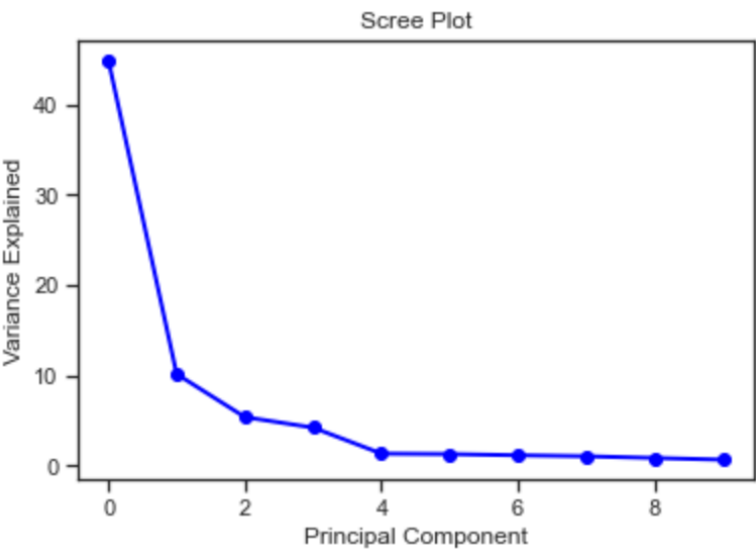
Plot 1. Pair-wise Document Distance (ordered by cosine)

		cityblock	euclidean	cosine	jaccard	correlation
doc_a	doc_b					
4	5	0.132357	0.002721	0.537538	0.561972	0.574346
1	5	0.213206	0.003640	0.745458	0.711566	0.817403
10	11	0.244392	0.005949	0.763034	0.792377	0.787753
1	4	0.221580	0.003914	0.766940	0.731581	0.830438
5	11	0.182118	0.004678	0.776476	0.822142	0.805607
1	6	0.261077	0.004737	0.822717	0.744549	0.896015
6	7	0.227006	0.005827	0.824539	0.725892	0.858146
0	1	0.298683	0.006592	0.826600	0.681647	0.888541
5	6	0.212825	0.004233	0.836858	0.668261	0.898447
4	6	0.216859	0.004450	0.837199	0.679465	0.890619
	11	0.192667	0.005063	0.862458	0.832768	0.891826

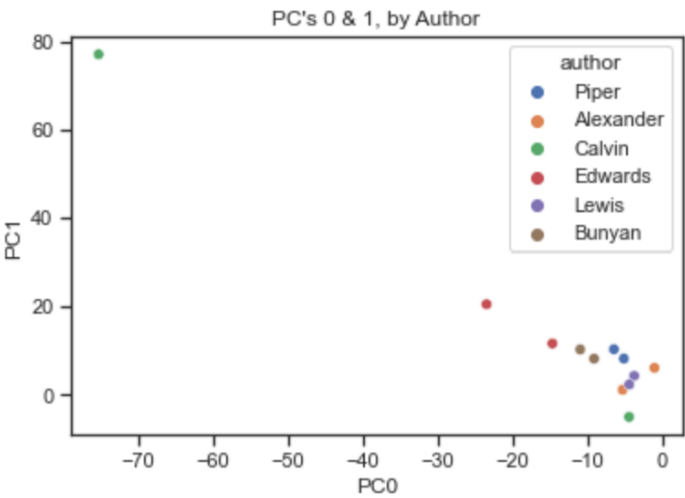
Plot 2. Hierarchical Clusters by Document Distance (cosine)



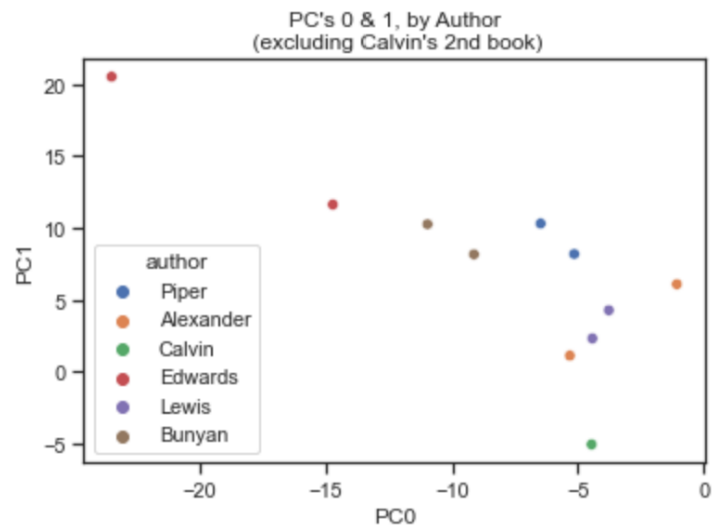
3. Scree Plot: Explained Variance by Principal Component



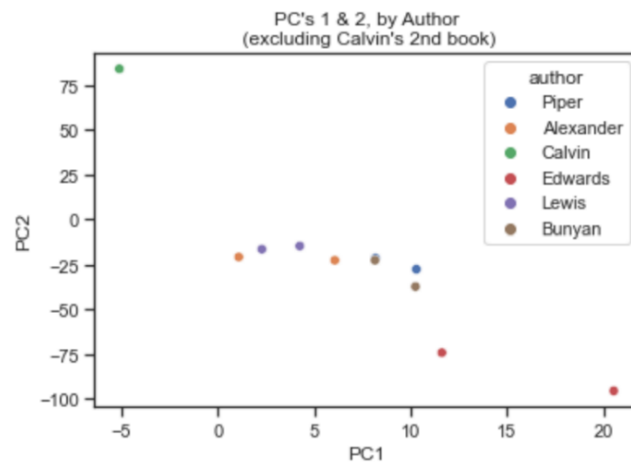
4. PC 0 vs PC 1, by Author



5. PC 0 vs PC 1, by Author (excluding Calvin's 2nd book)



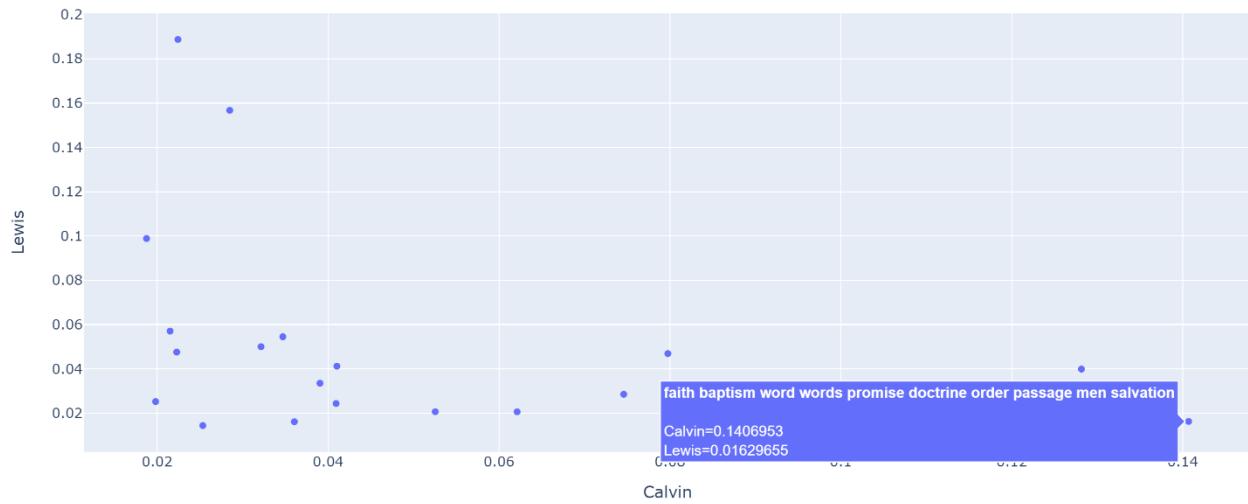
6. PC 1 vs PC 2, by Author (excluding Calvin's 2nd book)



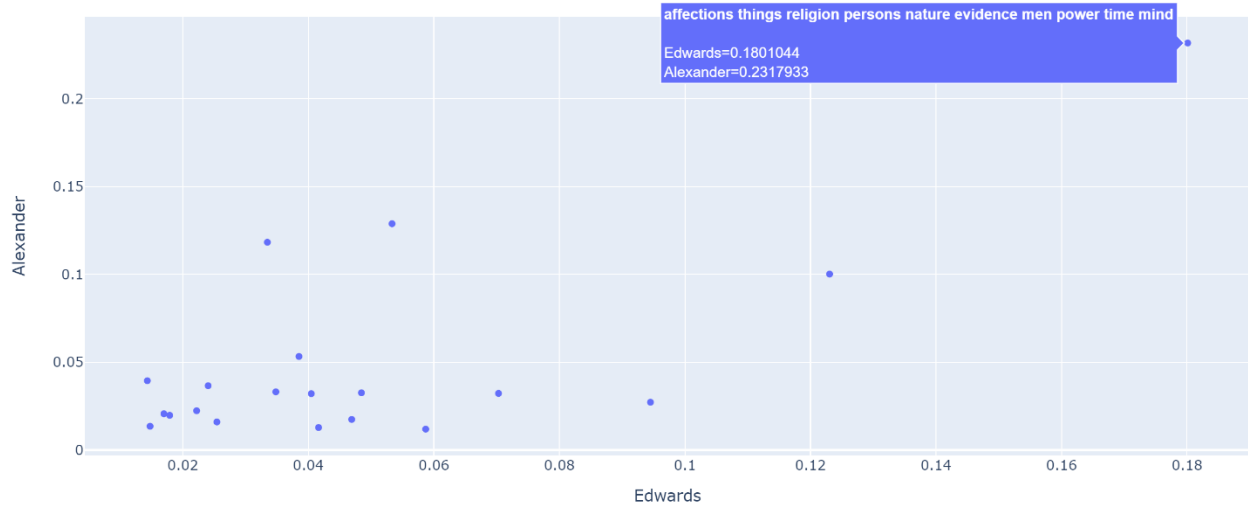
7. Positive & Negative Loadings for PC's 0, 1, and 2

```
Books PC0+ christ god moral anabaptist hermann jesus church letter bible law
Books PC0- persecutions policy frankfort protestants conferences catholic opening roman property assembly
Books PC1+ anabaptist hermann v discussion understanding mention good persecutions policy intellectual
Books PC1- christ god frankfort alexandre moral jesus germany lecture princes negotiations
Books PC2+ persecutions christ moral god jesus bible sin romans law faculty
Books PC2- alexandre negotiations lecture henry details protestant germany reconciliation princes bucer
```

8. Topics: Calvin vs Lewis



9. Topics: Alexander vs Edwards



10. Topics: Edwards vs Lewis

