The (Protestant) Bible, the (printed) sermon, and the word(s): The semantic structure of the Conformist and Dissenting Bible, 1660–1780☆

Mark Anthony Hoffman[a], Jean-Philippe Cointet[b], Philipp Brandt[c], Newton Key[d], Peter Bearman[a],*

[a] Department of Sociology, Columbia University, United States
[b] Sciences Po, médialab, Paris, France
[c] Department of Sociology, University of Mannheim, Germany
[d] Department of History, Eastern Illinois University, United States

ARTICLE INFO

ABSTRACT

Using co-occurrence methods for identifying semantic structure in texts, we first describe the structure of the Protestant Bible, focusing on the ways in which contents of the Bible are organized in both the New and Old Testaments. We introduce a strategy for capturing the co-occurrence of nouns and verbs in windows defined by verses that progressively move across the text, from start to finish in a manner similar to reading. We then consider how Dissenters and Conformists used the Bible by locating Biblical verse in sermons printed in England during the period from 1660 to 1780. We describe how chapters are linked by themes over time, by dissenting and conformist religious communities, and map Dissenter and Conformist uses of the Bible onto its semantic structure. We show that it is possible to induce a semantic network image of the Bible, that this structure serves as a skeletal frame for interpretation, thereby highlighting different contents as central to denominations' religious inspirations and concerns.

1. Introduction

Critical commentary on the Bible has led to widely differing religious, social, and political movements. Analyses of Biblical verses have energized competing intellectual traditions. And translations of the original Hebrew, Greek, and Aramaic have shaped various vernacular interpretations. Heterodox readings or interpretations of the otherwise shared Old and New Testaments have also led to lengthy and bitter disputes, separations (or migrations), and imprisonment and slaughter of those found to interpret Biblical meaning different from that accepted by the majority or by those in power.

Yet most studies of disputed Biblical commentary focus on this or that interpretation—how do groups differ on the *adiaphora* (things indifferent to salvation), the proper meaning of the rite of communion or order of the Mass, the larger implications of terms such as *ecclesia* (church, congregation) or *presbyteros* (elder, minister) in the original Greek. To our knowledge, no one has tackled the

Bible's semantic structure as a whole. This is due in part to the complexity of the task (although there are certainly numerous works which provide a book-by-book guide to the Bible). As a result, commentators are unable to reveal, in any systematic manner, how the semantic structure of the Bible links to the structure and content of theological dispute.

This article offers a method to identify the semantic structure of the Bible. It considers how concepts, for example vices and virtues, are linked. It maps out the interwoven structure of the Old and New Testaments. It then considers how the various nouns and verbs in the Bible are related to one another *as it was used*. To do this, we consider the Bible as expounded in sermons printed in English, specifically by comparing those delivered by Protestant preachers identifiable as Anglican or Nonconformist (Dissenters) between 1660 and 1780, that is over the course of the long eighteenth century from the Restoration of Charles II to the eve of both American independence and an industrial revolution based on cotton, coal and iron.

The meaning of the Bible has been a serious business conducted by serious people with serious stakes in the outcome. Uncovering the semantic structure of the Bible is a very different project than the religious project of revealing the Bible's ultimate meaning. The distinction between ultimate meaning and the structure of conversation is not trivial. The fact that countless numbers of people have been killed, had their careers cut short, their opportunities blocked, their properties seized, and their children murdered in the name of the truths believed revealed within the Bible is perhaps sufficient reason to insist that we are not trying to ascertain ultimate meaning. We are instead first interested in structure, and secondly in how actors navigate the structure of texts to communicate and contest critical issues of identity and belief. In this regard, then, structure, as we will show, is the skeleton on which are hung a wide variety of answers to questions of meaning and action.

1.1. Roadmap

We first discuss the historical context in which the sermons we rely on were written and delivered. We then discuss our data and methods. We then identify the semantic structure of the Bible, inducing a partition that neatly disentangles (in a semantic network perspective) the Old and New Testament. We then identify the communities of this semantic structure, focusing explicitly on the cluster of concepts we associate with sins and virtues. We refer to these communities as topics. This viewpoint provides one kind of mapping of the Bible, a mapping based entirely on the pattern of co-occurrence of nouns and verbs found across a moving window of verses in the Bible, from Genesis to Revelation, much as someone reads. We then turn to the Bible in action. By drawing on the explicit linkage between verse and sermon theme for over 20,000 sermons delivered in England (by Dissenters and Conformists) from 1660 to 1780, we induce a dual mapping of Biblical chapters tied by shared themes, and themes tied by Biblical chapters. We then embed the theme structures we observe into the semantic space defined by our first analysis, first for Conformists and then for Dissenters. This enables us to observe Dissenting and Conformist semantic network positions, and sheds light on the critical religious issues that distinguished each party.

We thus make two independent contributions. First, we induce a semantic structure for the Bible. Second, we show how this structure provides a skeleton for revealing religious movements in action. The latter is a historical contribution − providing a new strategy for thinking about what mattered to Dissenters and Conformists. The former is a structural contribution − showing how the analysis of structure can reveal new meaning in arguably one of the most studied texts in the world − the Protestant Bible. Linking these two foci is the idea that how things are used reveals both their content and their structure, and provides a mapping for understanding contestation amongst real groups of people (Bearman, 1997; Bearman, Faris, & Moody 2009; Sewell, 1989).

2. Historical context for preaching in the long eighteenth century

England and Wales saw the Restoration of both Church and State. Monarchy was restored through the accession of Charles II, whose father Charles I had been executed in 1649 during what is sometimes termed the Puritan Revolution of the 1640 s and 1650s. The Anglican Church of England, which had basically ceased to exist when the bishops had been abolished in the 1640s, returned from 1660 onwards. Soon bishops were in each diocese, and Anglicans sought ways to eject Presbyterian, Independent (Congregational), and even Baptist ministers who had been "intruded" during the said Puritan Revolution. According to the Act of Uniformity, by Bartholomew Day (24 August) 1662 all ministers were to subscribe to the 39 Articles of the Church of England and to hold services according to the newly revived Book of Common Prayer. If they refused to do so, they were to resign or be ejected. Between 1660 and 1662 nearly 2000 of the 9000+ ministers in England and Wales were ejected and replaced (Appleby, 2014).

Between the Restoration (1660) and the Revolution of 1688-89, harsh laws (those known as the Clarendon Code, as well as the Test Act, etc.) attempted to further silence the ejected ministers, a situation relieved only briefly when Charles II in 1672 and then James II in 1687 and 1688 issued Declarations of Indulgence which allowed in effect limited freedom of worship for both Protestant Nonconformists and Catholics, although in both cases, most feared that it was the latter that they most sought and both periods of "indulgence" were overturned—one by the Test Act of 1673 and the other by the Glorious Revolution (Key, 1990, 1994, 2014). The Test Act of 1673 legislated that all government officers had to take communion (according to the Church of England) and an oath against transubstantiation. After the Revolution of 1688-89, under the new monarchs William and Mary, a limited Toleration Act of 1689 allowed all Trinitarian Protestants (but not Unitarians) to worship in peace. The Test Act regarding office holders, however, remained in force, and was not fully repealed until 1828. The combination of the Test Act and the limited toleration after 1689 did create a gray zone for some Dissenters (mainly Presbyterians, or those earlier termed Puritan) who remained closeted, occasionally conforming, or, as their opponents feared, as hidden Dissenters (Haydon, 2002).

Against this background, throughout this early period and beyond, different religious groups pursued varied strategies in response to state repression. Some moderate Dissenters were reluctant nonconformists, and hoped that efforts towards "comprehension," or
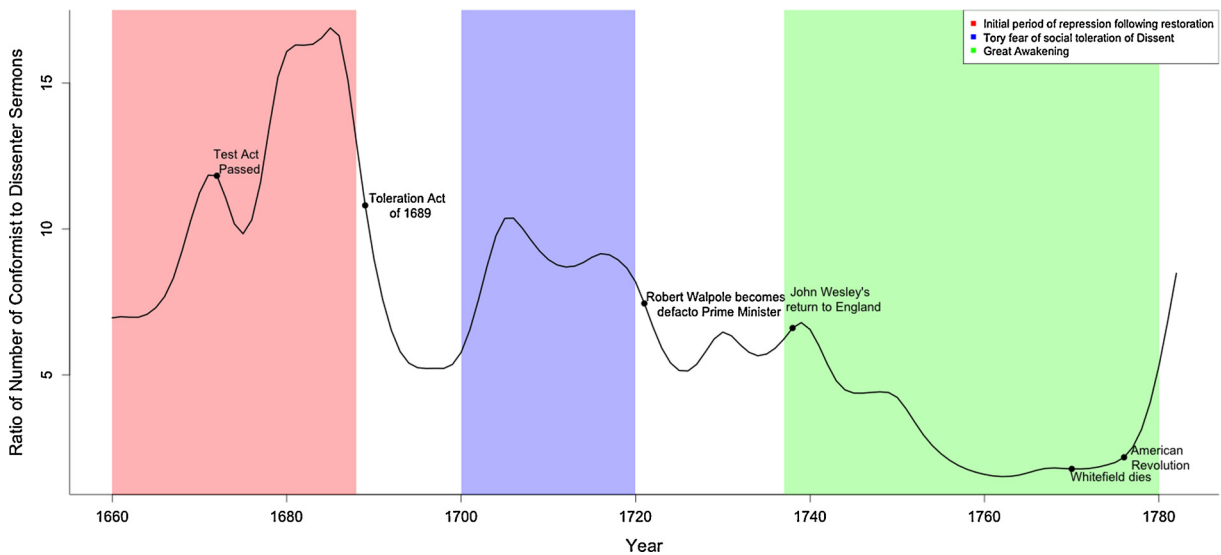
Fig. 1. Turning points in religious repression track landmark events in English religious history.

making the Church of England's bounds less strict, would bring them back into the fold once more. Others, and this group grew as efforts at comprehension stalled, embraced nonconformity. That is, they no longer attempted to be Puritan within the established church, but to separate, and identify openly as English Presbyterians or Congregationalists. A third group, Quakers and most Baptists (including the various Anabaptists, which would today be termed Mennonites) portrayed and thought of themselves as separate and distinct from the beginning. Later, a new Protestant sect, Methodism, would grow out of an attempt to re-energize the Conformist or Anglican Church of England. Only towards the end of the period considered could Methodism, sometimes termed the "New Dissenters," be considered Nonconformist.

Fig. 1 plots the ratio of Nonconformist to Conformist sermons published over time in our sample. Relatively stable ratios of Conformist (5 times as many) to Dissenting sermons published over the long 18th century indicate that the Church (Anglican) and Chapel (Nonconformist) categories were a permanent part of English society. Peak repression is indexed by energetic enforcement of the penal laws after the Test Act of 1673; the fall in the ratio after 1689 indexes the softening of this repression. From 1689 onwards, Nonconformist meetings were no longer proscribed officially, and consequently, the number of Dissenter sermons printed rose relative to that of Conformist sermons (Claydon, 2011). The Tory perception in the 1710s that the Anglican Church was in danger because of social toleration of Dissent resulted in a large number of Anglican sermons on this subject, increased repression, and consequently a corresponding decline in printed Nonconformist sermons. But this too was short lived, and by mid-century, associated with the "Great Awakening", there were new attempts associated with the ministries of George Whitfield and John Wesley to proselytize a more inclusive message and, for Wesley, a less Calvinist one, to a wider public.[1]

For our purposes, the key feature of the religious history of the period is that Dissenters and Conformists differed on issues deeper than the reins of power and that these profound differences between Conformists and Dissenters on fundamental religious issues reflect their quite different constituents. A plurality of English Dissenters in the long eighteenth century arose from specific occupational groups decoupled from traditional landed or agricultural communities (individuals in such communities, at either end of the social ladder, were unreceptive to evangelical messages) located initially in the small and medium sized craft-based towns, some of which would later appear as industrial centers (Crawford, 1987). Separated from the traditional dependency and surveillance relations that bound their rural neighbors into chains of mutual obligation, early Dissenters arguably emotionally were disposed to messages that stressed both frugality, industriousness, creditworthiness and the internal governance of behavior broadly construed and embedding into a community of conscious faith (Hughes, 2013). There was thus an affinity between the social position of Dissenters and religious experience. Separated from the organic cycle of life, early Dissenters were deeply concerned with death and damnation, sin and living the virtuous life.

Their obsession with death led them to embrace entirely new strategies for assessing the likelihood of their salvation. This embrace took both enthusiastic and sober forms, from the "religion of the heart" espoused by Methodists to the conservative rationalism of Presbyterians. Whichever form, enthusiastic or sober, they took, Dissenters − relative to Anglicans − emphasized sin, forgiveness, and shared a belief in re-birth tied to the capacity to directly experience God. With respect to living the Christian life, dissenters focused on their hostility to leisure time, and activities associated with leisure, especially drinking and gambling, but even dancing and sports. Dissenters were consequently more likely to see any form of secular "amusement" to be sinful (Francis, 2012).

Of course, Conformists were Protestants as well. Relative to Catholic orthodoxy both Anglicans and Dissenters were heterodox,

---

[1] Formal Methodism outside Anglicanism dates from the 1780s. But Wesley was preaching much earlier, from 1737. Our source for identifying Dissenter sermons (discussed below) delineates many of these earlier Methodist-Anglican sermons as Dissenter, and we follow that decision.

unbelieving of transubstantiation and emphasizing the parishioner's individual relationship with God. But the concerns of Conformist/Anglican ministers and congregations tended to be the concerns of the ruling class, often of the party in political power: suppressing religious and social rebellion, emphasizing the "natural order" of things, and the rules of comportment (manners) which sustained that order. All of these tendencies we observe in the different use of the Bible by Dissenting and Conforming ministers.

Three points stand out in this brief contextual narrative. First, Conformity as well as Nonconformity was constructed. There were conforming Puritans that moved to Nonconformity sometime after 1662. And, there were Puritans that remained conformists. Methodists began as part of an Anglican revival, and then became Nonconformists. Second, Conformists—responsible for most of the printed sermons considered below—were never completely unified, at times disagreeing over the nature of the Church and over the meaning of Anglican theology. During the long eighteenth century Anglicans and Anglican priests could be divided politically between Whigs and Tories. One would not predict that Conformist/Nonconformist to be the only or even the most salient ideological divide. Third, despite penal laws and the Test Act, English Protestants had a degree of choice over their religious affiliation during the long eighteenth century, at least after 1689. We have made the assumption that the sermons they heard reflected to some degree the beliefs of those who listened to (or in this case, read) them. This paper does not prove that assumption, but we have provided our rationale for thinking this.

2.1. The Protestant Bible and the printed sermon

For Christians, Scripture is the rule of faith. Protestant reformers rejected Roman Catholic tempering of Scripture with tradition. The word of God in the Bible was necessary and sufficient. Preaching was seen as necessary because it was the best way to make known "the sacred authority of the Scripture" (Hooker, 1875) and to apply its teaching to the auditory's understanding that they might find the means necessary for salvation. Indeed, the 19th Article in the Anglican Book of Common Prayer defines the "Church of Christ" as "a congregation…in the which the pure Word of God is preached." For Protestant English sermons between 1660 and 1780, that Word of God was found in the 39 canonical books of the Old and New Testaments, using the authorized or King James version.[2]

It is important to note, nevertheless, that preaching necessarily involves a double translation or interpretation of the word of God. First, the verse chosen would be read, not in the original language, but in the vernacular, in the case of our *corpus* English, and according to various standard versions. Second, the preacher would invariably direct his listeners from the text, which might have to do with Pharisees or pharaohs or otherwise arcane subjects known only in the ancient world, to its "use" or application in contemporary life, in this case that of 17th and 18th-century English men and women. This translation is what makes sermons important. It is through the sermon that the Bible as the word of God becomes activated for a congregation.

Thousands of sermons were given each week and only some small proportion of those delivered were printed (Farooq, 2013; Ferrell & MacCullough, 2000). That said, the printed sermons were significant; the market for them was very large (they were "by far the biggest literary genre of that period" Tennant, 2015); and many more people read sermons than heard them originally. The sermon in print served in part the role that radio or television preaching does now. And they set the standard. As Gilbert Burnet (an Anglican bishop, best-selling theologian, and renowned preacher–though one who claimed to preach from loose notes only), in 1692 noted "The Nation has got into so good a Taste of Sermons, from the vast Number of those excellent ones that are in Print, that a mean Composition will be ill heard." (McCullough, Adlington, & Rhatigan, 2011; Burnet, 1766).

3. Data and methods

Our sermon data arise from the John Gordon Spaulding's modernized and digitized version of *The Preacher's Assistant*, an index of all sermons published from the Restoration (1660) to 1782, originally constructed by Sampson Letsome up to 1753, later extended to 1782 by John Cooke and enhanced by Spaulding (Spaulding & Cooke, 1988). *The Preacher's Assistant* (Hereafter, PA) was designed, as the title suggests, to help preachers craft sermons. Fig. 2 provides an example of how the index was intended to work: the aspiring preacher considers possible sources of text, or themes, and searches for the sermons that correspond to the element that he is seeking. If preachers were interested in discussing, say, Romans 6:23 — "For the wages of sin is death; but the gift of God is eternal life through Jesus Christ our Lord." — they could use a sermon index to identify the themes, and locate printed texts, of dozens of sermons printed between 1660 and 1782 that discuss the verse. Alternatively, if they were interested in the theme of "Absolution", they could find all of the verses or printed sermons, which relate to that theme (although that search would involve thumbing through hundreds of pages).

From such an index, preachers and theologians could gather the contemporary interpretation of a given verse and identify sources for further research. These tools, which made the work of a pastor less dependent on detailed readings of the Bible, were criticized in the 17th and 18th centuries in the same way that Wikipedia, and more generally the Internet, are criticized today, as enablers of lazy, uninvolved scholarship. For example, French theologian François Fénelon's posthumously published *Dialogues on [Pulpit] Eloquence* notes: "All this puts me in mind of a preacher, a friend of mine, who lives, as you have it, from day to day. He does not ponder

---

[2] Occasionally, 18th-century preachers used other versions or translated anew from the Hebrew or Greek themselves. But the differences between those verses quoted and the King James version were slight, and have been ignored here. Moreover, many of the natural language processing scripts used to analyze the Bible and construct it as a semantic network were designed with 20th and 21st century English in mind. As a result, most modern parts-of-speech taggers have a difficult time identifying archaic words and verb forms. In order to resolve the problems produced by using modern scripts on archaic language, we modernized troublesome parts of the Bible, verbs and pronouns in particular. As an example, we converted "speaketh" to "speakest". We assume that no loss of meaning occurred in doing so. Problematic pronouns, including thee, thou and thine, would be ignored by our method even in modern form, so we dropped them from the text before modernization.
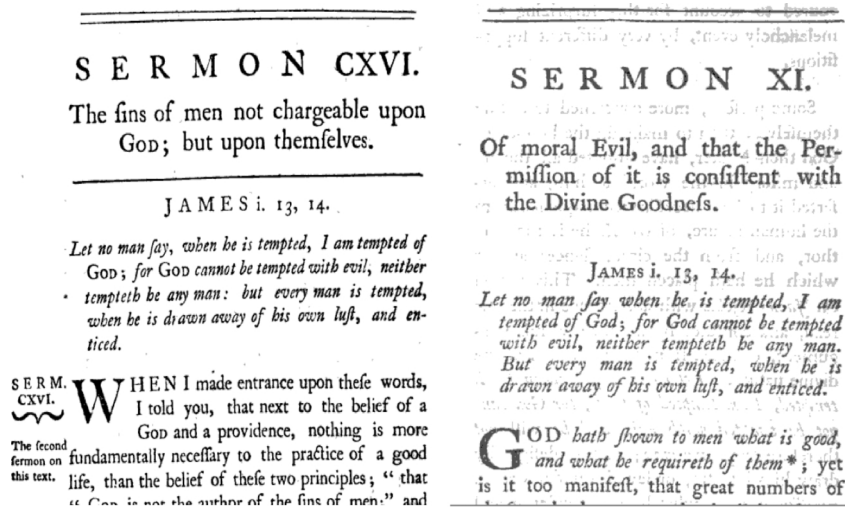
Fig. 2. Two sermons drawing from the same text, linked to different – but related – topics.

anything until he is scheduled to preach upon it. Then he closes himself in his study, thumbs his concordance … and some sermon books he has bought, and various collections he has made of purple patches wrested from their context and hit upon by good luck" (1722 translation, quoted in Deconinck-Brossard, 2009).

This would be most helpful when matched with the prescribed Old and New Testament verses for that week and service as outlined in the official lectionary − calendars for Christian worship, which appoint for each day or week of the year specific Scripture verses to be read and discussed. But it could also be used for occasional or state sermons before specific bodies or meetings, where verse choice was more open. In this case, the "lazy" or "busy" preacher could think of a theme − say "the righteousness of authority" − and identify with the index a sermon on that theme and the associated text from the Bible used to support it. Because each line of the index accords with a sermon and includes the Biblical book, chapter, verse reference, the name of the preacher and dignity (D.D., Bp., etc.—by which we have been able to label, following Spaulding and the compilers, preachers Conformist or Nonconformist), date published and format (4to, 8vo, etc.), whether a single sermon or part of a collection, and an abbreviated reference to the occasion or subject, the preacher using the PA can also align himself with specific preachers or interpretations. The result is a dataset which connects preachers to the themes and verses of their sermons over a 160-year period in English history.[3] It contains over 24,000 sermons, 5000 preachers, 6000 unique themes and 7000 verses.

Thousands of sermons were delivered on the specific days, and preachers did not have free reign over what verse of the Bible they would use to as text, those verses were already dictated by the lectionary for the year. Thus, for example, sermons given on the 3rd Sunday after Epiphany were in part dictated by a lectionary, such as the Book of Common Prayer. Included was an Old Testament reading, one from a Psalm, one from an Epistle, and one from a Gospel. The preacher would select one of those upon which to preach. Occasional or state sermon preachers had more leeway on which verse to choose. The selection of the verse itself limited what part of the word of God would be the focus of the preacher's message. The whole Bible was the inspired word of God, but obviously, only specific parts of it could be the subject of any sermon's message.

Most Biblical verses were never referred to in sermons, at least not printed sermons. The Bible contains 31,000 verses so less than one-quarter of the verses in the Bible are featured in the PA. To give one example, 2 Kings chapter 14 (with 29 verses), the PA only includes a few sermons. Verses 8 and 9 are mentioned only once: in Richard Arnald's citation of "the thistle that was in Lebanon sent to the cedar that was in Lebanon," for a thanksgiving sermon for victory over the Jacobites in Scotland (this citation is presumably inspired by the thistle!). Elsewhere, the PA might list dozens of sermons. More than 40 sermons are cited for Romans 13, 1 "Let every soul be subject unto the higher powers.".

Against this background, one thing we can assess is whether or not dissenters and conformists referred to the same texts as each other, or whether they drew from different parts of the Bible in constructing their sermons. Fig. 3 reports the results of such an analysis focusing on chapters of the Bible. Conformists and Non-Conformists obviously draw from different texts.

Compilers Letsome and Cooke assigned themes to many of the sermons. These sermon themes offer the best resource for identifying the subject and purpose of sermons in England during this period. Further, since Letsome and Cooke were contemporary

---

[3] The Letsome index serves as a representation of theological discourse following the Restoration (Letsome 1753; expanded by Cooke 1783; and now in Spaulding, 1996(Spaulding & Cooke, 1988)). A reviewer in 1754 (Anonymous, 1754) noted that Letsome had assembled many of the sermons listed, and, indeed, dozens of volumes bound as the Letsome Sermons reside in the British Library today. A newspaper article commenting on the original publications specified its coverage of published sermons at 95% and there is some supporting evidence today that this estimate is correct: searching for "sermon" in the title for books 1660–1780 in the most complete listing of printed works of the period, the current English Short Title Catalogue (http://estc.bl.uk) locates 17,507 items. It is fair to state that the amended Letsome is the most complete list of printed sermons in this period that can be obtained at present. And Tennant, "a former associate of Spaulding," is developing "Spaulding +" extending it to 1815 (Tennant, 2015).
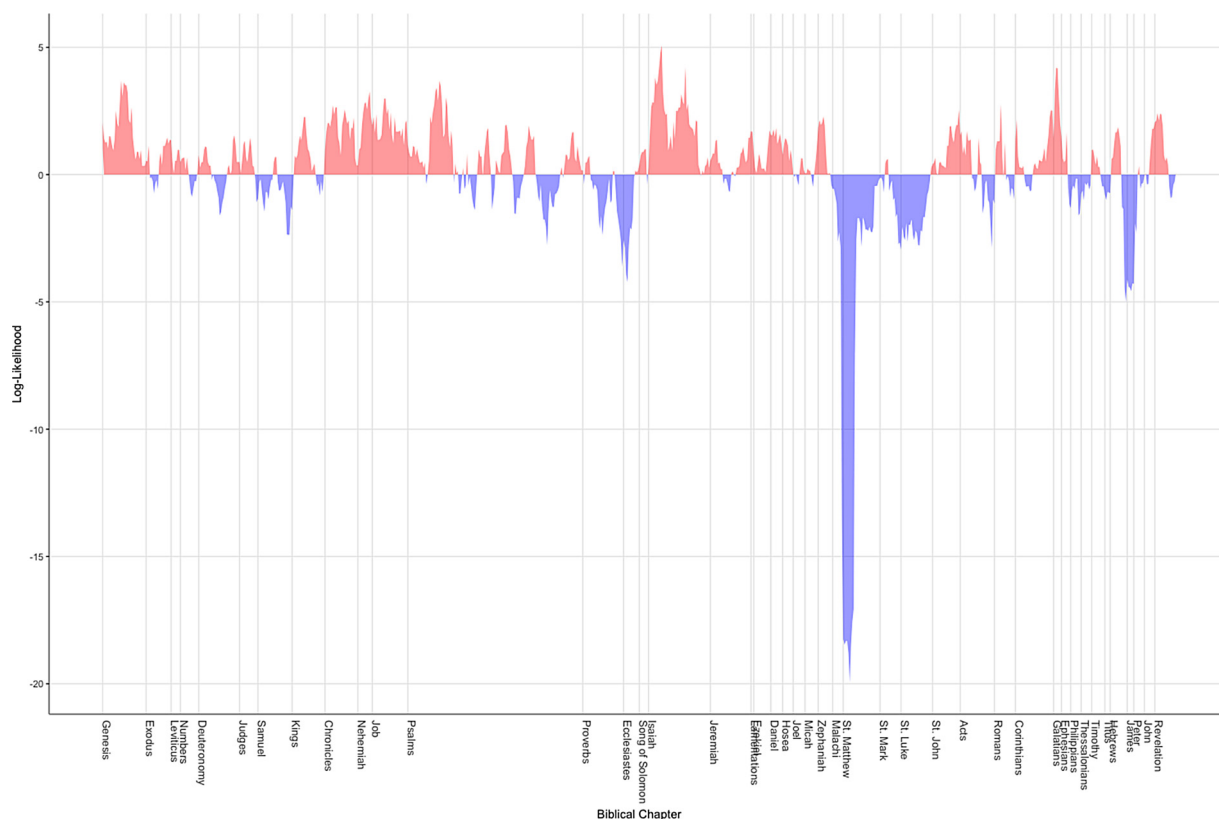
Fig. 3. Chapter use by Dissenters and Conformists.

observers, versed in the salient religious themes and debates of 18th century England, the sermon themes are congruous with the concerns of their time.

Sermon themes appear to relate to national and local events, often in what would seem to be a haphazard way. Chamberlain (2012:51-52) reports on the discovery of a preacher named Samual Rand whose diary describes the motivations for particular sermons he gave on specific days. One (13 Hebrew 4) arises from concern with "fornication and adultery occasioned by the great lewdness of the kind in my parishes"; another (26 Mathew 40) against "sleeping at church" (perhaps the previous sermons were not that inspiring), and so on.

Of the sermons in PA, roughly 21% were not assigned a theme. It is unclear why no theme was assigned; however, these sermons are evenly distributed across the period, so it doesn't appear to be due to a loss of information about sermons' content over time. Of the remaining 79% of the sermons, 48% were given to mark an occasion or holiday; for example, the date of Charles II's ascension to the throne, a sermon delivered to Parliament on the occasion of its opening, the consecration of a bishop or Easter Sunday. The remaining sermons, making up 52% of the sermons with themes and 40% of sermons overall, are general sermons that refer to Christian life, strategies for living the Christian life, sins to avoid, virtues to inculcate.

Importantly, even when they draw from the same chapter in constructing their sermons (OR 1.19). From this observation, and from Fig. 2 there is thus strong evidence that the two communities differ with respect to their use of the Bible.

The PA is thought to have been particularly important to dissenters, since roughly 20% of the sermons indexed are non-conforming (Gibson, 2012). The PA, even if it includes almost all of the sermons published from 1660 to 1782, of course, covers only miniscule fraction of the number of sermons read over this time period (Dixon, 2011; Francis, 2012).[4]

3.1. Text analysis

Theologians have long sought to reveal both the semantic and practical structures of the Bible. The first efforts towards understanding its semantic structure were made by Hugh of St. Cher, a French friar of the Dominican Order, who developed the first verbal index of the Bible in 1230 CE. Verbal indices, now commonly referred to as concordances, list in three columns: each word in the Bible, the verses where a word appears, and sometimes, the clause or sentence in which a word is embedded. Scholars researching a

---

[4] Given that there were more than 9000 parishes in England, that each parish had at least one sermon given every week (more often two or three), and the Preacher's Assistant covers more than 100 years, there were more than 100 million sermons read, of which 25,000 is just 0.00025 (%?) of the total. Gibson has made a different estimate of the number of sermons delivered and sermons printed, but the disparity between the two populations remains (Gibson 2012).

specific character, concept or place, either while writing an essay or preparing a sermon, would search for a word in the index and find a complete list of the locations in the Bible where that word could be found. Concordances were therefore a deconstruction of the Bible, from its familiar linear narrative form to a structure oriented around word occurrence, serving as a tool that allowed interested parties to locate words in the text. But like any index, by removing the word from the population of words to which it is linked, removes an important source of its meaning. In constructing the semantic structure of the Bible, we apply network analytic techniques to the Bible using co-occurrence methods. We save both the meaning that arises from context − the specific set of words in a sentence or verse to which a focal word is attached − and from co-occurrence across the text. The latter is not precisely accessible to the reader when reading. Consider for example, that we can for any word (say harlotry) model the changing community of words around it at any moment in the text. Of course, one could do this by hand for some words; the key to the concurrence strategy we develop is that occurs simultaneously, for all words, at all moments.

Our analysis strategy rests on a co-occurrence approach that relies on terms' joint appearance over a particular unit of text to induce topics or categories (Rule, Cointet, & Bearman, 2015). The goal is to endogenously identify interpretable higher-level units of meaning as communities − that is clusters of words and concepts that capture a topic or set of interrelated topics. This intent is common to other text analysis approaches that also rely on the count of co-occurrences in a text, for example correspondence analysis (Benzécri, 1973) and topic modeling (Blei, Ng, & Jordan, 2003). In topic modeling, generally, the meaning of text (at a higher level) is inferred from probabilistic models based on observed word distributions. For a given text − in this case, the Bible − the analyst defines a number of topics that are then distinguished by a probability distribution over words specific to each topic. Topics consist of a ranked list of terms whose weights are associated with their likelihood of being used when a topic is drawn. Similarly, Correspondence Analysis (CA) − in line with more recent word embedding approaches − assigns each word a position in a low-dimension space that can then be used to visualize clusters of close terms.

Our co-occurrence method directly solves some problems associated with text analysis that require additional steps to resolve in a topic model framework.[5] The most straightforward of these is that in topic modeling generally, a second step (namely a visualization of the co-occurrence network) is required to see the underlying structure. In the co-occurrence approach we adopt here, the structure of the corpora is revealed directly, and the visualization that is produced is, if not directly interpretable, easier to "read."

For clearly structured corpora such as scientific article abstracts, the topic modeling approach is extremely powerful (Wang & Blei, 2011). However, for corpora less strongly structured other strategies are required. This is the case with the Bible, where it is not clear what meaningful units of text are as there is not a clear paragraph structure to rely on. The seemingly obvious units − verses, sentences or chapters − are all problematic. Verses have too few words to allow for a meaningful co-occurrence network. Sentences are too heterogeneous. Some chapters, for example, consist of only one sentence. And the fact of co-occurrence at the chapter level is meaningless since on average each chapter consists of roughly 300 verses. Against this background, we run a sliding window of five verses across the Bible within books[6] (Erikson & Bearman, 2006) to capture the semantic relationship between terms using a simple decay function (Danowski, 2009) to capture the idea that co-occurrences between terms falling in the same verse are more meaningful than co-occurrences between terms several verses apart. More precisely, we assign a value of 1 for two co-occurring terms within the same verse, 0.5 when found one verse apart, 0.33 if two verses apart, etc. and 0 if greater than five verses apart. The results reported here are similar to those obtained under a wide arrange of decay functions and window lengths.

The result is a W by W co-occurrence matrix, where W includes a selection of 1000 words in the Bible[7] and the value W_{ij} is the sum of the number of times words i and j co-occur within five verses of each other in the text, inversely weighted by distance. We use cosine similarity to produce a semantic similarity matrix, where S_{ij} captures the similarity between words i and j in terms of the co-occurrence with other words. We conceptualize and analyze this similarity matrix as a weighted network, where an edge between words i and j is weighted by value S_{ij}. In its raw form, this similarity network is very dense − in very rare cases will i and j have 0 similarity. We therefore sparsify the similarity matrix by specifying a threshold T, such that if S_{ij} is less than T, it is set to 0; otherwise, it is set to S_{ij}. Words are dropped from the network if they have no ties to any other word and the specified threshold is chosen to maximize log(n) * modularity, in attempt to simultaneously maximize the number of connected words in the network and ability of the detection algorithm to identify groups.

In the plots below, we draw the sparsified similarity matrix as a network using the Fruchterman-Reingold layout algorithm. Communities of words, or topics, are identified using the Louvain community detection algorithm (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008).

---

[5] Topic modeling assumes documents are mixtures of topics, with each topic corresponding to a likelihood distribution of words. It treats words in each document as having been generated by a process that selects topics according to their mixture proportion and then selects words from a given topic according to their probabilities of belonging to that topic. Topic modeling refers to an estimation strategy, which attempts to infer from a set of documents the topics that generated their content, jointly estimating the distribution of word likelihood distribution in topics and topic likelihood distribution in documents.

[6] Cursory readings of the Bible, and technical experiments, show that chapter divisions do not demarcate meaningfully or semantically distinct units of text. Often, a single idea or story will be carried across chapter lines. We therefore ignored chapters when sliding windows across the text. Only book ends are used as stoppage points.

[7] Word selection involves a series of language processing scripts, which remove stop words, stem words to standardize related forms, perform column-wise chi-squared tests to select words with distinct co-occurrence patterns to reduce semantic redundancy between words. The words with the top 1000 chi-squared scores (summed over all other words) are included in the analysis. Experiments using other measures of inclusion (such as log-likelihood) and different amounts of words produce similar network structures, signaling robustness to analytic decisions.
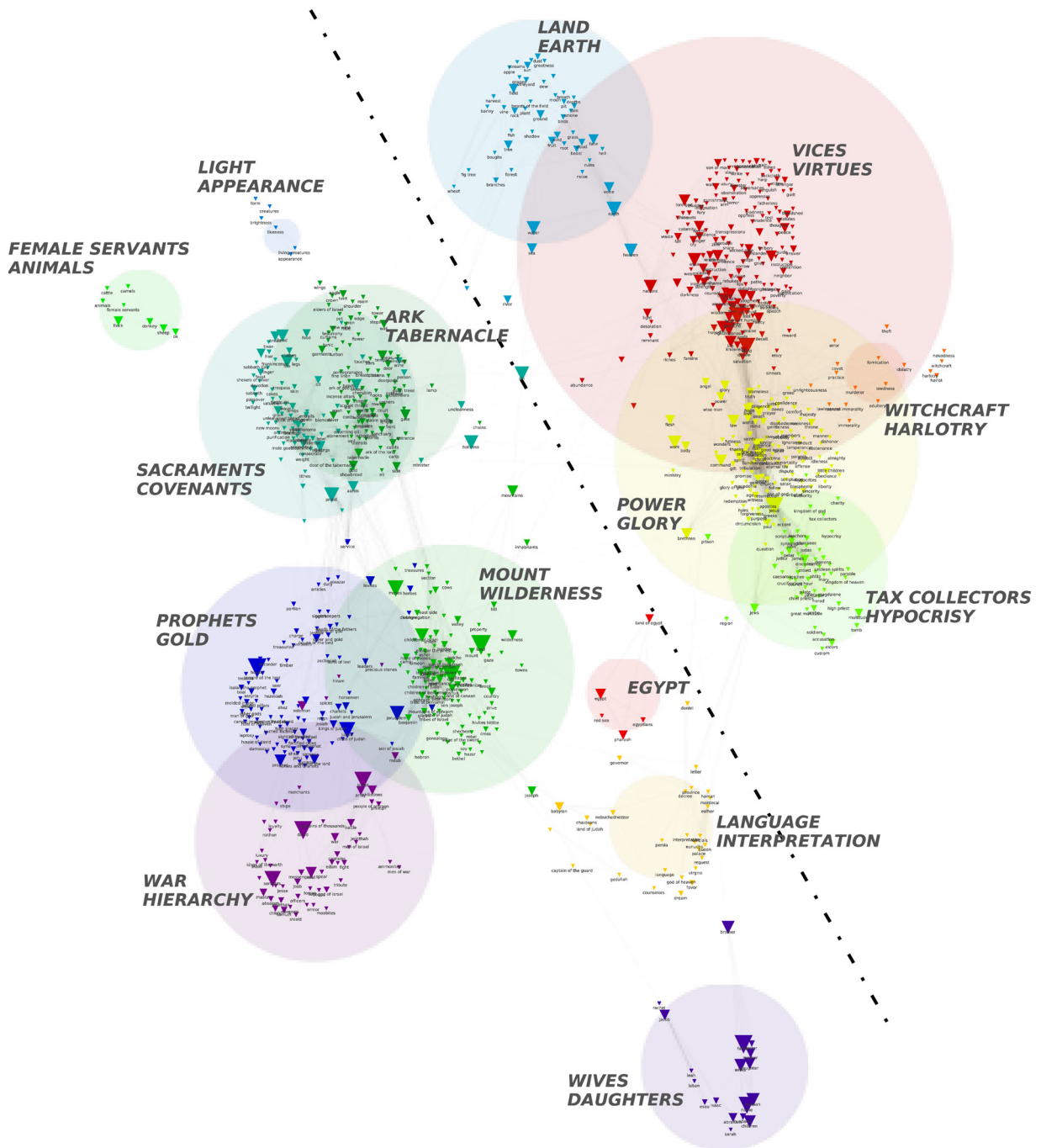
Fig. 4. The Semantic Structure of the Bible.

4. The semantic structure of the Bible

Clusters − or communities of words are identified in a semantic mapping that is characterized by a large partition separating two distinct regions composed of relatively tightly and therefore overlapping clusters as shown in Fig. 4. The major partition runs down what might be conceptualized as a main diagonal, from the top left to the bottom right.

Colors are associated with unique communities. The small green cluster on the far left of the graph identified "Female Servants and Animals" provides a window into how to interpret the fact that words are clustered together into communities. This cluster contains the words: *animals, cattle, flock, camels, donkey, sheep, ox, and female servants*. This is not a mistake. These are all things that men subjugate. Female servants are not co-located with male servants, wives or daughters, but instead with animals. The fact that the

writers of the Bible distinguish female from male servants and female servants from daughters, or that female servants are seen as equivalent to beasts of burden, animals, things without agency of their own is not the surprise here. The closeness of words to one another provides insight into their meaning, since the meaning of words is conditional on their position in the larger semantic network. The surprise, if there is one, is that the semantic mapping reveals that fact so elegantly.[8]

We can move systematically across this semantic space from *Female Servants and Animals* to the other clusters. To the right is "*Light and Appearance*" which reflects the core of creation; *likeness, form, brightness, appearance, creatures*. Continuing clockwise, and crossing the main partition "*Land and Earth*" is composed of all the words of the tangible material world; the Bible bleeds materiality: here we have *plants, roots, vines, boughs, harvests, grapes, figs, birds, smoke, beasts, fruit, forests, trees* all found together at the top of the graph. Heaven and earth are tightly coupled; hell is linked to depths, surrounded by *ruins, clouds, smoke, rain, and beasts*.

Below is the central community (*Vices and Virtues)* which contains all of the sins and virtues related to law and truth, *blasphemy, disobedience, ignorance, forbearance, idleness*. Here is found Satan, God, prayer, work, linked to a dense cluster of virtues (*joy, mercy, salvation knowledge, praise, integrity, righteousness, honor*, and sins that are at their core relational: *deceit, gain, strife, vengeance, anger, oppression*. To the right, in the cluster labeled "*Witchcraft and Harlotry*" are the sins and virtues of the Old Testament; they are sins of women or sins specifically related to women: *witchcraft, harlotry, coveting, judging, and lewdness*. These are the mechanisms for disaster, separate from but adjacent to the sins of conduct: *lust, murder, sodomite, sexual misconduct, conceit, envy, frivolity, anguish boastfulness*, along with their comparable virtues: *prudence, trust, profit*. Coupled together are sins of deception: *murmuring, hypocrisy, usury, extortion, self-indulgence*. In the center of these are *officials and tax collectors*.

The world of moral action is mainly to the right of the main partition. The story that the Bible tells − of course also directed towards a world of action − is on the left. At the very bottom of the left side of the graph in a large light purple cluster war and hierarchy) are found a community of words associated with structure, hierarchy, and war: *war, battle, horsemen, shield, weapons, armor, master, servant, servants, messengers, spear, fight, lord of Israel, David, captains, men of war, chariots*, and so on. Note that male servants are embedded in this community. Above the war community is a cluster of words centered on the exodus, the slaughter of the first born, Moses, Aaron, Tribes of Israel, the wilderness. To the left are the found treasures of all kinds: *treasures, silver, luxuries, treasuries, gold, the Kings of the Earth, merchants, tithes, overseers*, and the tools and stuff of trade, *ships, spices, recorders*. Above the arc, the covenant, the communities they brought together as congregations; lands and locations, the variety of rituals of cleanliness, washing. Wives and daughters are in the bottom-most cluster.

Any close investigation of any of the terms co-occurring across the Bible reveal deep subtleties of meaning. While we may have all known that female servants were subjugated − that they were structurally equivalent not to male servants but to animals − here we discover an especially crisp confirmation. Likewise, the decoupling of vices associated with sexuality from others is striking. Throughout the mapping we discover, for example, that Passover is associated with twilight; that tithes are associated with Priests, abundance with riches with righteousness. The mapping of words in the semantic structure provides new ways of understanding what they meant.

5. The Old and the New Testament

Where are these clusters arising from? To see how the New and Old Testament differ, in Fig. 5 we project a heat map that captures the distribution of words in the Old and New Testament across the skeletal semantic mapping. To do so, we simply measure a score for each term that measures the extent to which it is specifically attached to the New Testament. In doing so, we are able to compare the observed number of occurrences of a word in the New Testament relative to the sheer number of occurrences in the Bible as a whole. Because the New and Old testaments are uneven in size we normalize for size variation. For a given term *w* with frequency $f^{\mathfrak{N}}(w)$ in the New Testament, and global frequency *f* in the Bible, the specificity score of w in the New Testament is $\frac{f^{\mathfrak{N}}(w) - \hat{f}^{\mathfrak{N}}(w)}{\hat{f}^{\mathfrak{N}}(w)}$ where $\hat{f}^{\mathfrak{N}}(w)$ stands for the expected frequency of *w* if it were evenly distributed between New and Old Testament meaning; viz $\hat{f}^{\mathfrak{N}}(w) = f^{\mathfrak{N}}(w) \frac{\sum_{i \in V} f^{\mathfrak{N}}(i)}{\sum_{i \in V} f(i)}$, where *v* is all terms included in the semantic map. Once this score has been computed for every word, we produce a heat map with a Gaussian interpolation over the original layout of the semantic network. The color map is logarithmically scaled.

Red is associated with the New Testament, blue with the Old Testament. The first thing to observe is the dense cluster of concepts associated with the New Testament on the right side of the graph. The New Testament is about righteousness and transformation that arises from a mandate to be further from sin − which concern not only action (adultery, for example) but thought (lust). It is also about *damnation, the Devil, Satan, vice, lawlessness, death*. It is also about *redemption, meekness, love, forbearance, sincerity, modesty, the*

---

[8] Of course, there is also textual evidence for this interpretation. Consider Benveniste and Lallot (1973): "First in Homer, *Il.* 23, 550: "you have in your house much gold, bronze, *próbata* and servants [handmaids]." … After scrutiny of all the examples in Herodotus we can affirm that it is applied to live-stock, large or small. … The expression [is] morphologically related to *próbata*…. It is the Homeric word *próbasis* (πρόβασις)… The Homeric expression denotes wealth: …Thus *próbasis* indicates the totality of *próbata*, and the opposition *keimḗlia/próbasis* refers to possessions of two different categories, a distinction which seems to be essential in the economy of the Homeric world: Immovable or "lying" *(keimḗlia* from κεῖμαι 'lie'), i.e. immovable property, and movable property (*hósa probaínei*). … This way of regarding property in its two categories has a rough resemblance to the French distinction between *meubles* (*mobilia*) and *immeubles* (*immobilia*). But *immeubles* are buildings, whereas *meubles* are chattels ["numéraire": money, coinage, face-value…]. In Homeric Greece the division was different: all that "lies" (*keîtai*), *keimḗlia*, precious metals in ingots, gold, copper and iron, is opposed to *tà próbata*, property on the hoof, consisting of the herds and live-stock in general. Such is the sense of *próbata* as we have established from the textual evidence." And therefore, the association of handmaids with livestock as different forms of moveable wealth, distinct from other forms. We thank Martin Barnay for drawing our attention to this fact.
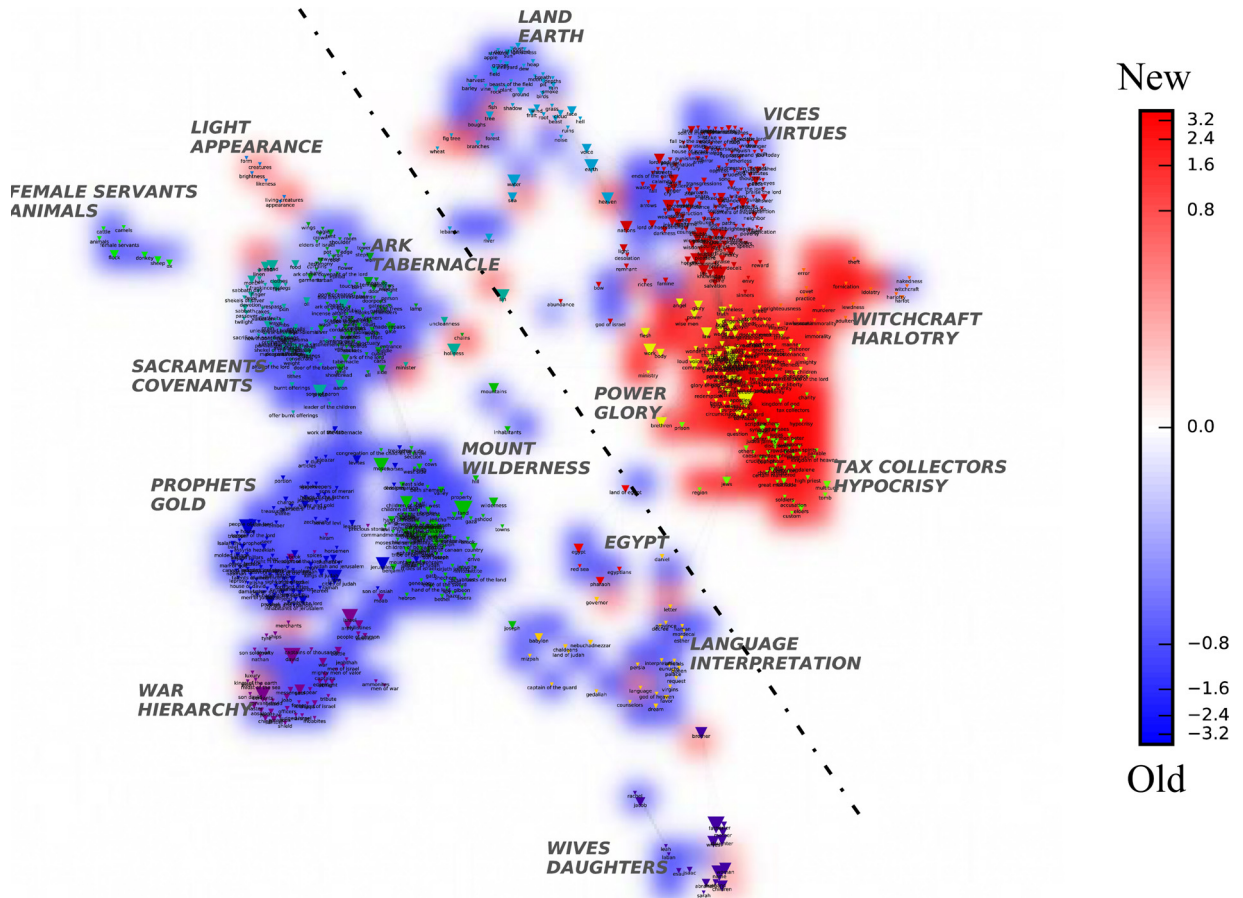
Fig. 5. Heat Projection of the Old and the New Testament.

flesh, supplication, seduction, immorality, sex, fornication, lust and *idolatry*. A small space of sin is specifically associated with the Old Testament: the sins of the action of women − the sins of *witchcraft, magic, harlotry*.

There is no cluster *not composed of both* the Old and the New Testament, not even the small cluster of things men subjugate; women and animals. The cluster about war is bordered by New Testament words: armor, weapons, master. Carved and molded images, prophecies, and prophecy are New Testament additions to the historical components of the Old Testament. But most striking is how interweaving the Old Testament with the New Testament holds the Bible together. Removing just six of the central concepts associated with the New Testament: *Sin, Light, Brethren, Darkness, Heaven and Hell*, breaks the Bible into multiple disjoint components.

6. Competing uses of the Bible

The semantic heat map is just that; a mapping of the people and places and concepts that make up the Bible. It is a skeletal structure: an opportunity for interpretation. We now turn to how the Bible was used. The context we explore is the long 18th century, from 1660 to 1780 in England. We exploit data arising from the Preachers' Assistant, which is transformed into a tri-partite edge list linking people, themes, and Biblical verses.

Each theme is employed in some number of conformist sermons and some number of dissenter sermons. Given the total number of sermons by Dissenters and Conformists, we can produce an expectation about how many conformist and dissenter sermons were likely to have used a given theme. We can then compare the observed number to the expected number, using the log-likelihood equation, to identify the themes and chapters of the Bible most uniquely associated with dissenter and conformist sermons (Rayson & Garside, 2000). Formally, for a theme or chapter, the log-likelihood of observing a specific number of Dissenters and Conformists that we observe:

$$-2\ln\lambda = 2\sum_i O_i\ln\left(\frac{O_i}{E_i}\right)$$

, where *i is a given topic,* O_i is the number of Dissenter or Conformist we observe using theme *i, and* the expected value $E_i = \frac{N_i\sum_i O_i}{\sum_i N_i}$. Performing this procedure for each theme results in a list of log-likelihoods, one for each theme. We split the log-likelihoods by

Table 1
Top Topics and Chapters in the Dissenting and Conformist uses of the Bible.

Panel A: Top Topics	
Dissenter Topics	Conformist Topics
Providence	Proper Gospel
Justification	Rebellion
Hear Word	Epiphany
Conversion	Irish Protests
Family Worship	Universal Obedience
Christ's Transference	Daily Bread
Life of Faith	Circumcision
Divine Meditation	Deliverance from Temptation and Evil
Free Grace	Resurrection
Ref Manners	Smallpox

Panel B: Top Chapters	
Dissenter Chapters	Conformist Chapters
Ephesians 1	St. Matthew 5
Romans 5	St. Matthew 6
Psalms 36	James 2
Romans 14	Romans 12
Isaiah 53	James 4
Genesis 22	Acts 24
Song of Solomon 2	St. Matthew 2
Exodus 33	St. Matthew 20
St. Mark 10	Proverbs 31
Revelation 3	Psalms 42

whether a given theme or chapter was more frequently employed by either Dissenters or Conformists compared to expectation, giving us two lists of themes or chapters, one for Dissenters and one for Conformists, which we sort by their log-likelihood value to identify the K themes and chapters most overrepresented for each group.[9]

Table 1 (Panel A) reports the top 10 themes most uniquely associated with Dissenters and Conformists. The two most salient themes for Dissenters, providence and justification, reflect the dissenting pre-occupation with salvation and the means through which one comes to be justified and saved before God. Some such means are themes themselves − such as hearing the word of God, converting, living a life of faith and divine meditation. Conformists are much more oriented towards worldly and political matters − Rebellion, Irish protests, Smallpox, and the practical aspects of a good Christian life, including propagation of the Gospel, breaking the daily bread, deliverance from temptation and evil, circumcision, and universal obedience. The two exceptions to their this-worldly focus are epiphany and resurrection.

Table 1 (Panel B) reports the top 10 Bible chapters most uniquely associated with dissenters and conformists. The top chapters for dissenters emphasize the necessity of Christ and justification through faith in order to be saved. Ephesians 1 declares Jesus the head of the Church, often cited by Dissenters as a critique of papal authority. Romans 5 discusses the purity and perfection of Jesus, as the only human to have never sinned, and justification by faith in Christ as the basis of salvation. Romans 14 emphasizes justification through faith as well. St. Mark 10 argues against divorce: "What therefore God has joined together, let not man separate" − a sensitive topic which divided Anglicans and Catholics and Dissenting Protestants.

The conformist chapters discuss Christian positions of traditional importance. Four out of the top ten conformist chapters are from the book of St. Matthew. St. Matthew 5 contains the well-known quote "the meek shall inherit the Earth", signifying the importance of the poor and downtrodden in Christian culture, and Romans 12 argues that we should love others regardless of their wealth or status. St. Matthew 6 discusses prayer and alms and James 4 discusses submitting oneself to God, traditional Christian duties. The dissenters and conformist lists both include chapters concerning the offering of Isaac by Abraham before God. However, they cite different chapters − Dissenters cite the initial Genesis 22, where it is first related, while the conformists cite James 2, where it is argued, using the offering of Isaac as an example, that faith and good deeds are co-implicated and both necessary to achieve salvation.

Of course, themes and chapters in the Bible do not occur independently; rather, certain themes and chapters tended to co-occur together at high probability, often higher than we would expect by chance. To explore differences between dissenter and conformist uses of the Bible in terms of the joint distribution of themes and chapters, we constructed a chapter-by-theme matrix, M, where M_{ij} is the number of times that any verse from chapter i is used as source scripture in a sermon discussing theme j. We create two versions of M, one for Dissenters and one for Conformists such that D_{ij} is the number of times a Dissenter cited scripture from chapter i when

---

[9] Given the descriptive purposes of this analysis, we do not conduct the hypothesis-test, normally computed by comparing the resulting log-likelihood values to a chi-squared distribution table. As Rayson and Garside (2000) show, without the hypothesis test, the equation given above is suitable for purposes of identifying relative differences between corpora.
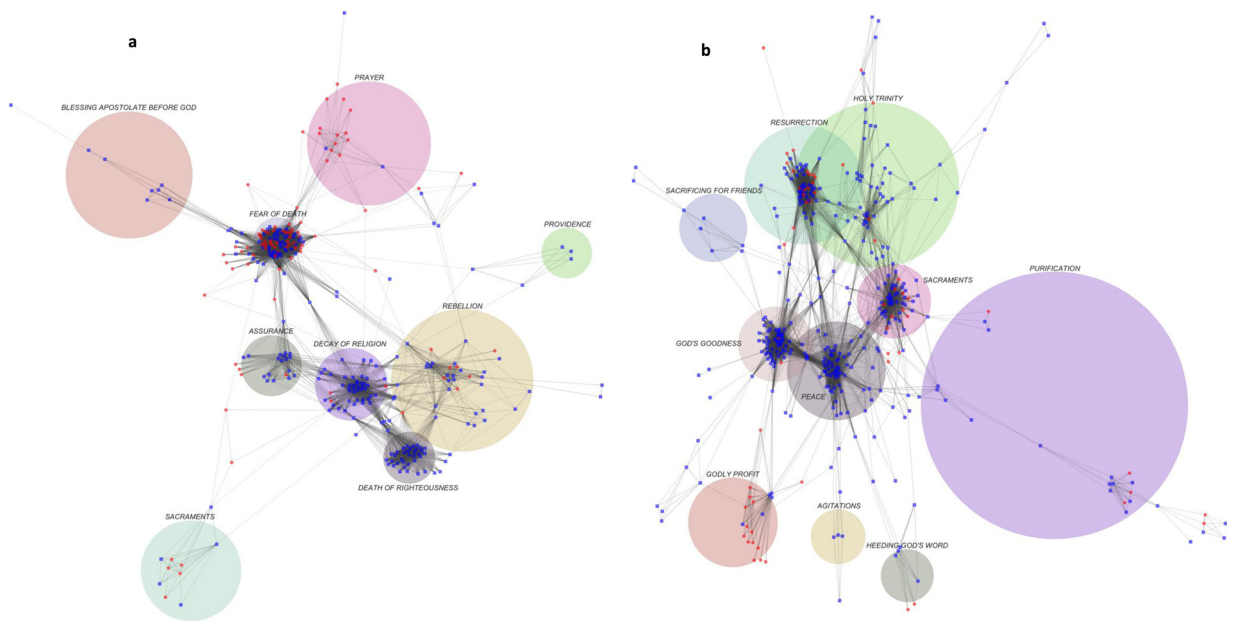
Fig. 6. Panel A: The Dissenting Structure of Topics. Panel B: The Conformist Structure of Topics.

giving a sermon concerning theme j, while C_{ij} is the same, but for Conformists. We then construct a similarity matrix between chapters for Dissenters and Conformists using cosine similarity, such that the similarity between chapter i and chapter j is the cosine of the angle of their frequency vectors over themes. We then plot a network, where an edge between two chapters in the Bible is drawn if their cosine similarity value exceeds a threshold of T, again chosen to maximize log(n) * modularity. Similarly, to the semantic network, we use a Louvain clustering algorithm to identify groups of chapters associated with related themes. For each cluster, we chose for a name the theme they most predominately and uniquely cited. In Fig. 6, Panel A we show the structure of themes addressed by Dissenters, coloring the verses by location in the New or Old Testament respectively. Panel B reports the same graph, but for Conformists.

In Panel A we see the core of dissenting belief highlighted dramatically. The center of the dissent network is the Fear of Death. At the bottom of the graph are five clusters (Decay of Religion, Death of Righteousness, Sacraments, Rebellion and Assurance) that summarize the dissenting critique of established state religion. Sermons about the decay of religion and the death of righteousness draw heavily from the Old Testament, including 2 Kings 9, in which Jezebel, Queen of Israel, is eaten by dogs as punishment for worshipping false gods. They figure centrally, serving as the central core linking to the justness of Rebellion, marked by Romans 13, which discusses the conditions under which a Christian should obey government, and Acts 7 which discusses rebellion against God. The sacrament and assurance themes are instantiations of the religious critique, emphasizing communion and "render[ing] to Caesar the things that are Caesar's, and to God the things that are God's" (Luke 22, Mark 12, Matthew 22).

Connected to Rebellion are sermons about Providence, about the end to which Rebellion is directed, the creation of Providence on Earth and deliverance by God from Evil. The core of dissenting belief is captured in the dense cluster of sermons disproportionately drawn from the New Testament about the Fear of Death (the deaths of Samuel (Samuel 25), Rachel and Isaac (Genesis 35), and the comfort of Christ's coming (1 Thessalonians 4, Matthew 13). The mechanisms for salvation − Prayer and blessing apostolate before God (the demand through example or missionary activity in the service of conversion of others so as to lead them to accept God's will) − are the inner-directed and other-directed forms by which salvation can be achieved.

The conformist map with respect to themes is strikingly different both with respect to form and content. At the center of the graph are three clusters, Peace, God's Goodness, and Sacraments. These clusters link at the bottom of the graph to Godly Profit, Agitation, and Obedience, or Heeding God's Word. Agitation is the conformist analog of Rebellion; it is what is not to be done. Godly Profit addresses the problem of theodicy, drawing on, for example: 1 Timothy 4.8 "For bodily exercise profits a little, but godliness is profitable for all things, having promise of the life that now is and of that which is to come." At the top of the graph are topics dealing with sacrifice, resurrection, and the holy trinity. Finally, a sparse cluster, largely disjoint from the rest of the graph, deals with purification. Here we see a conversation oriented not towards Dissenters but to Catholics, since Purification in this context refers to the Protestant critique of the Catholic idea of Purgatory as a transitional moment/space that can be transcended through prayers for the dead. Such prayers, and the costs associated with them, provided one of the elements of the Protestant critique of the Catholic Church. With the exception of a New Testament verses under the topic of Resurrection such as 1 Corinthians 15:3-4: "For I delivered unto you first of all that which I also received, how that Christ died for our sins according to the scriptures; And that he was buried, and that he rose again the third day according to the scriptures", the conformist Bible at the level of topics and books is disproportionately drawn from the Old Testament. We can, however, delve a little bit deeper than this dual representation.
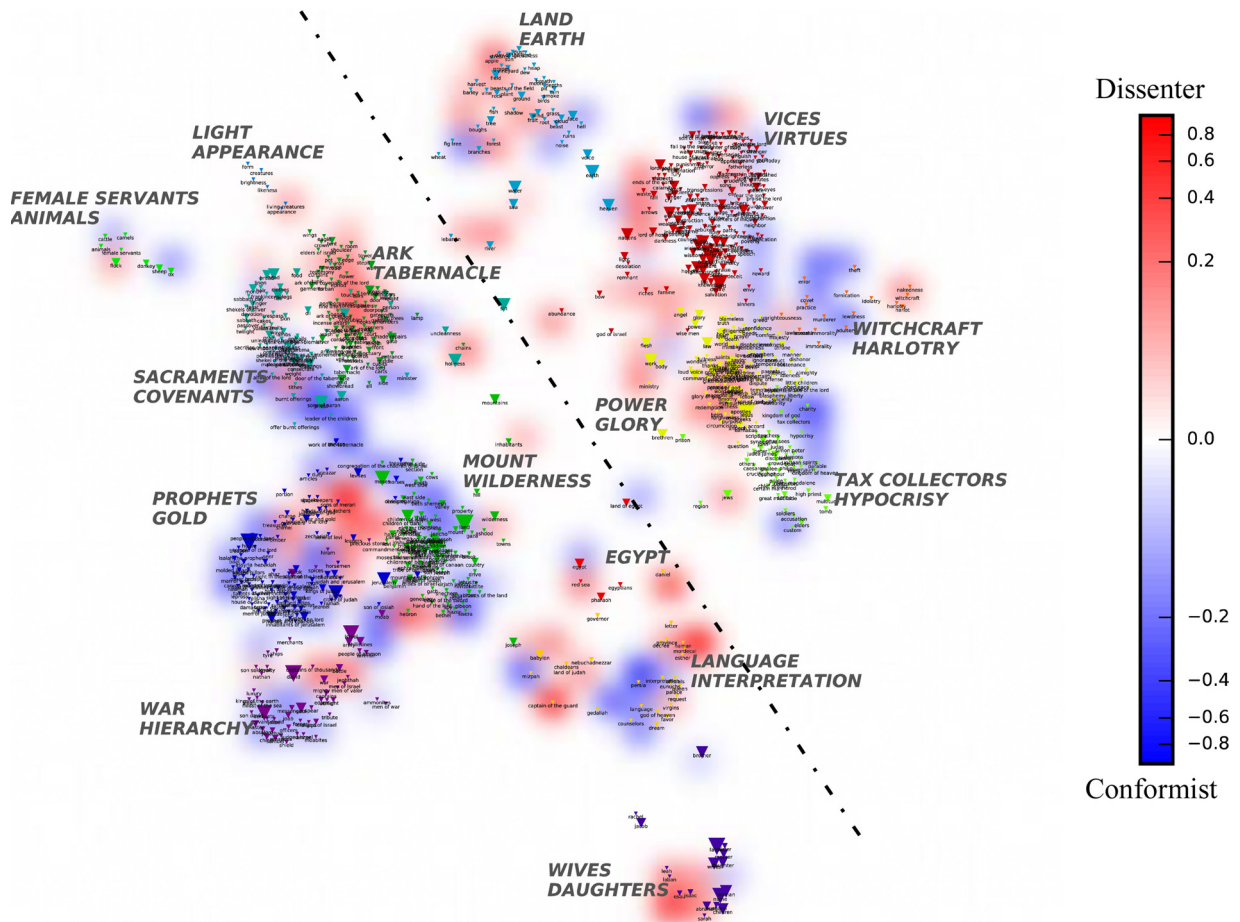
Fig. 7. Heat Map of the Dissenting and Conforming Bible.

7. The conforming and dissenting sermon heat map

We now link methodologically and conceptually the global semantic map of the Bible with the conformist and dissenter uses of the Bible. To do so, we follow the same methodology as the Old/New Testament heat map except we replace the frequency of words in each part of the Bible, with the cumulated number of times a word appears in a chapter cited either by a dissenter or conformist sermon. We again measure a score for each term of the extent to which it is specifically attached to Conformists or Dissenters and create a heat map using Gaussian interpolation over the original layout of the semantic network, where the color of words accords with their normalized use frequency by either Dissenters or Conformists. This heat map then shows how Dissenters and Conformists read the Bible differently focusing on strongly divergent vocabularies (Fig. 7).

Words associated with Conformists occupy central positions in communities below the central axis, while words associated with Dissenters occupy central positions in communities above the axis. This conforms with an understanding of Dissenters as emphasizing moral action. Dissenters were also superstitious and worried about witches. Fear of witches reflects a central aspect of dissenter theology, the subjugation of women to their husbands. The perceived need to subjugate women arose from the dissenting belief that women were inferior to men, and that because they were inferior they were at greater risk of ending up in error, in following the work of the devil, in part because they were not as close to God as men (Westerkamp, 1997). We can see all of these themes, the concern with witchcraft and the specific sins of women − harlotry, lust − and the desire for subjugation in the dissenters' use of the Bible. Note specifically the very strong dissenting region of the Bible on the far right of the graph, containing the Old Testament concern with harlotry and witchcraft, as distinct from the conformist concern with vices that distort or sustain ordinary social relations; for example, usury, slander and bribery. This focus on the sins of women is matched by a correspondingly strong focus − relative to conformists − on the proper relationship of women to men. We see this in the very dark dissenting cluster at the bottom of the graph − on wives and daughters.

Dissenters relative to conformists are also much more this worldly in their orientation; that is salvation as the end is achieved through the methodical giving over to oneself to the practical activity of adding values to the world, as a mechanism for enhancing God's glory. We see this focus strongly in the emphasis on words arising in the Arc and Tabernacle community and at the top of the graph in the Land and Earth community. Not surprisingly, dissenters' sermons are heavily over-represented in the center of the graph,

where the themes of God's glory and power are paramount. Contrast this focus on the expression of God's power with the conformist focus on the prophets, the sacraments, and covenants − key characters and elements of the Judeo-Christian story.

None of these different emphases are surprising given the historical and theological record. That they are revealed so crisply from our sermon data is rather remarkable, however. It tells us that the distant reading of the Bible we have put into place makes sense of the ways in which different communities of users found in the Bible material for the serious business of building identity.

8. Discussion

For two millennium scholars have interpreted the Bible. Such interpretation often seeks understanding by locating text in context − where context is defined by the situation in which it was written. What could be meant by a phrase for the people who first penned it? For those who first heard it? Read it? For some scholars, the relevant context is characterized by competing religious visions circulating when texts were written, heard, and read. This is not the strategy we invoke in this article. Nor do we try to suggest that the strategy we do invoke ought to replace that kind of interpretive effort.

Our interest, though, is in seeing how through "formalizing culture" − the special focus of this issue − we can come closer to formalizing a new way to study content/ion, by which we mean debates between groups over the deployment of text as symbols of group identity and group difference. The formal analysis of texts proceeds from within the text, often ignoring − by dis-embedding from − use. Texts are not just bags of words, encountered in one off readings by individuals, but tools and totems, sites of contention, and symbols of group difference employed strategically to induce and sustain group boundaries. Nowhere is contention over text more salient than in religious debate, where the interpretation of religious texts has real stakes for which people are willing to give their lives. Religious texts like the Protestant Bible are lexically heterogeneous − replete with histories, parables, commandments, and premonitions and as a result, they are rife with possible interpretations. The fecundity of the material makes it that quite distinct groups can differ in their use of the same text. By focusing on how groups used the same text we begin to put forward a strategy for formalizing the study of cultural and literary contention, of moving text analysis from a study of abstracted content to the study of content/ion.

Contention is at the end of the day about meaning, and the analytic strategy we deploy in this article is to search for meaning by defining the relative context for words − for concepts − as the community of words that they are associated with. When "red" is most often in association with "green" and "yellow" and "orange" we can think that its meaning is related to some color palette. But when "red" is found with "communist", "brigade", "diaper", "Lenin", "scare", and so on, we would know that its meaning is related to a political palette. Likewise, we know that the context for a word is the verse or the sentence or the paragraph in which it is found, and that this specific, local context tells us how to make sense of it. These bases for understanding are so routinely part of our everyday understanding that they almost feel not worth remarking on. But the fact that they are so obvious is beneficial for us, since the analysis strategy undertaken in this paper relies on our tacit acceptance of the idea that words mean different things depending on which other words they are associated with.

In doing so, we induce a fixed mapping − a skeletal structure. The skeletal structure of the Bible is revealed for the first time using such co-occurrence methods. It is revealed to be a beautiful object. What more could science want to achieve but to find a way to induce new, aesthetically rich, deeply textured objects? Just possibly, that, *and* a cogent demonstration of a new, more sophisticated, formal approach to how those objects − texts in this case − are used by other texts.

References

Appleby, D. J. (2014). *Black bartholomew's day: Preaching, polemic and restoration nonconformity.* Manchester: Manchester University Press.
Bearman, P., Faris, R., & Moody, J. (2009). Blocking the future: New solutions for old problems in historical social science. *Social Science History, 23*(4), 501–533.
Bearman, P. (1997). Generalized Exchange. *American Journal of Sociology, 102*(5), 1383–1415.
Benveniste, E., & Lallot, J. (1973). *Indo-European language and society.* Miami: University of Miami Press.
Benzécri, J.-P. (1973). *L'Analyse des Données, Vol. II*Paris, France: Dunod [L'Analyse des Correspondances].
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3,* 993–1022.
Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008.
Burnet, G., 1766. A discourse of the pastoral care. C. Bathurst, J. Rivington, etc, London.
Chamberlain, J. (2012). Parish preaching in the long eighteenth century. In K. Francis, & W. Gibson (Eds.). *The oxford handbook of the british sermon, 1689–1901* (pp. 47–62). Oxford: Oxford University Press.
Claydon, T. (2011). The sermon culture of the glorious revolution: Williamite preaching and jacobite anti-Preaching, 1685–1702. In P. McCullough, H. Adlington, & E. Rhatigan (Eds.). *The oxford handbook of the early modern sermon* (pp. 481–487). Oxford: Oxford University Press.
Crawford, M. J. (1987). Origins of the eighteenth-Century evangelical revival: England and New England compared. *Journal of British Studies, 26*(4), 361–397.
Danowski, J. A. (2009). Inferences from word networks in messages. In K. Krippendorff, & M. A. Bock (Eds.). *The Content Analysis Reader* (pp. 421–429). CA: Sage, Thousand Oaks.
Deconinck-Brossard, F. (2009). The art of preaching. In J. V. Eijnatten (Ed.). *Preaching, sermon and cultural change in the long eighteenth century* (pp. 95–130). Boston: Brill.
Dixon, R. (2011). Sermons in print, 1660–1700. In P. McCullough, H. Adlington, & E. Rhatigan (Eds.). *The Oxford handbook of the early modern sermon* (pp. 481–487). Oxford: Oxford University Press.
English short title catalogue. http://estc.bl.uk.
Erikson, E., & Bearman, P. (2006). Malfeasance and the foundations for global trade: The structure of English trade in the East Indies, 1601–1833. *The American Journal of Sociology, 112*(1), 195–230.
Farooq, J. (2013). *Preaching in eighteenth-Century London.* Suffolk: Boydell Press, Woodbridge.
Ferrell, L. A., & MacCullough, P. (2000). *The English sermon revised: Religion, literature and history 1600–1750.* Manchester: Manchester University Press.
Francis, K. (2012). Sermons: Themes and developments. In K. Francis, & W. Gibson (Eds.). *The Oxford handbook of the british sermon, 1689–1901* (pp. 31–44). Oxford: Oxford University Press.

Gibson, W. (2012). The british sermon, 1689–1901: Quantities, performance and culture. In K. Francis, & W. Gibson (Eds.). *The oxford handbook of the british sermon, 1689–1901* (pp. 3–30). Oxford: Oxford University Press.

Haydon, C. (2002). Religious minorities in england. In H. T. Dickson (Ed.). *A companion to eighteenth-century Britain* (pp. 241–251). Wiley-Blackwell.

Hooker, R. (1875). In R. Hooker (Ed.). *The works of that learned and judicious divine, Mr. Richard Hooker*. Oxford: The Clarendon press.

Hughes, A. (2013). The 'Chalk' and the 'Cheese': David underdown, regional cultures and popular allegiance in the english revolution. *History Compass, 11*(5), 373–380.

Key, N. E. (1990). Comprehension and the breakdown of consensus in Restoration Herefordshire. In T. Harris, P. Seaward, & M. Goldie (Eds.). *The politics of religion in restoration England* (pp. 191–215). Oxford: Basil Blackwell.

Key, N. E. (1994). The political culture and political rhetoric of county feasts and feast sermons, 1654–1714. *Journal of British Studies, 33*(3), 223–256.

Key, N. E. (2014). The 'Boast of antiquity': Pulpit politics across the atlantic archipelago during the revolution of 1688. *Church History: Studies in Christianity and Culture, 83*(3), 618–649.

Letsome, J. (1753). *The preacher's assistant.* [London printed for the author].

McCullough, P. E., Adlington, H., & Rhatigan, E. (Eds.). (2011). *The Oxford handbook of the early modern sermon.* Oxford: Springer.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the workshop on comparing corpora. Association for computational linguistics,* 1–6.

Rule, A., Cointet, J. P., & Bearman, P. (2015). Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Science, 12*(35), 10837–10844.

Sewell, W. H. (1989). Toward a theory of structure: Duality, agency, and transformation. *The American Journal of Sociology, 98*(1), 1–29.

Spaulding, J. G., & Cooke, J. (1988). *Pulpit publications 1660–1782: Being a new edition of The Preacher's assistant.*

Tennant, B. (2015). Sermons as evidence: 'Holdings questions' and 'corpus questions'. *The Journal of Religious History, Literature and Culture, 1*(1), 3–31.

Wang, C., & Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining,* 448–456.

Westerkamp, M. J. (1997). Engendering puritan religious culture in old and New England. *Pennsylvania History, 64*, 105–122.

Mark Anthony Hoffman is a PhD candidate in Sociology at Columbia University. He is interested in the relationship between culture and social structure. His dissertation analyzes the reading habits of New York and American elites after the Constitutional Convention to reveal the concerns and interests of elite factions as they vied for political and economic power.

Jean-Philippe Cointet is an Associate Professor at Science-Po, whose principal appointment is in the Media Lab. His work focuses on socio-semantic dynamics in knowledge communities, textual corpora analysis, knowledge diffusion, opinion dynamics, knowledge network morphogenesis, and network mapping.

Philipp Brandt is a postdoctoral researcher at the University of Mannheim. His research analyzes the emergence of the data science profession, using both computational methods and qualitative field observations, to understand how expert roles are collectively constructed.

Newton Key is a Professor of History at Eastern Illinois University. He is currently at work on various aspects of an Association Crisis in late-Stuart London: mapping feasts, plotters, and the rhetorical construction of conspiracy in political sermons and addresses.

Peter Bearman is the Cole Professor of the Social Sciences at Columbia University. He is currently working on the seam of social science and cognitive social neuroscience, qualitative research designs, and a book on America in the 20th century. He is the author, with Adam Reich, of Working for Respect: Community and Conflict at Walmart.