

The Idea of Liberty, 1600–1800: A Distributional Concept Analysis

Peter de Bolla, Ewan Jones, Paul Nulty,
Gabriel Recchia, and John Regan

This article investigates the idea of liberty across the seventeenth and eighteenth centuries. Much contemporary political thought is rooted in a tradition established by intellectual historians of early modern and enlightenment Europe. Of course, political concepts such as “republicanism” were invented before this period, and the histories of most if not all of our contemporary political ideas can be traced back to classical times. Notwithstanding such *longue durée* accounts, more time-constrained analyses are warranted. By focusing on the two-hundred-year period between the beginning of the seventeenth century and the end of the eighteenth century, this article operates within a well-delineated tradition of scholarship.¹ Many of the ideas that contribute to our senses of contemporary social, legal, and political life were given explicit and extensive attention during these two centuries in Britain.

The scholarship and scholarly legacies of Quentin Skinner, John Pocock,

¹ This tradition is extensive but see *inter alia* Quentin Skinner, *The Foundations of Modern Political Thought*, 2 vols. (Cambridge: Cambridge University Press, 1978); J. G. A. Pocock, *Virtue, Commerce and History: Essays on Political Thought and History, Chiefly in the Eighteenth Century* (Cambridge: Cambridge University Press, 1985); C. B. Macpherson, *The Political Theory of Possessive Individualism* (Oxford: Oxford University Press, 1962); Alan Craig Houston, *Algernon Sidney and the Republican Heritage in England and America* (Princeton, NJ: Princeton University Press, 1991); Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford: Clarendon Press, 1997).

and Reinhart Koselleck establish a context for the observations that will follow. This scholarship is certainly not uncontested. In keeping with the deep and powerful traditions of intellectual history it has produced revision and recalibration.² That said, this article does not set out to adjudicate any of the local arguments that appear in this tradition; rather it aims to outline the potential contributions of a new method for the history of ideas based upon computational modes of inquiry recently developed within the Cambridge Concept Lab. Throughout, our attention will be directed at how the terms “freedom” and “liberty” have operated within the English language. The data upon which our investigations are based has been extracted from two vast digital archives of printed materials: Eighteenth Century Collections Online (ECCO) and the manually transcribed texts of Early English Books Online Text Creation Partnership (EEBO-TCP).

Comprising some 180,000 titles, 200,000 volumes, and more than 33 million pages of text, ECCO is well known as the world’s largest digital archive of books from the eighteenth century. It contains “every significant English-language and foreign-language title printed in the United Kingdom between the years 1701 and 1800.”³ The entire corpus has been scanned and optical character recognition (OCR) has been applied to the texts, resulting in a “machine-readable” version of each that can be subjected to computational analyses. A limitation of this resource is the high degree of error in the recognized text. The Early Modern OCR Project (EMOP), a project aiming to build a bespoke process for applying OCR to early modern texts to achieve high levels of accuracy, ultimately was able to achieve only 86% accuracy, and even the most up-to-date version of the OCR-based ECCO texts offered by Gale Cengage have been estimated at only 89% accuracy.⁴ OCR errors are far more likely to result in nonwords than they are to transform words to other valid words, so digital searches for particular terms (e.g., “freedom”) will underestimate the frequencies of these words. That said, because our analyses primarily *compare* frequencies and associations of particular terms and phrases, proportions are more

² Among others see: Annabel Brett and James Tully, with Holly Hamilton-Bleakley, *Rethinking the Foundations of Modern Political Thought* (Cambridge: Cambridge University Press, 2006).

³ *Eighteenth Century Collections Online*, Gale, accessed July 6, 2018, <https://www.gale.com/primary-sources/eighteenth-century-collections-online>.

⁴ Laura C. Mandell, Matthew Christy, and Elizabeth Grumbach, *EMOP Mellon Final Report*. Initiative for Digital Humanities, Media, and Culture, Texas A&M University, September 30, 2015, accessed July 6, 2018, <http://emop.tamu.edu/news>. Laura C. Mandell et al., “Navigating the Storm: IMPACT, eMOP, and Agile Steering Standards,” *Digital Scholarship in the Humanities* 32, no. 1 (2017): 189–94.

important than absolute frequencies. As there is no reason a priori to believe that one of these words or phrases will be vastly more subject to OCR error than the other, we have confidence that the OCR error is not having a disproportionate impact on our conclusions, but it must be kept in mind as a source of “noise.”⁵

EEBO consists of over 125,000 books published in English, primarily between the years of 1600 and 1700, and drawn from Pollard and Redgrave’s Short-Title Catalogue (1475–1640), Wing’s Short-Title Catalogue (1641–1700), Thomason Tracts (1640–1661), and the Early English Books Tract Supplement (sixteenth and seventeenth centuries).⁶ While access to OCR-applied ECCO texts are accessible only to subscribers, EEBO’s texts are available without restriction. EMOP (the only research group of which we are aware that has attempted to apply OCR to the full EEBO corpus) has achieved 68% word accuracy. Its files can be searched online at 18thconnect.org. Researchers who make manual corrections using this online interface are permitted to download the specific files that they have corrected. However, EMOP’s license prevents them from making a machine-readable version of the whole of EEBO available to the wider research community in a format suitable for text and data mining. For this reason, our analyses of EEBO are necessarily restricted to the manually transcribed texts of EEBO-TCP. EEBO-TCP grows as new texts are transcribed and added to the dataset. At the time we obtained access to the corpus, it contained 52,915 texts in total, over 90% of which fell between the years 1600 and 1700 and were therefore used in this analysis.

Regarding the ECCO dataset, it is important to consider the impact of a text’s reprints or editions.⁷ Word-usage counts that are based on the entire content of the dataset need to be tempered with respect to the vagaries of eighteenth-century publication. Our own analyses of the dataset, however, indicate that, for our purposes, the noise that is produced by multiple editions or printings of the same text is not significant. ECCO contains approximately 207,628 texts. The number that are alternate editions of another work with the same title and volume is 7,679 (or 3.70% of the

⁵ We have also applied a bespoke “clean up” method in order to improve accuracy. A full explanation of this method can be found at <http://www.crash.ac.uk/programmes/the-concept-lab-cambridge-centre-for-digital-knowledge>.

⁶ See *Historical Texts*, a partnership between Jisc and Knowledge Integration, Gooii, Sero Consulting and the British Library, <https://historicaltexts.jisc.ac.uk/collections#eebo>.

⁷ On the book trade in the period see James Raven, *The Business of Books: Booksellers and the English Book Trade 1450–1850* (New Haven, CT: Yale University Press, 2007); James Raven, *Judging New Wealth: Popular Publishing and Responses to Commerce in England, 1750–1800* (Oxford: Oxford University Press, 1992).

total number of texts). The number that are identical to another work with the same title, volume, and edition but have a different publisher's or printer's imprint is 6,482 (3.12%). The number of outright duplicates is 1,362 (0.66%).

Our aim in this study is to test the efficacy of computational text-mining techniques for the history of ideas. More specifically, we use digital tools and computational and statistical methods developed within the Cambridge Concept Lab to effectively read an archive in its totality, offering a complement to mainstream histories of ideas that are based upon the close reading of a small sample of texts. No one scholar can read the entire printed archive of this period, and it is important to recognize that computers (accurately speaking) cannot do so either. But data analysis can build a picture of the culture at large insofar as it has been preserved or sedimented within the archive of printed books represented by our two datasets. Our approach, then, turns away from grand theory, or the master tradition of thinking about ideas, in order to explore their dispersal and traction within the culture at large. To some extent the analytical approach we take here can be compared to some current projects in historical linguistics such as LDNA, based in Sheffield, which also uses methods developed in corpus and computational linguistics. That project focuses on the early modern period, using a transcribed subset of EEBO-TCP in combination with a thesaurus categorization of word senses from the period to examine the change over time of raw word association frequencies and pointwise mutual information scores between pairs of terms of interest.⁸ A different project based in Amsterdam, Texcavator, allows users to explore the development of sentiment around issues in newspaper text, presenting results in the form of histograms of word clouds and word and sentiment dictionary frequencies, alongside document metadata.⁹ And a project based in Brussels has created a method for multi-dimensional scaling of distributional semantic change, in order to analyze a change in meaning in positive evaluative adjectives in American English from 1860–2000, using Pointwise Mutual Information (PMI)-weighted co-occurrence scores derived from ten word windows around the term of interest.¹⁰ Lastly, there is a project based in

⁸ See Susan Fitzmaurice et al., "Linguistic DNA: Investigating Conceptual Change in Early Modern English Discourse," *Studia Neophilologica*, 89 sup 1 (2017): 21–38.

⁹ See Joris van Eijnatten, Toine Pieters, and Jaap Verheul, "Using Texcavator to Map Public Discourse," *Tijdschrift voor Tijdschriftstudies* 35 (2014): 59–65.

¹⁰ See Kris Heylen, Thomas Wielfaert, and Dirk Speelman, "Tracking Change in Word Meaning: A Dynamic Visualization of Diachronic Distributional Semantic Models," DGfS 2013 Workshop on the Visualization of Linguistic Patterns, University of Konstanz, Germany, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.724.4739&rep=rep1&type=pdf>.

Helsinki that aims to analyze publication trends in the field of history in early modern Britain and North America from 1470–1800, based on English Short Title Catalogue data.¹¹ The major difference between the work we present here and these other projects is our emphasis on conceptual structure or behavior as opposed to semantic shifts. We do not deny, however, that changes in structure are likely to be congruent with changes in the meanings of terms.

An initial observation of the following kind helps orient our approach: let us say that Hobbes had a theory of liberty which directed his thinking with this idea. Did his fellow citizens mirror or adopt this thinking? Of course, we cannot answer that question with very fine-grained detail since those citizens may have thought about the idea in numerous ways that never fell into print transcriptions of such thinking. Nevertheless, a full-scale survey of the printed text archive does provide us with valuable insights into the ways a culture formulated and used ideas.

To grasp the trajectory of this endeavor, consider the vigorous debate about the two guises of liberty sparked by Isaiah Berlin's 1958 lecture on the "Two Concepts of Liberty."¹² The first, positive liberty, is based upon our freedom to choose what we do. The second, negative liberty, is based upon our acceptance of constraints on how we act—freedom *from* slavery, for example. This debate has a very clear contemporary relevance: it helps us understand our own attempts to work within (or against) received theories of government and democracy based upon "liberalism." Our article concludes with an outline of ways in which computational methods can illuminate the emergence or incubation of such theories, effectively mapping the shifting lexical terrains within which the two terms "freedom" and "liberty" operated in English at the end of the eighteenth century. This kind of research helps us understand how concepts cohere or constellate over time and provides the basis for the articulation of complex political ideas. The evidence leads us to conclude that any close-grained historical account of what has become a contested but nevertheless widely accepted truth—broadly speaking the identification of liberalism's triumph over republicanism,¹³ or more narrowly the interdependence of liberty and individual rights—based on English language sources (as this study is and recognizes as a limitation due to comparability of available datasets) is likely to find

¹¹ See Mikko Tolonen, Leo Lahti, and Niko Ilomäki, "A Quantitative Study of History in the English Short-Title Catalogue (ESTC), 1470–1800," *Liber Quarterly* 25 (2015): 87–116, <https://comhis.github.io/outputs/>.

¹² Isaiah Berlin, *Two Concepts of Liberty* (Oxford: Oxford University Press, 1958).

¹³ See for example Pettit, *Republicanism: A Theory of Freedom and Government*, 41–50.

the last thirty years of the eighteenth century of particular importance. Indeed, a strongly formulated revision to the prevailing orthodoxy notes that the longer *durée* history of the political idea of liberty is likely to pass over the intense work of conceptual formation and adaptation that occurred in this thirty-year period. The sweep of our essay, then, moves from the well-embedded accounts of liberty both historically and philosophically, that is from Cambridge School accounts and the post-Isaiah Berlin philosophical tradition to a data-supported conceptual micro-history that identifies forces active in the last decades of the eighteenth century.¹⁴

It is useful to note that, in our view, “surface” or “distant” reading (the terms that have become established for describing methods of interrogating digital text archives through computational means) is not an end in itself. Indeed, the very terms are misleading because machine modes of data extraction are not in any sense “distant” from the texts to which they are applied: such methods when applied at scale read exponentially closer than humans are capable of doing. Moreover, as the concluding sections of this essay suggest, reading at scale can have the effect of identifying very local effects that otherwise are unperceivable. When we uncover such spikes in a general trend they should be understood as diagnostic with respect to further interrogation of the underlying data.¹⁵ Thus, we propose this essay as an invitation to return to the more generously constructed historical context of our thirty-year period, thereby combining the new methods used here with more conventional modes of the history of ideas. Our hope is to extend and expand the field as it has evolved, not to supplant it. We begin, however, in the pre-history of this emergent political category, “liberalism,” by asking the extent to which the dominant account of two types of liberty (as mapped by Berlin and his interlocutors) might have been recognizable to, say, an English gentleman in 1660.

I. RAW FREQUENCY

This study follows an incremental procedure: first, some rudimentary data extraction from two datasets, followed by more sophisticated parsing of

¹⁴ See Quentin Skinner, *Liberty before Liberalism* (Cambridge: Cambridge University Press, 1998).

¹⁵ For an example of such an inquiry see: Ryan Healey, Ewan Jones, Paul Nulty, Gabriel Recchia, John Regan, Peter de Bolla, “The Uses of Genre,” *Representations* 149 (Winter 2020).

conceptual forms. Ultimately, the evidence yields a clear picture which accords with extant accounts.

We begin our data analysis with a simple inquiry: during the first one hundred years of our dataset, what other words are used in association with the noun *liberty*? And with what frequency? Did agents in the seventeenth century speak of “liberty from servitude”? Did they think of themselves as free from persecution? We searched the EEBO-TCP dataset to find all uses of the phrase *liberty from*. The results indicate that liberty was most commonly understood to be from *sin* (or *sinne* in its variant spelling), a total of 57 occurrences across the seventeenth century. The next most common was *bondage*, a total of 36 occurrences. *Law* (24), *prison* (15), *necessity* (13), *God* (12), *power* (11), *king* (10), *oppression* (9), *tyranny* (9), *imprisonment* (8), and *coaction* (8) are the next most frequent terms used in association with liberty.

For comparison, we ran the same search using the phrase *freedom from*. *Freedom from* was most commonly attached, once again, to *sin* (including *sinne*)—a total of 339 occurrences across the century. And, again, *bondage* was the next most common, with 71 occurrences. *Law* (55), *guilt* (49), *punishment* (47), *death* (38), *arrests* (36), *evil* (35), *curse* (33), *power* (31), *pain* (31), *condemnation* (30), *persecution* (27), *misery* (27), and *trouble* (25) are the next most common.

Beyond *sin*, *bondage*, and *law*, the two phrases had very few other nouns in common. While the terms *liberty* and *freedom* operated in similar ideational terrain, so to speak, their habits of usage were not identical. Their differences are useful in determining the extent to which *liberty* and *freedom* diverge as two distinct ideas or concepts across the two centuries.

Our second data extraction inspects the two variants of the phrases *liberty to* and *liberty from* and compares them with *freedom from* and *freedom to*. The data is presented in the following table:

Table 1	
	1600–1640
liberty to	3143
liberty from	234
freedom to	511
freedom from	1285

Data from EEBO

The data indicates that *freedom from* was far more common than *liberty from*, and, correspondingly, the frequency of the phrase *freedom to* is far

smaller than *liberty to*: 511 occurrences compared to 3,143. This difference clearly marks a distinction in the uses for the two words, and one might begin to hazard that the difference is determined by the positive or negative senses of the concept of liberty. Although one could think of freedom in its positive inflection—freedom to choose what one might do—that conception was much more commonly articulated in the verbal expression *liberty to* do something. Conversely, the negative inflection—liberty from restraint—was more commonly articulated as *freedom from*.

Extracting the data for these uses across the two centuries begins to reveal how the idea of liberty slowly but surely became distinct from the idea of freedom:

Table 2

	1600–1640	1660–1700	1700–1740	1760–1800
liberty from	234	1272	2791	5934
liberty to	3143	19913	68788	154122
freedom from	1285	431	5890	15667
freedom to	511	189	5639	12597

Data from EEBO and ECCO.

These raw frequencies of the phrases indicate clearly that over the two centuries the uses of *freedom* in both the positive and negative senses evens out: while *freedom from* in the early seventeenth century is clearly more common than *freedom to*, by the end of the eighteenth century there is no clear preference. The story with *liberty* is markedly different: *liberty to* is far more common across the two hundred years. If we are to understand liberty as a distinct idea from freedom these data suggest that liberty was articulated in the positive sense: liberty to act as one wished. In the case of freedom, there seems to have been no clear preference for the positive or negative sense.

II. DISTRIBUTIONAL PROBABILITY

In the next data extraction, we used more sophisticated statistical methods for analyzing very large datasets of language use and predicting the likelihood of two terms co-associating. The use of the term “co-associating” is intended to signal that our approach is slightly different from most corpus linguistic studies which use the term “co-occurrence.” First, the use of

“association” underlines our focus on ideas which are said to be linked in the mind through a process of association: when we derive data on words or terms appearing in a text stream at different proximities we mean to be directing attention away from their purely linguistic attributes or functionality, that is away from grammar or syntax, toward an underlying conceptual architecture. Secondly, when we derive data from co-occurrence at wide spans—say one hundred words away—we are not likely to be picking up on grammatical or syntactic coherence. Rather, this kind of stretched linkage may be more likely explained by the fact that the text is concerned with a particular topic. By considering co-association across varying spans of text, we aim to develop a sense of a unit of thinking or understanding that in common speech we call an idea.

Our measure of distributional probability, *dpf*, is created by first observing the raw frequency of occurrence of the target term and then calculating the statistical probability of such a term co-associating with every other term in the dataset. This enables us to create a measure against which we can compare the actual occurrences of every co-associated pair of terms. We generate a numerical value from these calculations, the *dpf*, which serves as an index to the degree to which lexis is statistically co-associated throughout the dataset. This measure can be plotted above a baseline, which is calculated by assuming that the target term could in theory be found in proximity to every other term were that term to be randomly distributed within a string of lexis. It is important to note that our measure is not sensitive to grammar or syntax, which allows us to inspect co-association at large spans or distances between terms. Thus, our tool enables us to inspect spans from close up (five words either before or after the target term) to far away (one hundred words either before or after). The purpose of extracting this data is to glean information about lexical behavior. A pattern of co-association between terms can be understood as a “conceptual signature,” a unique identification for any concept based upon data derived from distributions in lexical use. Most linguistic research that utilizes similar techniques based on neo-Firthian distributional semantics is interested in the features of a language that enable or construct coherence.¹⁶ In work of this kind aimed at understanding conceptual relations, statistical

¹⁶ See John Sinclair, Susan Jones, Robert Daley, and Ramesh Krishnamurthy, *English Collocational Studies: The OSTI Report* (London: Continuum, 2004); Michael Hoey, Michaela Mahlberg, Michael Stubbs, and Wolfgang Teubert, *Text, Discourse and Corpora: Theory and Analysis* (London: Continuum, 2007); and for a review of the field see Tony McEnery and Andrew Hardie, *Corpus Linguistics: Method, Theory and Practice* (Cambridge: Cambridge University Press, 2012).

regularities in grammatical structure are a key component. Our approach differs in that it does not use this method, and does not try to detect relations like meronymy and hypernymy; rather we detect a general association relation from supra-sentential co-occurrences.¹⁷ Thus the co-association data we have captured in the following analysis helps us identify the widest lexical terrain within which a target term operates without regard to immediate syntactic placement or grammatical aspect. By inspecting both close-up and increasingly distant behavior of two co-associated terms, we can begin to assess the strength of “binding” that occurs between any two terms. In this way we can move from strictly semantic or syntactic binding—as in phrases that are common in the English language—to a different kind of binding that we think of as more narrowly “conceptual.” That is, we can start to map the lexical terrain within which ideas circulate and are given shape, structure, and form.

In the following analyses we selected a number of sample decades over the two centuries and inspected the number of highly co-associated terms for a selected target term as distance from the target increases. The first line in the table indicates the number of new terms that appear in the co-association list for any given span and the second line indicates the percentage of terms that are preserved from the previous span. This is the first data for the term *liberty* in the decade 1620–30:

Table 3

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	175	158	125	153	160	132	151	143	164	149
% preserved	12.9	9.2	10.7	5.6	4.8	7.0	5.6	6.5	4.1	8.6

Data from EEBO

Liberty does not preserve a common set of co-associations: at each span increase, new lexis enters the co-association list. Thus, between the distances ten and twenty, 158 new terms appear in the list. Between the distances ninety and one hundred, 149 new terms appear. The percentage report tells the same story from the other angle: very small amounts of lexis are preserved as the distance lengthens. For comparison, consider the report for the later decade, 1680–90:

¹⁷ See J. R. Firth, “The Technique of Semantics,” *Transactions of the Philological Society* 34 (1): 36–73; Sinclair, *Corpus, Concordance, Collocation* (Oxford: Oxford University Press, 1991); and for a good overview of historical semantics see Christian Kay and Kathryn L. Allan, *English Historical Semantics* (Edinburgh: Edinburgh University Press, 2015).

Table 4

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	57	46	41	44	42	51	48	43	42	49
% preserved	37.4	29.2	30.5	21.4	23.6	17.7	15.8	12.2	14.3	9.3

Data from EEBO

Here we can see that the preservation of the same words as distance varies is greater than in the earlier decade. Comparing what happens at close range, between five and ten terms away in the two time segments, we see that the increase from 1620–30 to 1680–90 is marked: 12.9% to 37.4%, and then in the next distance window 9.2% to 29.2%.

Consider the reports from the two corresponding decades in the eighteenth century. This is the data for 1720–30:

Table 5

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	20	15	10	13	7	13	9	6	8	4
% preserved	71.8	70.6	67.7	59.4	70.8	50	59.1	68.4	65.2	76.5

Data from ECCO

In this decade of the eighteenth century we can immediately see a very different pattern: *Liberty* attracts hardly any new co-associated lexis as distance or span increases. Or, put another way, the preservation of the same co-associated lexis is around 70% for most of the distance markers. This report for 1780–90 shows preservation closer to 80% across the spans:

Table 6

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	26	13	14	9	10	12	9	8	7	6
% preserved	74.8	82.4	78.1	82.7	80	77.8	81.6	81.4	83.3	84.2

Data from ECCO

The data provides a clear picture of the shape of binding for the word *liberty* across two centuries. In the early seventeenth century *liberty* operates in a varied lexical terrain. At close span it is bound with a wide variety of other terms, and as distance increases this varied binding persists, adding new lexis to its operational terrain at each distance marker. At the end of the eighteenth century the picture has reversed: the varied pattern of binding is maintained within close spans, but as the span increases the number

and variation of bound terms falls dramatically, indicating that the bound terms are *predominantly the same* as distance increases.

We can compare this binding profile with that of *freedom*. This is the report for 1620–30, using the variant spelling *freedom(e)*:

Table 7

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	253	263	274	267	266	288	266	283	274	272
% preserved	12.8	10.8	11.6	7.9	8.3	5.6	5.7	6.6	7.1	4.6

Data from EEBO

The following report for *freedom* in 1680–90 (the variant spelling had become rare by the end of the seventeenth century) demonstrates a very similar profile:

Table 8

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	121	101	102	110	111	100	85	100	103	94
% preserved	18.8	9.0	10.5	7.6	8.3	9.9	14.1	9.9	8.0	8.7

Data from EEBO

Like *liberty* throughout the seventeenth century, *freedom(e)* binds with a wide range of other terms, and as distance increases it attracts new and different terms. A more complete picture emerges with comparison to the eighteenth century. This is the data for 1720–30:

Table 9

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	130	97	78	80	95	79	83	72	63	63
% preserved	24.4	21.1	19.6	14.0	10.4	9.2	12.6	13.3	14.9	12.5

Data from ECCO

And this is the data for 1780–90:

Table 10

Span	5–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
No terms	82	82	45	50	42	38	32	31	35	20
% preserved	50.3	41.0	51.6	41.2	40.8	44.1	44.8	45.6	30.0	44.4

Data from ECCO

The preservation of terms bound over distance increases diachronically but to a lesser degree than for *liberty*. Where *freedom* has a maximal preservation between any distance marker of 51.6% in the date range 1780–90, “liberty” preserves 84.2% of bound lexis in the same date range. These initial data lead us to suppose that the behavior of our target terms alters or comes under pressure during the last decades of the eighteenth century. This is developed further in section six. Another measure helps us even better understand the difference between the two terms with respect to change over time. This table compares the preservation of lexis at the same distance (ten terms apart) between two dates in the eighteenth century, 1720–30 and 1780–90, for each of the two terms:

Table 11

	liberty	freedom
% preserved	70.40%	22.70%
no terms	50	24

Data from ECCO

A consistent pattern has emerged: *liberty* operates in a very stable lexical terrain over the course of the eighteenth century. The opposite is the case for *freedom*.

III. THE CONVERGENCE OF THE IDEAS OF “LIBERTY” AND “FREEDOM”

Thus far, we have identified differences in the behavior of the two terms *liberty* and *freedom* across the seventeenth and eighteenth centuries. This difference in lexical behavior appears to map onto a difference in conceptual structure: although they share a lexical terrain—and as shall become evident below, these terrains converge by the end of the eighteenth century—*liberty* and *freedom* nevertheless appear to be structured differently as concepts. This difference and convergence are clearly observable if we track the common co-associated lexis at distances between five and one hundred over the two centuries. The following table provides the data for the overlap between the co-association lists for the two terms. The number of co-associated terms on these lists is given in columns two and three, and the number of terms that are common to both lists is given in column four.

Table 12

	liberty 1600–40	freedom 1600–40	No shared
D:5	183	443	30
D:10	116	443	10
D:50	106	435	5
D:100	101	453	2

Data from EEBO

This data can be compared with the later time segment:

Table 13

	liberty 1660–1700	freedom 1660–1700	No shared
D:5	177	179	54
D:10	101	98	21
D:50	41	76	6
D:100	39	67	4

Data from EEBO

There is a very narrow common terrain between the two terms. At best they share 54 terms.

The following tables present data for the eighteenth century, first the earlier date range 1701–40:

Table 14

	liberty 1701–40	freedom 1701–40	No shared
D:5	276	295	105
D:10	113	180	45
D:50	55	91	19
D:100	44	51	11

Data from ECCO

It is important to note that at the distance of one hundred terms, the 11 common words in the co-association lists contain five “stop” words: *the*, *and*, *to*, *of*, and *that*. Thus, there are only six terms of any significance in common.

The following table presents the data for the last forty years of the eighteenth century:

Table 15

	liberty 1760–80	freedom 1760–80	No shared
D:5	383	451	175
D:10	222	275	99
D:50	133	137	55
D:100	114	84	43

Data from ECCO

By the end of the century the overlap in lexical terrain between *liberty* and *freedom* has been transformed: the 43 terms that are held in common between the co-association lists at distance one hundred represent 51.2% of all of the terms on the list for *freedom*. Or, from the other perspective, the 175 common terms at distance five comprise 45.6% of the terms in the list for *liberty*. The lexical terrain in which the two terms operated had converged by the end of the eighteenth century.

IV. COMMON BOUND LEXIS

The stability of the lexical terrain within which “liberty” operates might be considered in terms of a network or constellation of terms that together compose the circumscribed semantic space that we call an idea. The code developed by the Cambridge Concept Lab enables us to drill down further into these spaces and discover the tight lexical networks operating in our datasets.¹⁸ And, given the fact that we can derive this data chronologically, we can track how these networks change over time. In the following data analysis we have constructed the network by identifying which terms are in the other’s lists of bound terms, thereby isolating the common bound lexis to all the terms in the network. Such networks or cliques are generally not large, containing no more than a handful of terms—on account of the rule that each term must be on each other’s list. This is indeed borne out by the data.

In the early seventeenth century, 1630–40, *liberty* can be found on the binding list of six other terms, each of which also contain the other terms in the set of seven terms. These terms are *liberty*, *bondage*, *freedom*, *slavery*, *thralldome*, and *freed*. In the later decade, 1690–1700, the largest set within which *liberty* operates is six terms: *thralldom* (in the modern variant spelling), *bondage*, *freedom*, *liberty*, *slavery*, and *free*. Once again note the stability of

¹⁸ The code is available at The Concept Lab, <https://concept-lab.lib.cam.ac.uk>.

this lexical terrain. The picture changes at the far end of the eighteenth century. In the period 1770–1800 *liberty* is a member of 82 sets of eight terms, and the six most strongly associated are a set of variations on the following terms: *anarchy*, *aristocracy*, *democracy*, *government*, *liberty*, *monarchy*, *republican*, *tyranny*, *equality*, *revolution*, *republic*. Interestingly, however, the profile for *freedom* is very divergent. In the same time period, the last decades of the eighteenth century, the largest set size within which this term appears is six, and there is only one such set: *democracy*, *freedom*, *government*, *liberty*, *revolution*, *tyranny*. Once again our earlier data analyses are confirmed: the uses of *liberty* and *freedom* converge over the two hundred years, and the tight lexical terrain within which *liberty* operates has, by the end of the eighteenth century, become very evidently established. Whereas the seventeenth century thought of liberty in conjunction with slavery—that is, it conceived of liberty as an adjunct of person—by the end of the eighteenth century liberty had become an adjunct of the state.¹⁹

V. LIBERTIES AS RIGHTS

What contribution can the preceding computational and statistical approach to the history of ideas make to the long tradition of inquiry into the foundations of the modern concepts of freedom and liberty that underlie the contemporary understanding of liberalism? For answers, we must look to the scholarship of Quentin Skinner, particularly his tenacious and compelling work on the genealogy of the concept. As is well known, Skinner began his long career as a scholar in the late 1960s when, immersed in the traditions of thinking about modern political concepts, he presented his Cambridge lectures that were the basis for *The Foundations of Modern Political Thought*. But it was in the 1980s that he turned most consistently to the historical reconstruction of the various traditions of thinking that developed the idea of liberty.²⁰

In his 1984 essay “The Idea of Negative Liberty,” Skinner gives an historical account of two opposing ideas. One is “negative liberty,” in which the individual’s social freedom is guaranteed only by the absence of

¹⁹ As Skinner notes, Hobbes was the first thinker to effect this change by constructing the state as a particular kind of person. See Skinner, *Liberty before Liberalism*, 4–5.

²⁰ The best account of the development of Skinner’s thought at this time is Marco Guena, “Skinner, pre-Humanist Rhetorical Culture and Machiavelli,” in Annabel Brett and James Tully, with Holly Hamilton-Bleakley, *Rethinking the Foundations of Modern Political Thought* (Cambridge: Cambridge University Press, 2007) 50–72, esp. 64–69.

limiting factors such as state intervention, responsibilities to one's communities, and other externalities. In this scheme, liberty can be defined only negatively, as Thomas Hobbes has it at the start of his chapter "Of the Liberty of Subjects" from *Leviathan*: "liberty or freedom signifieth (properly) the absence of opposition."²¹ Skinner contrasts negative liberty with an ideal of liberty in which the operative factor is the virtue and value of public service. According to this view, to be free, individuals must acknowledge their social responsibilities and carry out virtuous acts of public service. These contrasting ideas of liberty are named by Canadian philosopher Charles Taylor as the "opportunity concept" and the "positive exercise concept."²² The former relies purely on the absence of constraint and prescribed social objectives (freedom from), whereas the latter involves positive action in the service of the state or community (freedom to).²³ Skinner sets out to demonstrate that the early modern period *combined* these two notions of liberty, writing, "I shall try to show that, in an earlier and now discarded tradition of thought about social freedom, the negative idea of liberty as the mere non-obstruction of individual agents in the pursuit of their chosen ends was combined with the ideas of virtue and public service in just the manner nowadays assumed on all sides to be impossible without incoherence."²⁴ In his inaugural lecture as Regius Professor of Modern History in 1997, Skinner returned to this material and subsequently published a short book on the topic entitled *Liberty before Liberalism*.²⁵ Once again he stressed the combination of negative and positive liberty in the neo-Roman tradition, which he claims was dominant in political discourse in

²¹ Thomas Hobbes, *The Clarendon Edition of the Works of Thomas Hobbes*, vol. 4, *Leviathan: The English and Latin Texts*, ed. Noel Malcolm (Oxford: Clarendon Press, 2012), 324.

²² See Charles Taylor, *Philosophy and the Human Sciences: Philosophical Papers*, vol. 2 (Cambridge: Cambridge University Press, 1985).

²³ The classic account of this distinction remains Berlin, *Two Concepts of Liberty*.

²⁴ Quentin Skinner, "The Idea of Negative Liberty," in *Philosophy in History: Essays on the Historiography of Philosophy*, ed. Richard Rorty, J. B. Schneewind and Quentin Skinner (Cambridge: Cambridge University Press, 1984), 197.

²⁵ Skinner, *Liberty before Liberalism*. The topic has, of course, been deeply embedded in much of Skinner's work. See, for example, Skinner, *The Foundations of Modern Political Thought*; Skinner, "Machiavelli on the Maintenance of Liberty," *Politics* 18 (1983): 3–15; Skinner, "The Paradoxes of Political Liberty," in *The Tanner Lectures on Human Values* (Cambridge: Cambridge University Press, 1985), 227–50; Skinner, "Pre-Humanist Origins of Republican Ideas," in *Machiavelli and Republicanism*, ed. Gisela Bock, Skinner, and Maurizio Viroli (Cambridge: Cambridge University Press, 1990); Skinner, "The Republican Ideal of Political Liberty," in Bock, Skinner, Viroli, *Machiavelli and Republicanism*; Skinner, "Thomas Hobbes on the Proper Signification of Liberty," *Transactions of the Royal Historical Society* 40 (1990): 121–51; Skinner, *Reason and Rhetoric in the Philosophy of Hobbes* (Cambridge: Cambridge University Press, 1996).

England immediately following the regicide in 1649. His aim, in both this short book and the original essay published in 1984, is to revise, even dissolve, our modern assumption that liberty is incoherent outside a theory of rights. Early modern republican writers, he insists, understood liberty from constraint within the context of behavior that was based in notions of virtue and public service.²⁶

If this sense of liberty was prevalent during the period—that is liberty as harnessed to or articulated around notions of virtue—one would expect the lexical terrain of the two terms to have intersections or commonalities. Such a common terrain derived from co-association data could, of course, be either supportive or critical of the notion that Skinner proposes—that is, co-association in and of itself does not come with an index to the senses in which terms qualify each other. Notwithstanding this caveat, while Skinner’s reading of the classic texts—those by Harrington and Sidney prime among them—certainly makes a convincing case, the extent to which this neo-Roman account of liberty penetrated the culture needs to be assessed. A first pass through the datasets we have been using suggests that the overlapping lexical terrain between *liberty* or *freedom* and *virtue* was negligible. In the following table we have tracked this overlap across the two centuries by creating *dpf* lists for the terms at a span of ten terms and have included data for the variant spellings *libertie* and *freedom*²⁷:

Table 16

	1620–30	1680–90	1720–30	1780–90
liberty/virtue % shared	0 (0)	0 (0)	7% (5)	4.9% (5)
libertie/virtue	0 (0)	0.2% (1)		
freedom/virtue	2.1% (5)	0.7% (1)	1.7% (3)	3.6% (6)
freedom/virtue	1.7% (5)	0.7% (1)		

Data from EEBO and ECCO

²⁶ Skinner returned to this theme in his London Review of Books lecture “A Third Concept of Liberty,” subsequently published in *London Review of Books* 24, no. 7 (April 4, 2002). There is also a large literature that engages with his argument. See, among others, Pettit, *Republicanism: A Theory of Freedom and Government*; Phillippe Van Parijs, *Real Freedom for All* (Oxford: Oxford University Press, 1995); Daniel Weinstock and Christian Nadeau, eds., *Republicanism: History, Theory and Practice* (London: Frank Cass, 2004); Cécile Laborde and John Maynor, eds., *Republicanism and Political Theory* (Oxford: Blackwell, 2008); Charles Larmore, “Liberal and Republican Conceptions of Freedom,” in Weinstock and Nadeau, *Republicanism: History, Theory and Practice*.

²⁷ The numbers in brackets indicate the total number of terms that are common to both *dpf* rank lists.

At a proximate distance, here ten terms apart, one would expect to pick up semantic behavior, but as the data indicates, the lexical terrain within which *liberty* and *freedom* operated did not have strong connections to the terrain within which *virtue* appeared. At the longer span of one hundred terms, where we expect to find a different kind of binding, the story is substantially the same:

Table 17

	1620–30	1680–90	1720–30	1780–90
liberty/virtue % shared	0.6% (1)	0 (0)	23.5% (4)	13.2% (5)
libertie/virtue	0.5% (1)	0 (0)		

Data from EEBO and ECCO

This data seems to contrast with Skinner’s argument, at least insofar as he supposes the neo-Roman account to have widespread currency. But it also supports another strand of his thesis which points out the virtual hegemony of a Hobbesian “negative liberty.”

If we drill down further into the datasets and create similar reports for some candidates for specific virtues the story is pretty much the same. This table reports a similar analysis run on the terms *benevolence*, *magnanimity*, *charity*, *generosity*, and *virtue*, this time across fifty-year segments of the two centuries:

Table 18

Number of shared terms	1600–50	1650–70	1700–50	1750–1800
liberty/benevolence	4	0	2	11
freedom/benevolence	13	3	15	26
liberty/magnanimity	1	1	4	19
freedom/magnanimity	8	2	10	31
liberty/charity	0	0	2	6
freedom/charity	4	3	1	11
liberty/generosity	4	1	3	7
freedom/generosity	8	5	11	21
liberty/virtue	1	0	2	13
freedom/virtue	0	0	0	0

Data from EEBO and ECCO

These data indicate that these virtues—represented here by the lexis that designates them—were not considered to occupy the semantic space as either *liberty* or *freedom*. If we inspect the actual terms that appear in these lists—remembering that the number of terms here is very small and entering due caution with respect to generalizations from such sparse data—another strand of Skinner’s argument moves into view. The three terms that appear on the co-association lists for *benevolence*, *magnanimity*, and *generosity* in the seventeenth century are *slavery*, *servitude*, and *arbitrary*. And these terms fall out of the lists in the eighteenth century. It is also noteworthy, given Skinner’s characterization of Hobbes’s difficulty in reconciling a theory of the state as person with the idea of negative liberty, that the following two terms enter these lists: *volition* and *rights*.

We believe this to indicate that the forces which fuse rights to liberty really only began to have effects within the conceptual architecture toward the end of the eighteenth century. For Skinner the *longue durée* account is more persuasive as he draws out the implications of the “Hobbesian claim that any theory of negative liberty must in effect be a theory of individual rights.”²⁸ In contrast we see the tectonics underlying the formulation of a linked or constellated set of terms which contribute to a theory of liberty in a slightly broader perspective outlined below. Let us stay with Skinner’s point as way of sharpening that observation: note that he claims that “*any* theory of negative liberty” must be congruent with, even inserted within, a theory of rights. As we have noted above, Skinner is certainly correct in stating that the Hobbesian version of negative liberty quickly became hegemonic and that our history of this idea is to some extent a history of forgetting, of the erasure of different ways of thinking about that idea. Noting the linkage of negative liberty and rights, he writes:

As we have seen, this has reached the status of an axiom in many contemporary discussions of negative liberty. Liberty of action, we are assured, “is a right”; there is a “moral right to liberty”; we are bound to view our liberty both as a natural right and as the means to secure our other rights. As will by now be obvious, these are mere dogmas. A classical theory such as Machiavelli’s helps us to see that there is no conceivable obligation to think of our liberty in this particular way. Machiavelli’s is a theory of negative liberty, but he develops it without making any use whatever of the concept of individual rights.²⁹

²⁸ Skinner, “Negative Liberty,” 218.

²⁹ Skinner, “Negative Liberty,” 218.

Inspection of the data reveals the accuracy of this account of the fusion of liberty and rights. When we tracked the shared lexical terrain between *liberty* and *rights* across the two centuries, we found this data for the percentage of common co-associated lexis at a distance of ten terms:

Table 19

	1620–30	1680–90	1720–30	1780–90
liberty/rights	1.5% (3)	4.7% (4)	25.4% (18)	42.7% (44)

Data from EEBO and ECCO

And if we inspect the data for a longer span of one hundred words, where we pick up what we consider to be weak syntactic binding and stronger conceptual binding, the picture is even clearer. Thus, the same inquiry but at the co-association at distance one hundred yields the following:

Table 20

	1620–30	1680–90	1720–30	1780–90
liberty/rights	1.2% (2)	0 (0)	35.3% (6)	63.2 (24)

Data from EEBO and ECCO

The data indicates that by the end of the eighteenth century the Hobbesian version of negative liberty was, effectively, the only game in town.

VI. FROM LIBERTY TO LIBERALISM

The data extraction presented thus far indicates that the theory of liberty based upon positive individual rights—what Skinner describes above as “liberty of action”—slowly emerged during the eighteenth century, no doubt framed by practical political action resulting from the two large constitutional events of the second half of the century: the war with the colonies and the British reaction to the French revolution.³⁰ In broad brushstrokes these forces have generally been examined within a longer time-frame that observes a European shift in political conceptual sensibility,

³⁰ See in particular in the vast literature on these topics J. C. D. Clark, *The Language of Liberty 1660–1832: Political Discourse and Social Dynamics in the Anglo-American World* (Cambridge: Cambridge University Press, 1994); Michael Zuckert, *The Natural Rights Republic: Studies in the Foundation of the American Political Tradition* (Notre Dame, IN: University of Notre Dame Press, 1996); Craig Yirush, *Settlers, Liberty, and Empire: The Roots of Early American Political Theory, 1675–1775* (Cambridge: Cambridge University Press, 2011); Pamela Clemit, ed., *The Cambridge Companion to British*

from roughly speaking a late seventeenth century formulation of republicanism to what Skinner takes to be a hegemonic modern concept of liberalism based on subjective rights. Liberalism replaces republicanism, and both are seen as opposed or antagonistic to each other. As J. G. A. Pocock notes, the tradition of republicanism is based upon a completely different set of principles and vocabulary from what emerged in the nineteenth century as the classic account of liberalism.³¹ Such a reading is no doubt supported by selective consideration of the major philosophical and political texts within this long period. But when we take a more holistic view from the position of the aggregated archive, another model for the establishment of modern liberalism becomes discernible. This account sees liberalism as effectively the genetic mutation of liberty as it tries to accommodate republicanism. And contrary to the longer historical sweep of a pan-European tradition of republicanism, our data analyses based on ECCO suggest a much narrower time scale in which a more forceful conceptual shift occurred between the 1770s and the end of the century when English language attempts to wrestle with or adapt and alter the concept of republicanism succeeded in transforming the idea of liberty. Republicanism was, effectively the catalyst for liberalism. The data, therefore, not only supports the revision to the Skinnerian account proposed by Andreas Kalyvas and Ira Katznelson, it also allows us to track with considerable granularity the decisive expansion of the lexical terrain at the core of the concept of liberty, essentially providing a window onto the insertion of the idea of republicanism within liberty. And this composite political idea, we contend, provided the means for the rapid development of what has come to be one of the most consequential Western political concepts since the nineteenth century: liberalism.³²

Using the same techniques for ascertaining distributional probability outlined in section two above we can create a conceptual signature we call a “core.” This is determined by aggregating the co-associated lexis for a target term at three distances: ten, fifty, and one hundred words both before and after the focal term. This concentrates the more populated networks within which a term operated at any time segment in the dataset so as to identify what persists, what lies at the heart of the concept. As our analyses

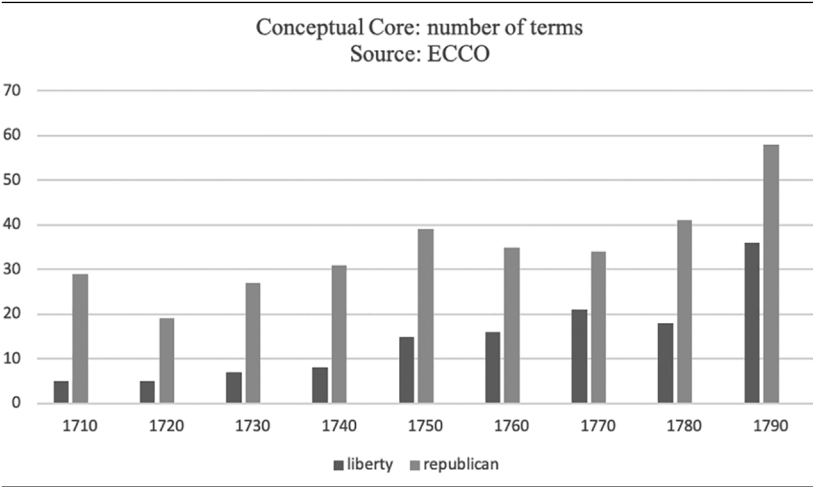
Literature of the French Revolution in the 1790s (Cambridge: Cambridge University Press, 2010).

³¹ J. G. A. Pocock, “Virtues, Rights, and Manners: A Model for Historians of Political Thought,” in *Virtue, Commerce, History* (Cambridge: Cambridge University Press, 1985).

³² See Andreas Kalyvas and Ira Katznelson, *Liberal Beginnings: Making a Republic for the Moderns* (Cambridge: Cambridge University Press, 2008), esp. 5–17.

above have already indicated, *liberty* is a very stable term over the eighteenth century, and its core comprises the following four terms until the 1750s: *slavery*, *volition*, *tyranny*, *freedom*. Five more terms enter into the core before that decade: *servitude*, *toleration*, *free*, *government*, and *licentiousness*. The story for *republican* is very different as can be seen from the following chart which plots the core for *liberty* against that of *republican* with respect to the number of co-associated terms that are common across the three distances:³³

Table 21



Data from ECCO

³³ The data for the adjectives “liberal” and “republican” tell a similar story. The core for liberal across the decades of the century never amounts to more than 11 terms, whereas the core for republican is 61 terms in the final decade and averages 33 terms.

But it is not simply the fact that the core of *liberty* began to increase in the decade of the 1750s; it is the overlap in lexical terrain between the two conceptual cores of *republican* and *liberty* that provides insight in this process of conceptual re-engineering. Here is the table of common terms from 1750:

Table 22

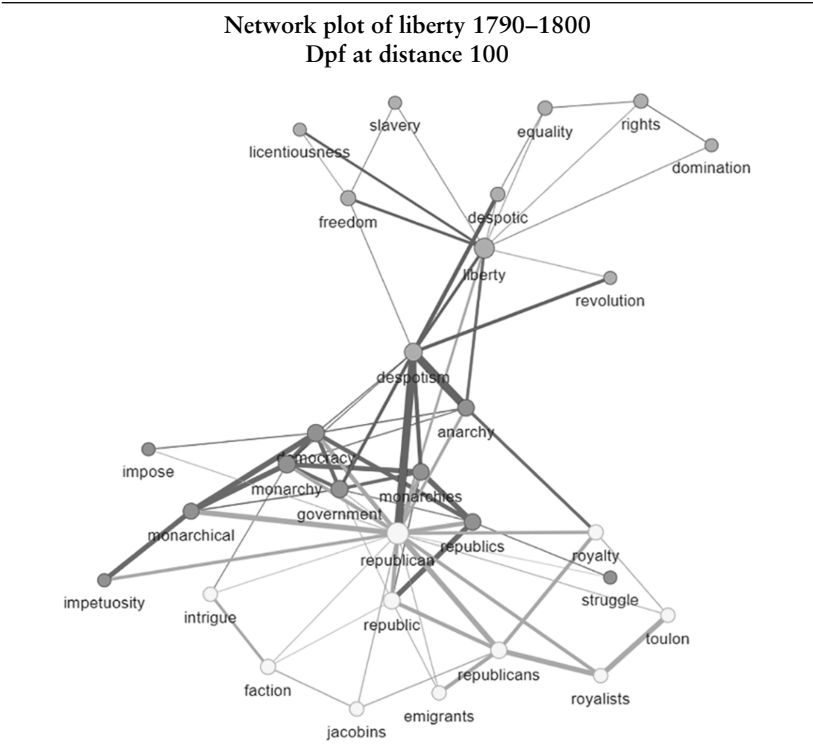
overlap cores liberty/republican	1750–60	1760–70	1770–80	1780–90	1790–1800
	liberty republican government	despotic government liberty	government despotism constitution laws despotic republican liberty	arbitrary constitution government political independence legislative	despotism rights tyranny tyrants constitution government equality republican revolution anarchy despotic citizens governments republic people monarchy convention national citizens
% of liberty's core	18.00%	17.60%	30.40%	33.30%	51%

Data from ECCO

The data clearly indicate that from the 1770s on, the idea of liberty, which for over a hundred years had remained stable and resistant to mutation, began to alter under pressure from the attempts within British political theory and debate to conceive of republicanism in a modern dress. This is borne out by the fact that for the first half of the eighteenth century there is no common lexical core shared by *liberty* and *republican* (the one term in common by 1740 is *government* which persists as the single term through the 1750s) and that by the end of the century 51% of *liberty*'s core is held in common with *republican*'s. The following map, based upon the same *dpf* information but now expressed within a network graph, indicates that this effort was in large part coincident with the attempts to understand or negotiate the concept of despotism, a word that first appears in English in 1708

but was hardly used for the first fifty years of the century, occurring only 189 times in all English printed text up until 1750.³⁴ During the last decade of the century it appears over 14,000 times.³⁵

Table 23



³⁴ This observation is based on the ECCO dataset which does not capture *all* lexical use across the Anglophone eighteenth century, so to this extent the claim is subject to qualification. For the first use see Rev Thomas Cooke, *The Universal Letter-Writer; or, New Art of Polite Correspondence* (London: A Millar, W. Law and R. Cater, 1708), 123. It is also worth noting that words with the suffix “ism” are rare in the period. The most common across the century are despotism, atheism, patriotism, fanaticism, and paganism.

³⁵ Melvin Richter dates the modern resurgence of “despotism” to Montesquieu’s *De l’Esprit des lois* (1748), which prefers the term to the then-current tyranny. See Melvin Richter, “A Family of Political Concepts: Tyranny, Despotism, Bonapartism, Caesarism, Dictatorship, 1750–1917,” *European Journal of Political Theory* 4 (2005): 221–48. We have discussed this at length in a second paper, “The Conceptual Foundations of the Modern Idea of Government in the British Eighteenth Century: A Distributional Concept Analysis,” currently under review.

Here we see that the re-engineering of the idea of liberty so as to include republicanism occurs by inserting despotism into its architecture. Seen from the other side, liberty can be thought of as resisting the insertion of despotism into its conceptual architecture. In this way, we suggest, the modern conception of liberalism held republicanism at bay. The history of these two ideas from the nineteenth century to the present day confirms this.

University of Cambridge.

Copyright of Journal of the History of Ideas is the property of University of Pennsylvania Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.