



Decoding British Empiricism:

A Digital Text Analysis of Locke, Berkeley, and Hume

Digital Companion Piece:
overt-analysis.com/DBE/DecodingBritishEmpiricism.html

Jason Bradshaw ~ Philosophy/Digital Humanities



Introduction

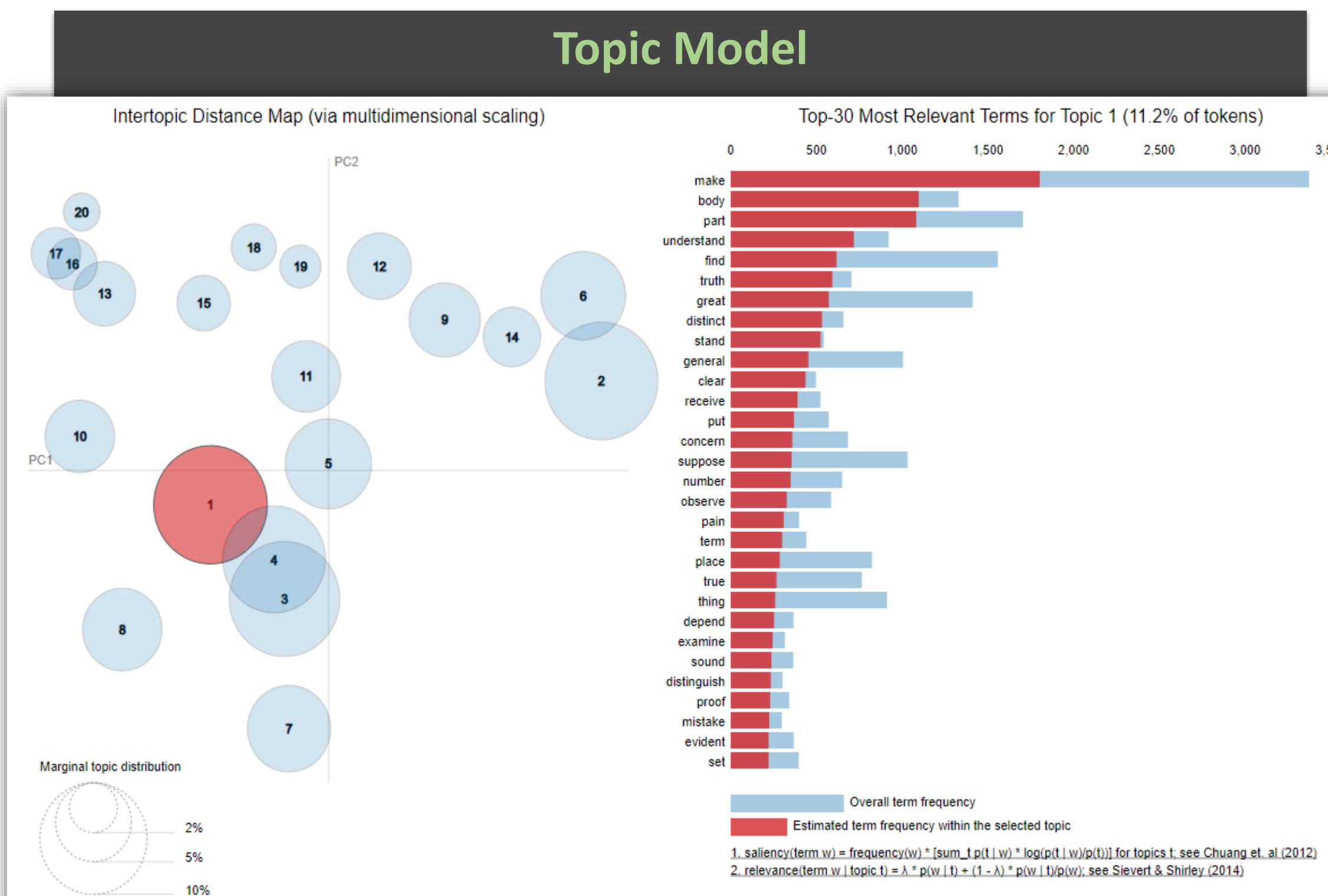
Decoding British Empiricism is the culmination of nearly a years worth of research, reading, coding, and more reading. This ongoing thesis work started in September of 2018 and is slated for completion by August, 2019. Presented here are some of the findings made after an initial pass of the text analysis results from the British empiricist corpora. For an extended version of these findings and more, please refer to the digital companion piece website. The full corpus is comprised of selected works from John Locke, George Berkeley, and David Hume—the original British empiricist trifecta.

The project started with a modest enough question: *where are all of the Digital Humanities philosophers?* This question is based upon my own, admittedly limited, experience with the Digital Humanities (DH) field. However, the sentiment has also been voiced by the likes of DH scholar, Lisa Spiro (digitalscholarship.wordpress.com/2013/02/26/exploring-the-significance-of-digital-humanities-for-Philosophy/), assistant professor of Philosophy, Laura Kane (digitalfellows.common.gc.cuny.edu/2014/10/08/no-room-for-digital-humanities-in-Philosophy/), and McDaniel College's associate professor of Philosophy, Peter Bradley (chronicle.com/blogs/profhacker/where-are-the-philosophers-thoughts-from-thatcamp-pedagogy/37408). This is not to suggest that Philosophy is completely absent from the field. There are a number of exciting DH projects that focus on the discipline. For example, Stanford University's Mapping the Republic of Letters (republicofletters.stanford.edu/index.html), the collaborative Homer Multitext Project (homermultitext.org/), and the dynamic ontologies created through The InPho Project (inphoproject.org/). While it's true that these and other important philosophical projects exist, there is no denying that Philosophy as a discipline is underrepresented at DH conferences. This is something that you may even notice at the conference that you are currently attending—Philosophy projects are few and far between.

Decoding British Empiricism does not attempt an answer at why it is that Philosophy is currently underrepresented in the field. This is a question with too many discipline specific factors to answer adequately. However, the question does provide sufficient justification to revisit some classical philosophical literature and introduce it to the 21st century by means of modern text analysis techniques. It will be shown that our understanding of British empiricism can be enhanced with the addition of a digital textual analysis. The techniques provide further evidence and supplement theories that have been put forth in prior close readings of the material, as well as producing a few new hypotheses of their own. *Decoding British Empiricism* may not be able to answer where the DH philosophers are, but the project undoubtedly shows that there is ample room for philosophers and their projects in DH.

* * *

A quick note on method. The majority of the visualizations and data analysis were conducted via Jupyter Notebooks, a browser-based GUI for Python 3. There are, however, a couple of exceptions. The British empiricist's Top 10 word frequency counts were created in Photoshop with the use of Python data. This was an aesthetic choice. There was really no way to visualize the data how I envisioned it using only Python. The topic model data was made possible by using MALLET on the command line, which was then visualized in the Jupyter Notebook environment. There are a number of modules for compiling topic models within Python, but none of them come close to the power of the models produced by MALLET. The complete methodology can be found in the *Decoding British Empiricism* thesis paper that should be publicly available by late September, 2019.



Empiricism

What first makes an impression from the model is how accurately the first 3 topics ("Physical Bodies", "Ideas", "Epistemology") reflect the shared empirical goals of Locke, Berkeley, and Hume. These term lists express the attainment of knowledge through bodily senses that then inform and create ideas in the mind. There is something to be said about the spaces represented by the pyLDAvis intertopic distance map as well. **Physical Bodies** and **Ideas** are separated significantly. This is appropriate given the distinction made by the empiricists between the mind and the body. While Locke may have rejected rationalist innatism due to his conviction "that all ideas are acquired from experience," he did have to concede to the mind-body dualism (a distinct separation between the processes of the mind and the physical body) popularized by rationalists like René Descartes (Carlin, 2009, p. 83). This is perfectly reasonable. Mind-body dualism was a result of the mechanistic view of the universe that was taking hold during this time period and it heavily influenced both rationalism and empiricism.

Looking again to the intertopic distance map, **Physical Bodies** and **Epistemology** are relatively close. This makes sense given that empiricism appeals to sensory experience as its primary source of knowledge. Unlike mind-body dualism, this aspect of empiricism is in stark contrast to the epistemology of the rationalists, who instead express "the view that reason (as opposed to sensory experience) is the ultimate justification of our knowledge" (Carlin, 2009, p. 1). It is interesting that these 3 topics take precedence as the full corpus included texts like dialogues, political essays, and moral philosophies. It suggests that empiricism played a large role in all of these author's philosophical endeavours.

Political Philosophy

Another category of topics that are prominent in the model are those relating to politics. **Rule**, **Ancient Politics**, **Geopolitics**, **Power Dynamics**, **Establishments**, and **Governmental Structures** all contain term lists that are highly politicized. These results do make sense given the scope of the full corpus. At least half of the empiricist's documents contain philosophical discussion on the government and politics of the time. Locke's *Two Treatises of Government* make up about a third of the philosopher's total word count contribution to the corpus. In the text, Locke's (2003) main tenet of governance stems from the expression of God's will made manifest through the Law of Nature: "Conceived of as a law (the law of nature), or almost as a power, it is sovereign over all human action. It can dictate to a man as conscience does and to more than one man in the social situation, since it is given by God to be the rule betwixt man and man" (p. 95). Locke's doctrine of the Law of Nature is reflected by the **Rule** term list ("power", "father", "man", "god", "children", "law", "authority", "dominion", "society", "nature") and **Power Dynamics** ("govern", "master", "death", "public", "live", "belong", "submit", "lay", "force", "possession"). Hume's *Essays, Moral, Political, and Literary* make up the majority of the philosopher's contribution to the full corpus. Although not strictly political in nature, there are many references to Hume's own ideas regarding governance. The term list for **Ancient Politics** contains words that are present in several of Hume's political essays from the anthology. In particular, Essay III ("That Politics may be Reduced to a Science") holds a number of references to the governance of medieval Europe, as well as the absolute rule and conquest by ancient Rome.

Moral Philosophy

Hume's moral Philosophy finds a voice in 2 of the topics: **Morality** and **Human Nature**. The Enlightenment revolutionized many perceptions and attitudes towards the external world. Copernicus revolutionized the way people saw their place in the universe after announcing that the earth moves around the sun, Newton's laws of motion opened the doors for a mechanistic world view, and some have argued that Hume marked the beginning of modern moral Philosophy. Nicholas Capaldi (1989) notes the shift from the belief in an external source of morality to an inward source during this time (p. 2). Modern moral philosophers debated over the faculty to which this source of moral insight belonged. One camp believed that morality was derived from reason, while the other believed that it originated with sentiment. Hume straddled the line between these two views, along with Joseph Butler and Francis Hutcheson, all of whom "opted for some combination but one in which sentiment was the dominant element" (Capaldi, 1989, p. 3). We see this view reflected in Hume's moral topics. "Sentiment" makes its way into the lists for both **Morality** and **Human Nature**, along with a number of associated terms like "passion", "moral", "feel", and "source". Capaldi (1989) notes that Hume was opposed to the utilitarianism favoured by some of his predecessors, like Thomas Hobbes, and voiced a "persistent rejection of the move to reduce morality to self-interest" (p. 6). There are echoes of this opinion in the topics, through terms like "utility" and "happiness". Hume rejected self-interest as the source of morals for a view of morality as a developmental process, that was both natural and artificial. Morals are artificial because they are "partly conventional, being social products over time" and they are natural due to our "capacity for sympathetic identification which can reconcile conflicts among self-interest, moral motives, and the social good" (Capaldi, 1989, p. 6). The emphasis placed by Hume on these artificial and natural origins are represented by topic terms like "interest", "character", "laws", "society", "situation", and "social".



One thing that is immediately apparent from this visualization is that LRI measures are not particularly useful on their own. There is clearly a distinction between the 3 philosophers. Hume has the highest lexically "rich" corpus, while Locke has the lowest, and Berkeley falls somewhere in between. But how does this discrepancy actually reflect their respective vocabularies? To answer this question, we need to turn to the word frequency counts. These counts play a pivotal role in the formation of LRI measures, as well as deciphering what the measures mean. For example, Locke, who has the lowest scoring LRI measures, also has the highest percentages of terms making up his total word counts. The 10th word in Locke's full corpus word frequency count ("word") makes up 1.82% of the entire corpus. Compare this to the empiricist with the highest LRI measure values, David Hume. Hume's first word in his full corpus frequency count ("reason") only comprises 1.50% of his entire corpus, a value lower than Locke's 10th term. Indeed, these types of percentages are consistent across each of the empiricist's documents. By no means does this suggest that Hume's writing is superior to Locke's in any way. Rather, it reflects stylistic differences. The LRI trajectory is intriguing. Locke's contributions to the corpus were written towards the end of the 17th century, Berkeley's around the beginning of the 18th century, and Hume's at the tail end of the 18th century. It implies a direct correlation between time period and LRI value—it increases in the corpus over time. There are some possible explanations for why we might see such a trend:

- The empiricist's (excluding Locke) had the gained advantage of their predecessor's material to draw upon and structure their own arguments against. Borrowing vocabulary could have led to increased LRI scores.
- It is a condition of the level of education each of the empiricist's received during their respective time frames.
- It is a result of the geographical region where they received their education. Locke and Berkeley both finished their formal educations in their respective homelands (England and Ireland), while Hume split his time between the university of Edinburgh in Scotland and the Jesuit college in La Flèche, France.

Another general finding worth mentioning from the word frequency counts is the prevalence of masculine language and pronouns. Words like "man" and "men" feature predominately in each of the empiricist's word frequency counts. This is not surprising given the time period that these major philosophical works were written. The 17th and 18th centuries were not particularly observant of the roles and needs of women. Even in the 21st century women and visible minorities all too often lack the respect that they deserve.

Multiple Correspondence Analysis (MCA)

The philosopher's main empirical documents are all clustered fairly close together in the center of the distribution point cloud, with minor deviations towards particular general inquirer categories. However, the distances are not significant and more likely reflect slight individual differences in the philosopher's conceptions of empiricism. The distance provides evidence in support of the similarities between the epistemological writing of the British empiricists. They also point to a trend that is seen with the rest of the documents. The majority of Locke's documents tend towards power, Berkeley's towards negative categories, and Hume's towards the positive categories. Indeed, each of the philosopher's have documents that define the extremes of the cloud in these directions.

Locke is Inclined to Power

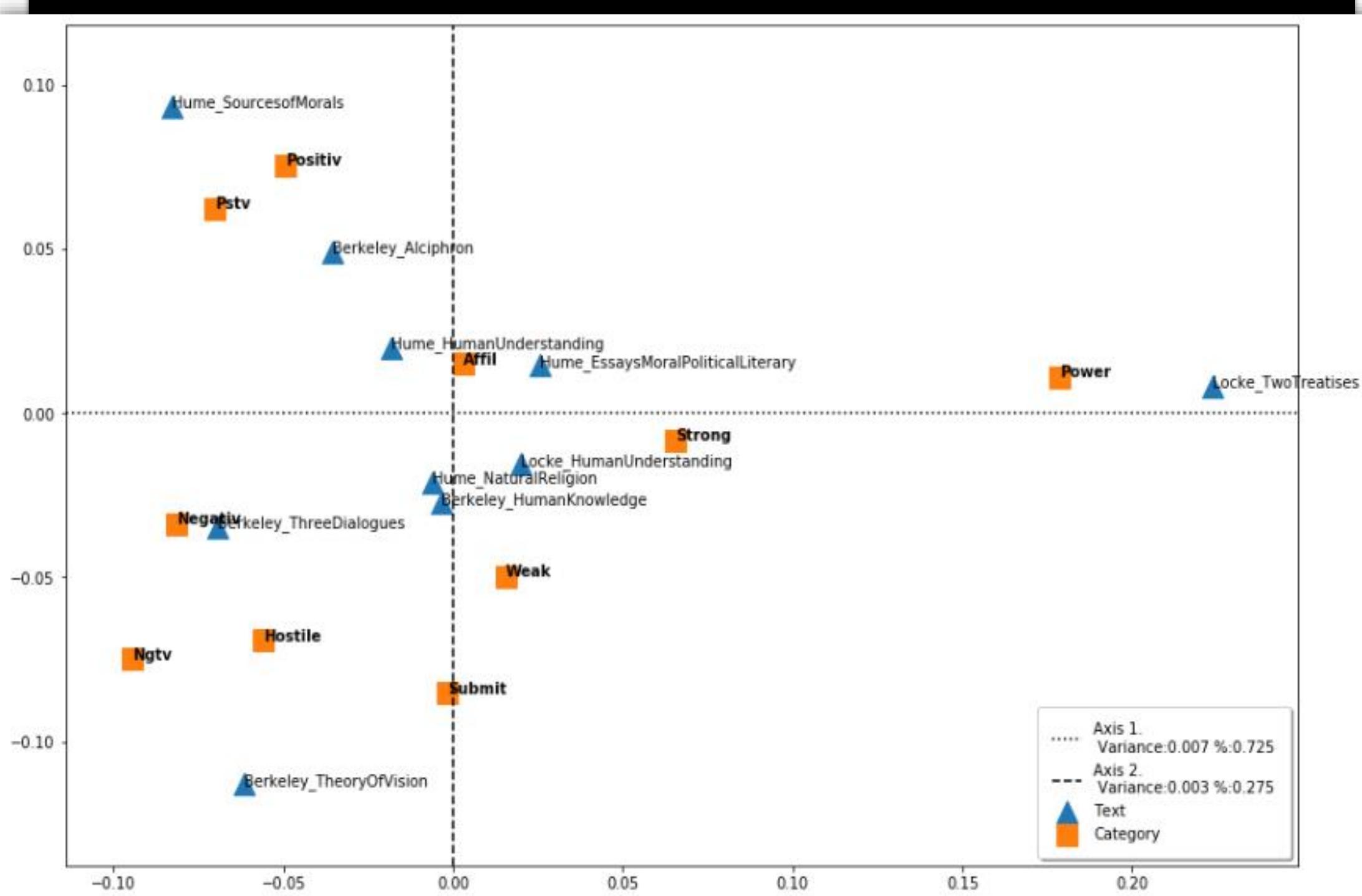
Locke's *Two Treatises of Government* and the **Power** category are removed significantly from the rest of the distribution points. There is a reasonable explanation for this. The document is primarily concerned with governance from God and from earthly governmental bodies, both of which are associated with terms that would fall under the **Power** category. A similar pull towards the **Strong** and **Power** categories can be seen in Hume's political addition to the corpus, *Essays, Moral, Political, and Literary*. However, Hume's document does still tend towards positive categories, while both of Locke's pull towards **Power**. The nature of the text in *Two Treatises of Government* certainly influences the distribution of Locke's document points, but there may be some other underlying function of Locke's lexicon or Philosophy implied by this trend.

Berkeley is Negative

Something about Berkeley's writing places him severely in the negative categories. Why might this be? *An Essay Towards a New Theory of Vision* introduces Berkeley's own scientifically grounded theory of vision, while critiquing widely held beliefs about vision from this time period. In order to situate his own theories, Berkeley must first argue against the pseudoscientific beliefs that preceded him. Berkeley's rhetoric consists of stating commonly held beliefs about vision, denying them, and then introducing his own theory. This critique and dismissiveness fill *Towards a New Theory of Vision* with negative terminology that is picked up by the general inquirer categories. The rhetoric of this text may account for its extreme placement in the negative quadrant of the visualization, but all of Berkeley's documents (with one exception) tend towards the negative. This is indicative of Berkeley's writing style. *In A Treatise Concerning the Principles of Human Knowledge*, Berkeley is concerned about the scepticism that may arise from Locke's conception of empiricism. This led to a critique of his predecessor's methods, much like his critique on theories of vision. In *Three Dialogues between Hylas and Philonous*, Hylas is highly sceptical and critical of Philonous' many arguments about matter and reality. Berkeley was an excellent critic of his predecessors. This style of writing can, at times, be confrontational and requires negation of other's views. Berkeley's documents being placed towards negative categories do not suggest that the philosopher was an unpleasant, pessimistic man. Rather, they reflect his magnificently critical mind and style of writing.

Hume is Positive

On the opposite end of the spectrum, there is Hume's mostly positive collection of documents. *An Enquiry Concerning the Principles of Morals* demarcates the extremes of the positive general inquirer categories, and for good reason. Recalling the earlier discussion on Hume's morality from the topic model, the philosopher's moral theory is, by most accounts, optimistic. Hume rejected the self-interest of utilitarian doctrines, opting instead for a view of morality as a developmental process driven by social conventions and the empathetic nature of human beings. Hume strongly believed that humans tend more towards benevolence than they do to self-love or self-interest. Moral philosophers who emphasized sentiment as the source of human morals, like Hume, express it in a positive light, more so than the cold, calculated logic of utilitarianists who viewed reason as the sole source of morals. This is precisely what the position of Hume's *An Enquiry Concerning the Principles of Morals* in the MCA reflects. The terminology is hopeful, optimistic, and highly positive. Hume's other positively aligned documents both straddle the **Affil** category. To tell why it is that these documents cluster around this category, it would be best to look at a couple examples of the terms contained within it. On one hand, **Affil** terms are things such as "amenable", "intimate", and "kindness". However, these are contrasted by **Affil** terms that are also tagged under the negative categories, like "loneliness", "mourn", and "outcast". The category may be a subset of **Positiv**, but the terms from this category capture both the positive and negative words for affiliation and supportiveness. Rather than viewing **Affil** as a positive category, it should be considered a category defining the intricacies of human relationships, further reinforcing readings of Hume as a modern humanist and moral philosopher.



References

- Capaldi, Nicholas. *Hume's Place in Moral Philosophy*. New York: Peter Lang Publishing Inc., 1989.
 Carlin, Laurence. *The Empiricists: A Guide for the Perplexed*. London: Continuum, 2009.
Harvard General Inquirer Categories. Retrieved from: wjh.harvard.edu/~inquirer/homecat.htm
 Locke, John. *Two Treatises of Government*. Edited by Peter Laslett. New York: Cambridge University Press, 2003.